

**Simulation in Nursing: Historical Analysis and Theoretical Modeling in
Support of a Targeted Clinical Training Intervention**

by

Joseph S. Goode, Jr.

BSN, University of Pittsburgh, 1997

BA, University of Michigan, 2003

MSN, University of Pittsburgh, 2000

Submitted to the Graduate Faculty of
The School of Nursing in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

University of Pittsburgh

2018

UNIVERSITY OF PITTSBURGH
SCHOOL OF NURSING

This thesis/dissertation was presented

by

Joseph S. Goode, Jr.

It was defended on

August 31, 2018

and approved by

Paul Phrampus, MD, FACEP, Professor of Emergency Medicine and Anesthesiology; Director, Winter Institute of Simulation, Education & Research (WISER); Medical Director for Patient Safety, UPMC; Vice Chair of Quality & Patient Safety, Department of Emergency Medicine

Susan M. Sereika, PhD, Professor of Health & Community Systems;
Director, Center for Research and Evaluation

Dissertation Committee Co-Chairs:

Judith Erlen, PhD, RN, FAAN, Professor, Department Chair Health & Community Systems

John M. O'Donnell, RN, CRNA, MSN, DrPH, Professor, Department Chair Nurse Anesthesia

Copyright © by Joseph S. Goode, Jr.

2018

Simulation in Nursing: Historical Analysis and Theoretical Modeling in Support of a Targeted Clinical Training Intervention

Joseph S. Goode, Jr., PhD

University of Pittsburgh, 2018

The use of simulation is widespread in healthcare education, and the potential impact of its use large. This is especially true for nursing education as we look to address problems with obtaining clinical experiences, develop critical thinking skills and create methods to measure the impact of simulation interventions. There is substantial empirical evidence in support of predictive relationships between simulation training interventions and knowledge acquisition. This has been extensively demonstrated with the use of a variety of simulation training modalities from standardized patients to human patient simulators. However, data to support changes in clinical practice and improved patient outcomes are quite limited, including attempts to measure the impact of simulation education on retention and transference of knowledge and skill for more complex healthcare process. Additionally, literature searches reveal that only a handful of authors have engaged in the types of foundational work that any emerging science needs. For example, while pieces of the simulation process have been examined in detail, few have attempted to describe what the *process* of simulation entails at a macro level. Within the past few years some researchers have begun to ask whether there is a causal or predictive relationship present, but few have explored what these associations may look like structurally and what the evidence for them is. The overall objectives of this current research were to: 1) perform an historical review of simulation in healthcare; 2) use this review to outline a new theoretical model of healthcare simulation; and, 3) conduct a small-scale study aimed at pilot-testing and describing part of that model. Hierarchical Task Analysis (HTA) was used to derive an optimum task set for the standard induction of general anesthesia (OTS-SIGA). New Student Registered Nurse Anesthetists (SRNAs) were trained to this task set, and their adherence to the process steps in the clinical setting was then assessed. We also attempted to measure whether repeating the HTA-derived OTS-SIGA simulation training would have an impact on knowledge and transference of simulation-developed skills to the clinical environment. These measures necessitated the development of associated data collection tools and processes for rater training.

Table of Contents

Preface..... 10

1.0 Introduction to the Overall Project..... 11

1.1 Purpose and Specific Aims 11

1.1.1 Historical Review and Analysis 11

1.1.2 Theoretical Model Development..... 12

1.1.3 Healthcare Simulation Pilot Study 12

1.2 Background and Significance..... 14

1.3 Preliminary Work 16

**1.3.1 Using Hierarchical Task Analysis for the Development of a Simulation
Intervention 17**

**1.3.2 Development of Tools for the Measurement of the Impact of Simulation
Interventions on Knowledge Gains and Clinical Outcomes..... 21**

2.0 Historical Analysis 24

3.0 Theoretical Model: Moving Towards a New Model of Healthcare Simulation 35

3.1 Introduction 35

3.2 The Importance of Theoretical Grounding..... 36

3.3 Current Models and Their Role in Simulation Based Research..... 43

3.4 A New Model of Healthcare Simulation..... 46

3.5 Discussion and Support for the New Model..... 53

3.5.1 Falsifiability 54

3.5.2 Curricular Integration..... 54

3.5.3 The central role of debriefing and reflection	55
3.5.4 Social and Clinical Contextualization	57
3.5.5 Explanatory Structure	57
3.5.6 Summary	62
4.0 Pilot Study.....	65
4.1 Methods and Development of the Study Protocol	65
4.1.1 Introduction	65
4.1.2 Background.....	67
4.1.3 Experimental Design, Setting and Sampling	68
4.1.4 Study Timeline and Methodology	71
4.1.4.1 Standard Preparation and Familiarization with the OTS-SIGA Protocol.....	73
4.1.4.2 Baseline Data Collection.....	74
4.1.4.3 Baseline Mock Induction Event.....	74
4.1.4.4 Clinical Observation Period.....	75
4.1.4.5 Refresher Simulation Event	78
4.1.5 Measures	79
4.1.5.1 Demographic data.....	79
4.1.5.2 Knowledge assessment.....	79
4.1.5.3 Adherence to the OTS-SIGA Process steps.....	80
4.1.5.4 Self-report of confidence in the ability to complete the OTS-SIGA .	82
4.1.5.5 Agreement of SRNA airway assessment with expert airway assessment	82

4.1.5.6 Self-report and clinical preceptor report of the ability to secure the airway (successful placement of an endotracheal tube or a supraglottic airway device)	83
4.1.5.7 Self-report of the incidence of oral soft tissue or dental injuries	83
4.1.6 Observational Rater Training.....	84
4.1.7 Statistical Analysis Methods	85
4.1.8 Summary	86
4.2 Results of the Pilot Study	87
4.2.1 Inter-rater Reliability	87
4.2.2 Demographic Characteristics of the Sample	89
4.2.3 Standard Induction of General Anesthesia-Knowledge Assessment Instrument	91
4.2.4 OTS-SIGA Live Observations	93
4.2.5 SIGA-CAA Preceptor Scoring.....	97
4.2.6 SIGA-CAA Preceptor Overall Rating and SRNA Self-Evaluation.....	110
4.2.7 Preceptor to SRNA Airway Assessment Comparisons.....	112
4.3 Discussion	114
4.3.1 Rater Training.....	115
4.3.2 Demographic Profile of the Study Group	116
4.3.3 Knowledge Retention: SIGA-KAI.....	117
4.3.4 OTS-SIGA Live Scoring and Clinical Preceptor Scoring	118
4.3.5 Clinical Preceptor Overall Rating and SRNA Rating of Self-Confidence.	121
4.3.6 Agreement with Airway Assessment	122

4.3.7 Oral Soft Tissue and Dental Injury	123
4.4 Limitations	124
4.4.1 Methodological Limitations	124
4.4.2 Pilot Study Data Interpretation Limitations	126
5.0 Conclusions and Summary.....	129
5.1 Conclusions	129
5.2 Summary	133
Appendix A Hierarchical Task Analysis Derived Optimal Task Set for Standard Induction of General Anesthesia (OTS-SIGA) © 2006 Goode, Jr. JS, O'Donnell JM; revised January 2016	134
Appendix B Example of the Standard Induction of General Anesthesia Knowledge Assessment Instrument (SIGA-KAI) © 2001 Goode, Jr. JS, O'Donnell JM.....	137
Appendix C Operational Definitions.....	142
Bibliography	155

List of Tables

Table 1 Cohen’s kappa for each rater, comparing their ten subgoal operational task scores for a post-training test video with those of a benchmark score. Strength of agreement assessment from Landis and Koch (1977).....	88
Table 2 Cohen’s kappa for each rater, comparing their fifty subgoal scores for a post-training test video with those of a benchmark score. Strength of agreement assessment from Landis and Koch (1977).....	88
Table 3 Demographic Characteristics for the total sample and by group.	91
Table 4 SIGA-KAI Score Means and Standard Deviations.....	93
Table 5 Linear Mixed Modeling Results for OTS-SIGA Live Scoring: Proportion of Subgoal Steps Scored Yes, No, Prompted or Not Available	95
Table 6 OTS-SIGA Live Scoring Effect of Time: Proportion of all Subgoal Scores by Time Block for the Entire Cohort	96
Table 7 Linear Mixed Modeling Results for Preceptor Scoring of Equipment Verification: Proportion of Subgoal Steps Scored Yes, No or Not Available	99
Table 8 Linear Mixed Modeling Results for Preceptor Scoring of Apply Monitors: Proportion of Subgoal Steps Scored Yes, No or Not Available.....	99
Table 9 Linear Mixed Modeling Results for Preceptor Scoring of Induction Steps: Proportion of Subgoal Steps Scored Yes, No or Not Available.....	100
Table 10 Linear Mixed Modeling Results for Preceptor Scoring of Mask Ventilation: Proportion of Subgoal Steps Scored Yes, No or Not Available	100

Table 11 Linear Mixed Modeling Results for Preceptor Scoring of Laryngoscopy & Intubation:
Proportion of Subgoal Steps Scored Yes, No or Not Available 101

Table 12 Linear Mixed Modeling Results for Preceptor Overall Performance Rating for
Performing OTS-SIGA Steps and for SRNA Self-Rating of Confidence in Performing OTS-SIGA
Steps..... 110

List of Figures

Figure 1. Screen Shot of a Portion of the OTS-SIGA Scoring Application.	23
Figure 2. Simple Continuum From Simulation to Change in Clinical Practice Pattern.	40
Figure 3. Proposed Model of the Impact of Healthcare Simulation Interventions on Clinical Practice.....	48
Figure 4. Causal Chain Involved in Mediation as Proposed by Baron and Kenny (1986).....	49
Figure 5. A priori power analysis for two groups (t-tests for two independent means) assuming equal sample sizes.	70
Figure 6. Power analysis for n = 12 for each arm of the study indicated a 60% chance of detecting a large effect size (defined by Cohen, 1992, as .95 of a population standard deviation between the means) with significance set at the 0.5 level(two-tailed).....	71
Figure 7. Study Timeline	72
Figure 8. SIGA Clinical Assessment Application Tool (SIGA-CAA) Preceptor Portal View.	77
Figure 9. SIGA Clinical Assessment Application Tool (SIGA-CAA) SRNA Portal View.	78
Figure 10 SIGA-KAI Scores Over Time	92
Figure 11 OTS-SIGA Live Scoring Change in Proportion of all Subgoal Scores by Time Block for the entire cohort (mean ± SE)	97
Figure 14 Change in Preceptor Scores Over Time: ‘Equipment Verification’ for the Total Study Group (mean ± SE)	101
Figure 15 Change in Preceptor Scores Over Time: ‘Apply Monitors’ for the Total Study Group (mean ± SE)	102

Figure 16 Change in Preceptor Scores Over Time: ‘Induction Steps’ for the Total Study Group (mean ± SE)	102
Figure 17 Change in Preceptor Scores Over Time: ‘Mask Ventilation’ for the Total Study Group (mean ± SE)	103
Figure 18 Change in Preceptor Scores Over Time: ‘Laryngoscopy & Intubation’ for the Total Study Group (mean ± SE).....	103
Figure 19 Change in Preceptor Scores: ‘Equipment Verification’ SIGA-O and SIGA+R groups (mean ± SE)	105
Figure 20 Change in Preceptor Scores: ‘Apply Monitors’ SIGA-O and SIGA+R groups (mean ± SE).....	106
Figure 21 Change in Preceptor Scores: ‘Induction Steps’ SIGA-O and SIGA+R groups (mean ± SE).....	107
Figure 22 Change in Preceptor Scores: ‘Mask Ventilation’ SIGA-O and SIGA+R groups (mean ± SE).....	108
Figure 23 Change in Preceptor Scores: ‘Laryngoscopy & Intubation’ SIGA-O and SIGA+R groups (mean ± SE)	109
Figure 24 Change Over Time in Preceptor Overall Rating of SRNA Ability to Perform OTS-SIGA Steps Compared to SRNA Self-Rate Confidence in Performing OTS-SIGA Steps for the SIGA-O and SIGA+R Groups and the Total Study Group.	111
Figure 25 Airway Assessment Agreement over time for the Total Cohort.	113
Figure 26 Airway Assessment Agreement Over Time for the SIGA-O Cohort.....	113
Figure 27 Airway Assessment Agreement Over Time for the SIGA+R Cohort.	114

Preface

The study was supported through funding from the Margaret E. Wilkes Scholarship Fund Award. I would like to acknowledge and thank my dissertation co-chairs, Judith Erlen & John O'Donnell for their constant support, encouragement and occasional reminders that someday I would *actually have to complete my dissertation*. Thanks as well to my other committee members, Paul Phrampus & Susan Sereika, for their contributions and guidance. Robert Batterman provided key insights into minimal models and Peter Machamer helped guide the historical analysis and my understanding of pragmatist philosophy. Many thanks to the members of the Nurse Anesthesia Program Class of Fall 2018 who agreed to participate in the study and to the student researchers who assisted in data collection and video analysis, especially Eman Bascal, Blake Jule, Sarah Kukura, Michelle Smith and Christina Yanicko. Jade Hasson played a particularly important role during the preliminary work phase and during the pilot study Mock Induction training sessions. John Lutz created the web-based tools based on our HTA process map that allowed us to easily rate study participant performance in the simulation and clinical settings.

The doctoral education process is challenging under any circumstances. Doing so while continuing my duties as an instructor with the University of Pittsburgh Nurse Anesthesia Program and as a clinical CRNA at UPMC-Presbyterian Hospital proved to be the most difficult task I have undertaken. Drs. O'Donnell and Erlen have been long-time mentors, collaborators and dear friends. Their contributions go far beyond serving as my committee chairs. Helen DeFranco and Dale Fleck, the Chief and Associate Chief CRNAs at UPMC-Presbyterian, were understanding and flexible in scheduling my clinical duties so that I might complete the pilot study that is the basis of this work. I'm grateful for my all of my closest friends who patiently listened to my ideas, concerns and frustrations, particularly Paul Cartman, Bob Evani, Timothy Schumann, Renée Ranier and Lauren Wohl Sanchez. Finally, thanks to my family members: my mother Gloria Goode-Barone, my sister Cynthia Surace-Volpe and my niece Sabrina Volpe. Their patience and support during this process has been limitless as is my appreciation.

1.0 Introduction to the Overall Project

1.1 Purpose and Specific Aims

There are three broad areas where gaps still exist in our understanding of the role of healthcare simulation: 1) the historical, philosophical and ethical grounding of simulation; 2) the theoretical constructs guiding the use of simulation methodologies; and 3) how learning transfers from the simulation setting to the clinical environment. This three-part program of research attempts to address, in part, these areas. Part one was an examination of the history of simulation in healthcare. Part two entailed leveraging the historical review in combination with the best evidence in the literature to develop a new theoretical model of healthcare simulation. Part three was the conduction of a pilot study attempting to measure the transference of skills from the simulation setting to the clinical environment, and also targeting components of the new model of healthcare simulation model.

1.1.1 Historical Review and Analysis

Area of Inquiry I: Do historical precedents exist that may help to inform our understanding of current approaches to healthcare simulation?

Specific Aim 1

An examination of the historical record for precedents to our current approaches to healthcare simulation –with a focus on its use in nursing education was undertaken. This aim of this review was to help define the boundaries of the science of simulation and provide perspective

and guidance for the development of a new model of healthcare simulation. Additionally, a preliminary exploration of past and emerging ethical and philosophical considerations was done.

1.1.2 Theoretical Model Development

Area of Inquiry II: Does the emerging science of healthcare simulation have any agreed upon theoretical model(s) to guide practice or further investigation?

Specific Aim 1

To describe currently used theoretical models and to determine how they have been used to guide both simulation interventions and investigations into issues such as transference and the timing of simulation interventions.

Specific Aim 2

To develop a new theoretical model that will seek to correct existing theoretical weaknesses and place components of the simulation process into a descriptive temporal matrix. The results obtained from Specific Aim 1 above, as well as the findings of the historical review, were used to inform this process.

1.1.3 Healthcare Simulation Pilot Study

Area of Inquiry III: Can we demonstrate and describe the process of transference of learning from the simulation setting to clinical practice?

The specific aim of the pilot study, *Adherence and Retention of an HTA Optimal Task Set for Standard Induction of General Anesthesia (OTS-SIGA): Measuring Transference of Skills Learned in a Simulation-Based Educational Intervention for First Term Student Registered Nurse*

Anesthetists was to assess the impact of a simulation intervention, based on a Hierarchical Task Analysis (HTA) derived protocol for standard induction of general anesthesia (OTS-SIGA), on clinical outcome performance markers of new Student Registered Nurse Anesthetists (SRNAs). This included determining whether repeating the HTA-derived OTS-SIGA simulation training would have an impact on transference of simulation-developed skills to the clinical environment. Specific research questions included:

Research Question 1: Will the use of an HTA derived standardized training protocol result in good adherence to the general anesthetic induction process steps for these SRNAs?

Research Question 2: Will repeating the HTA-derived induction process training (refresher training) improve adherence to the HTA-derived induction protocol process steps?

Research Question 3: Will repeating the HTA-derived induction process training (refresher training) improve clinical outcome performance markers of SRNAs doing standard general anesthetic inductions? These markers included:

1. ability to perform preoperative airway assessments
2. ability to secure the airway (successful placement of either an endotracheal tube or a supraglottic airway device)
3. competence in completing key steps of the OTS-SIGA
4. confidence in their ability to complete the OTS-SIGA
5. incidence of oral soft tissue or dental injuries

1.2 Background and Significance

The use of simulation is widespread in healthcare education, and the potential impact of its use large. This is especially true for nursing education as we look to address problems with obtaining clinical experiences, develop critical thinking skills and create methods to measure the impact of simulation interventions. There is substantial empirical evidence in support of a predictive relationship between simulation training interventions and knowledge acquisition. This has been extensively demonstrated with the use of a variety of simulation training modalities from standardized patients to human patient simulators. (Cook et al., 2011; Crofts et al., 2007; Hoffmann, O'Donnell, & Kim, 2007; Seibert, Guthrie, & Adamo, 2004) Over the last decade evidence has begun to emerge that supports a long-held contention that this educational modality has the potential to change how providers practice with resultant improvement in patient outcomes. Among the first to document improved clinical outcomes was DeVita et al. in a retrospective analysis which reported a reduction in code-related mortality after implementing highly structured simulation team training for an in-hospital Medical Emergency Team (MET). (DeVita & Minnini, 2004) Since then other authors have pointed to areas such as simulation team-training and highly specific task training, which are generating promising and compelling evidence for a positive impact on both educational and patient outcomes. (Barsuk et al., 2015; Barsuk, McGaghie, Cohen, O'Leary, & Wayne, 2009; Crofts et al., 2008; Draycott et al., 2008) One powerful example is the simulation-based mastery learning (SBML) approach to teaching central venous catheter (CVC) insertion skills. Barsuk et al. in a series of studies regarding the use of mastery learning techniques in central venous catheter insertion training have demonstrated retention and transference of simulation acquired skills, improved patient outcomes and healthcare system cost savings. (Barsuk, Cohen, McGaghie, & Wayne, 2010; Barsuk, McGaghie, Cohen, Balachandran, & Wayne,

2009; Cohen et al., 2010) They reported immediate post-intervention internal jugular CVC insertion skill retention of 100% with impressive 6- and 12-months post-training skill retention rates (82.4% and 87.1% respectively). (Barsuk et al., 2010) What is still missing from the literature is an attempt to directly measure the impact of simulation education on retention and transference of knowledge and skill for a more complex healthcare process (e.g. induction of anesthesia).

Perhaps the best overview of the state of healthcare simulation as a science has come from a series of meta analyses and systematic reviews undertaken by Cook et al. These were aimed at examining the educational efficacy of simulation interventions as well as their impact on patient outcomes. (Cook et al., 2012; Cook, Brydges, Zendejas, Hamstra, & Hatala, 2013a, 2013b; Cook, Erwin, & Triola, 2010; Cook, Hamstra, et al., 2013; Cook et al., 2011; Cook, Levinson, et al., 2010) This work provided meaningful insight into the current state of simulation science, most notably that the measurable impact of simulation education was large enough for the authors to recommend that there was little value in continuing to do studies comparing simulation with no simulation. However, most of the studies available for review had serious methodological limitations, clearly identifying significant theoretical and structural weaknesses for *the science* of simulation. As Cook (2011, pp. 987-988) stated, “The important questions for this field are those that clarify when and how to use simulation most effectively and cost-efficiently. Unfortunately, the evidence synthesized herein largely *fails to inform the design of future simulation activities*. [emphasis added]”

These are critical, core questions, the answers to which will help to determine the direction of future research in healthcare simulation. To be sure, there has been an explosion of publication in the field over the last fifteen years. (Issenberg, McGaghie, Petrusa, Lee Gordon, & Scalese, 2005) But structured literature searches have revealed that only a handful of authors have engaged

in the types of foundational work that any emerging science needs.¹ For example, while *pieces* of the simulation process have been examined in detail, few have attempted to describe what the *process* of simulation entails at a macro level. Additionally, while it is tacitly implied, no one has explicitly asked whether there are causal or predictive relationships present, what these associations may look like structurally and what the evidence for them is.

Given this background, the overall objectives of this current research were to: 1) perform an historical review of simulation in healthcare; 2) use this review to outline a new theoretical model of healthcare simulation; and, 3) conduct a pilot study aimed at testing and describing part of that model.

1.3 Preliminary Work

Preliminary work that supported and informed this proposal has occurred in the areas of 1) leveraging hierarchical task analysis (HTA) for the development of targeted simulation interventions; 2) measurement of the impact of these simulation interventions on knowledge gains and clinical outcomes; and 3) the qualitative analysis of trainee perceptions of the impact of their simulation experiences on their later clinical practice.

¹ An example of a structured search for articles relating to models of the simulation process: (("Models, Nursing"[Mesh] OR "Models, Educational"[Mesh]) OR "Models, Psychological"[Mesh]) AND (("Patient Simulation"[Mesh] OR "Computer Simulation"[Mesh:noexp]) OR "Manikins"[Mesh]). This search was also run separately looking for these terms as components of the title or abstract, and run again in combined form with the mesh terms. These efforts yielded potential 1010 articles of which after review only 23 related to either proposing or critiquing an existing model of the healthcare simulation process.

1.3.1 Using Hierarchical Task Analysis for the Development of a Simulation Intervention

The principle investigator and his colleague John O'Donnell have demonstrated over a series of studies the utility of using HTA to deconstruct a process into its component steps and then using those process steps to both design and measure outcomes of simulation educational interventions. (Goode, Schumann, Klain, & O'Donnell, 2007; O' Donnell J, Goode Jr, Odonohoe, & Choe, 2007; O' Donnell, Goode, et al., 2012) The first attempted use of HTA was in the description of the process of inserting an intravenous catheter. (O' Donnell J et al., 2007) In a later study, O'Donnell et al. reported this approach in the implementation of a team-based training model intended to reduce the incidence of musculoskeletal injury associated with the physical transfer or repositioning of patients. (O' Donnell, Goode, et al., 2012; O' Donnell et al., 2011) HTA methodology was used to deconstruct the process of transferring a patient, allowing for the development of an idealized 10-Step Patient Transfer Protocol. Two physically separate hospitals, each with similar neurologic rehabilitation units, and no movement of personnel between them, were utilized for the control and intervention cohorts. All personnel on the intervention units were trained with the simulation intervention. The outcome measure was adherence to the idealized 10-Step Patient Transfer Protocol with the protocol itself becoming the objective scoring tool by which trained observers rated personnel on both the control and intervention units as they carried out actual patient transfers. (O' Donnell, Goode, et al., 2012; O' Donnell et al., 2011)

The target of this current project was a more complex process, the standard or routine induction of general anesthesia. The University of Pittsburgh Nurse Anesthesia Program has long recognized the need to appropriately train and prepare SRNAs for entry into their first clinical rotation. For more than 15 years a Mock Induction simulation training educational exercise has been conducted in some form and is a core component of this preparation process. The goal of Mock Induction is

to prepare first term SRNAs to perform the process of doing an induction of general anesthesia. In the past, as with most nurse anesthesia and residency training programs, this has been done without a specified process task list. The description of an optimal task set for a standard induction of general anesthesia (OTS-SIGA) was deemed essential to allow for measurement of SRNA performance in the simulation setting. Additionally, development of this task set would allow for measurement of the same clearly defined tasks in the clinical setting. Standard induction of general anesthesia (SIGA) is a complex process involving a minimum of fifty-one major steps, with many of those having component steps of their own. Well over one hundred combined total steps are involved. To capture this complexity, the description of the SIGA process followed a nine-step HTA protocol. These steps (and the specific processes that were followed) are derived from the work of Annett, Shephard and Stanton and are described below. (Annett et al., 2000; Shepherd, 1998; Stanton, 2006)

1. *Define the purpose of the analysis*

The purpose was to describe the component steps of a standard induction of general anesthesia, which were then used as a template for the structure of a simulation education intervention to teach this process. The goal is to be able to teach this process in a simulated environment in a manner that replicates actual clinical processes as much as is possible. The component steps will also form the basis of an evaluative tool for use in measuring adherence to the protocol steps in both the simulated and real-world settings.

2. *Define the boundaries of the system description.*

All procedures and tasks that an anesthesia provider would need to perform to safely complete a standard induction of general anesthesia.

3. *Access a variety of information sources about the system to confirm the reliability and validity of the analysis.*

A wide range of sources were accessed, including but not limited to: extensive review of the anesthesia literature, focused interviews with clinical experts (certified registered nurse anesthetists (CRNAs), anesthesiologists, nurse anesthesia faculty), review of relevant guidelines from anesthesia professional organizations and clinical observations of the process. Additionally, the best evidence from the simulation education literature as well as input from experts in healthcare simulation education was used to critique the optimal task set for OTS-SIGA.

4. *Describe the system goals and subgoals; define a subgoal hierarchy for the task at hand.*

The system (or superordinate) goal is “perform a standard induction of general anesthesia following accepted guidelines of anesthetic practice and patient safety”. An iterative process of consulting information sources as outlined in Step 3 above, defining and ordering process steps (subgoals) and then subjecting the proposed task set to subject matter expert review was followed. At the end of each iterative segment, clinical validation was attempted through observations of both simulated and actual anesthetic inductions.

5. *Try to keep the number of immediate subgoals under any superordinate goal to between 3 and 10.*

The HTA literature suggests that it is optimal to keep the number of subgoals to a maximum of 10. Keeping this limitation in mind, we ultimately decided on the number of process steps by consultation with subject matter experts.

6. *Link goals to subgoals and describe the conditions under which subgoals are triggered.*

Each step of the protocol was operationally defined, again referring to the reference sources in Step 3. Each operational definition clearly defines the beginning of each system subgoal (task step), triggering actions to begin the next step or, for the final subgoal, the indicators that the system goals had been completed.

7. *Stop redescribing the subgoals when you judge the analysis is fit for the purpose.*

By convention in HTA, the re-description process stops when the level of description is deemed appropriate to achieving the system goals, in this case safely completing a standard induction of general anesthesia. Ultimately termination of the re-description process was determined by consultation with content experts and validation through clinical observation.

8. *Verify the analysis with subject-matter experts.*

Subject matter experts were asked to perform a final review of the protocol to verify that the analysis is fit for the purpose and that operational definitions of subgoals are clear and in concordance with the best evidence in the literature.

9. *If necessary, revise the analysis based on feedback.*

Per HTA process guidelines, revisions were required based on results of Step 8.

The current Mock Induction simulation is structured around the final iteration of the OTS-SIGA (**Appendix A**).

In our experience, HTA has proven to be a viable, valuable approach to evaluating the impact of simulation training, being that it is scalable from the level of the individual to that of whole, complex systems. Outside of simulation, Annett has used HTA to describe the complex task management sets encountered in the military nuclear submarine setting. (Annett et al., 2000; Shepherd, 1998) This provides confidence that the process of SIGA can be adequately described and measured.

1.3.2 Development of Tools for the Measurement of the Impact of Simulation Interventions on Knowledge Gains and Clinical Outcomes

As described above, O'Donnell's prospective, longitudinal observational study of the effect of simulation on patient transfer in the clinical setting demonstrated improved scores on a knowledge assessment tool (baseline score of 65% correct to 95% correct post-simulation intervention). Patient transfer success in the clinical setting, as measured by adherence to the idealize protocol, also improved (66% adherence pre-simulation, 88% at 4-weeks post-simulation and 71% at 12-weeks post-simulation). (O' Donnell, Goode, et al., 2012; O' Donnell et al., 2011) This study provided valuable experience which has been leveraged in the development of this current work. Additionally, an exploratory and developmental study with a cohort of 22 first term SRNAs (HRPO Protocol #12090179, Adherence and Retention of an HTA Optimal Task Set for Standard Induction of General Anesthesia (OTS-SIGA): A Simulation-Based Educational Intervention for First Term Student Registered Nurse Anesthetists) allowed for the development of data collection tools. This included the development and trial of a knowledge assessment tool,

the Standard induction of General Anesthesia Knowledge Assessment Instrument (SIGA-KAI) (**Appendix B**).

Additionally, an OTS-SIGA scoring tool was derived from the HTA process described above with the assistance of the Winter Institute for Simulation, Education and Research (WISER) information technology personnel. This tool was deployed as a web-based application accessible from any tablet or smartphone platform (see **Figure 1**). The application was trialed during the Mock Induction simulation training sessions with this exploratory study cohort and evaluated for ease of use and reliability. Each step of the induction process was scored ‘live’. Possible scores for each step include ‘completed’, ‘completed after prompting’, ‘not completed’. Review of our initial use of the OTS-SIGA application in this simulation setting resulted in some modifications. A key decision was to add a ‘no opportunity/not applicable’ scoring option in recognition of differences between the simulation event and the real-world induction experiences SRNAs would participate in. The OTS-SIGA attempted to include all possible steps and processes for a standard induction. During the simulated Mock Induction sessions, the standard of training is to have each participant complete every step, regardless of how much prompting is required. However, in the clinical setting it is possible that the clinical preceptors may not allow the SRNA to do a process step, may complete it themselves or that step may not always be applicable for a given patient.

Event Log App - NURSAN 2700

Participant :
Reviewer :
Media : Time Frame :
Scenario :

For each of the following categories, select one of the following:

Y = Yes they completed this task independently

P = Completed this task with prompting

N = Did not complete

1. Equipment verification- Verbalize or demonstrate

1.1. System pressure check Y P N

1.2. Prepare machine (flows, open APL valve) Y P N

1.3. Place suction, oral airway, tongue blade, laryngoscope, ETT, ETT tree at bedside Y P N

1.4. Manual Resuscitator, manual cuff, independent light source, alarm system Y P N

2. Interview and Airway Assessment- Ask pertinent questions (keep short-get facts)- Demonstrate

2.1. Patient ID Y P N

2.2. Oral Opening Y P N

2.3. Mallampati Classification Y P N

2.4. Thyromental distance Y P N

2.5. Cervical range of motion Y P N

2.6. Mandibular length Y P N

2.7. Mandibular mobility Y P N

2.8. Identification of cricothyroid membrane Y P N

2.9. Other exam and questions appropriate to case Y P N

3. Apply Monitors- Demonstrate

3.1. Check that safety strap is in place on patient Y P N

Figure 1. Screen Shot of a Portion of the OTS-SIGA Scoring Application.

2.0 Historical Analysis

Most contemporary histories of simulation in healthcare begin with the fortunate meeting and subsequent collaboration of Peter Safar and Asmund Laerdal. (Cooper & Taqueti, 2008) However, a strong case can be made that attempts to leverage a variety of types of technology to assist in training and teaching about anatomy, medicine and the care of patients go back much further. In the 17th century Jacques de Vaucanson (1709-1782) created several automata that were widely known and discussed. If his own account to the French Royal Academy of Sciences is to be believed, de Vaucanson attempted to replicate not just the *appearance* of human –and animal- function, but the actual coordinated *movements* themselves. His description of one of the two automaton flute players he created details his intent to replicate in as much detail as possible the human mechanical functions associated with this task, as well as how he believes he has improved upon the process:

*“In this the Automaton surpasses all of our players of the tambourin, who cannot move the tongue with enough nimbleness, to make a whole measure of 16th notes all articulated. They flow at half and my Tambourin plays an entire tune with a flick of the tongue at each note”*² (Vaucanson, 1738)

While flute players are not themselves of interest to our discussion of healthcare education, several sources indicate that de Vaucanson had the idea of, and perhaps actually embarked upon,

² While Vaucanson refers to this automaton as the ‘Tambourin’ it was actually a complex machine in the guise of a rustic shepherd and was reported to have been able to play 20 different songs using the galoubet wind instrument with one hand and a Turkish-style tambourine with the other.

constructing an automaton that would replicate certain biologic functions such as the circulation and perhaps the function of the lungs. ("Littell's Living Age," 1852; Wood, 2002) His musical playing creations may have been a precursor to this work, but the project does not appear to have ever been completed. Further investigation is planned which may provide additional information.

There were of course other attempts to teach through 'simulation', especially in the area of anatomy. Wax models reproduced anatomic structures in exquisite detail, and some clearly were designed to be touched and taken apart, allowing the learner to see how the anatomic pieces fit together. (Chen, Amar, Levy, & Apuzzo, 1999; Dacome, 2007; Messbarger, 2001) However, it wouldn't be until the 20th century that any Vaucanson-like attempts to reproduce human anatomy and physiology would be made, especially in terms of the mannequin-based simulation training that has become the predominant model of healthcare simulation outside of the realm of surgical simulation.

Although often overlooked in historical overviews, one of the earliest examples of mannequin-based training occurred in the nursing discipline in the form of 'Mrs. Chase'. Introduced in 1910, this mannequin was designed with the intent that it "would make teaching demonstrations easier and would afford students an opportunity to practice, thereby sparing patients possible discomfort." (Herrmann, 1981) This is a prescient echo of the current National Patient Safety Movement call for the use of simulation in healthcare education. The conceptual mother of Mrs. Chase, Lauder Sutherland, contracted with a doll manufacturer (the M.J. Chase Company) to produce the mannequin. Fifty years later Peter Safar –considered the father of Cardiopulmonary Resuscitation (CPR)- would collaborate with another doll manufacturer, Asmund Laerdal, to produce the widely known CPR trainer Resuci®-Anne. (Cooper & Taqueti, 2004b; Herrmann, 1981) The emergence of more readily accessible computer technology, at least

at some large academic centers, allowed for the development of mathematical models of human physiology and the pharmacologic effects of drugs on that physiology. An example of this was the screen-based simulator SLEEPER developed by Smith at the University of California, San Diego. (Cooper & Taqueti, 2004b; Lawson, 1990) During this same period in the 1980s, David Gabba at Stanford and Michael Good/JS Gravenstein at the University of Florida were developing full-body mannequins that incorporated some features of a modeled physiology. (Cooper & Taqueti, 2004b; Gaba & DeAnda, 1988) The costs of these one-of-a-kind systems hindered wide-spread dissemination, and although both concepts eventually were licensed to industry, there simply wasn't a viable market until the introduction of Laerdal's more affordable SimMan®.

In reviewing the simulation literature that has accumulated over the past eighteen years, one can sometimes experience a sense of déjà vu. For example, if we look at typical list of the advantages of simulation as a teaching modality, we find the following:

- Strong preference by trainees
- Allows for experiential learning in a safe environment
- Can convey information in multiple learning domains at one time
- Avoids the problems of not being sure healthcare trainees will ever see a particular event or type of case
- Allows for reflection after the event. (JM O' Donnell, Beach, & Hoffman, 2005; JM O' Donnell, Fletcher, Dixon, & Palmer, 1998; JM O' Donnell & Goode, 2008)

These are essentially the same advantages that were being made about education through gaming (both computer and non-computer) in the 1960s and 1970s. (Clark, 1976, 1977; Coleman, Livingston, Fennessey, Edwards, & Kidder, 1973; Greenblat, 1971, 1977) Other key concepts that are found in the current healthcare simulation literature also abound in the older gaming discussion.

Coleman, et al discussed the issue of “translation” from a symbolic framework to action, the importance of post-game reflection and discussion (read that as ‘debriefing’ in the current simulation literature) and the fact that games that involve teams or groups of people can “change students’ attitudes towards the real-life persons whose roles they take”. (Coleman et al., 1973) Greenblat lamented the lack of good guidelines for conducting post-game debriefings or for integrating this educational modality into curricula. (Greenblat, 1971, 1977) Clark discussed the possibility that gaming education might be valuable for trainees to go through just prior to beginning clinical rotations with real patients. (Clark, 1977) Every essential issue regarding healthcare simulation today was presaged by the computer gaming movement in the 60s and 70s, but current mannequin-based simulation researchers appear to be unaware of this body of literature. These core references are never cited, resulting in a literal, and wasteful, example of ontogeny recapitulating phylogeny in healthcare education. (JM O’ Donnell & Goode, 2008)

One of the most important issues, if not the most important, is that of translation to clinical practice. That is, does the investment of time and resources in simulation actually change how healthcare providers practice and, if it does, are there measurable benefits in terms of patient outcomes? Not surprisingly, both Greenblat and Clark frequently mention this in the earlier gaming literature:

In addition, the field has moved from the early days of great enthusiasm without hard data to realization of the importance of testing claims about the efficacy of games. Numerous claims have been made about the effectiveness of gaming-simulations in teaching, including increases in (a) motivation and interest, (b) cognitive learning, (c) skills of various sorts, (d) affective learning (e.g. attitude change, empathy), and (e) creation of a more effective learning environment; but few methodologically sound attempts have been made to test these claims. In recent years, more

systematic attempts have been undertaken and further data have been accumulated. While questions of whether and how games work remain to be answered, there is increasing support for the claims... (Greenblat, 1977)

Transfer of simulation game learning to clinical practice should be studied. It can be speculated that since instructional objectives were developed through the Goal Analysis Method and tied to game plays, transfer to clinical practice could be high. Whether, in fact, students would perform one way in the simulation game and another with patients is yet to be studied. (Clark, 1977)

And just like the gaming movement proponents, contemporary simulation researchers have been slow to attempt to address this central issue. This is no small academic quibble. From the start of this current era of healthcare simulation in the late 1990s, costs for implementing a program of simulation training are substantial in terms of both money and faculty time. In a 1998 article, O'Donnell, et al estimated \$200,000 - \$300,000 for initial start-up of a simulation center and 345 hours of actual faculty time commitment for development and implementation of an ACRM course. (JM O' Donnell et al., 1998) In the era of evidenced based practice, it would seem to be important to confirm the utilities and the areas of strength and weakness of an educational intervention before committing large quantities of academic resources.

There is a vigorous and ongoing debate in healthcare simulation circles regarding the level of fidelity (or realism) necessary for any given simulation exercise. This includes efforts to use immersive virtual reality, at times with some level of predictive or artificial intelligence. Whether necessary or not, there is little doubt that the future of simulation is linked to both past and current efforts to reproduce life-like functionality. The recent history of artificial intelligence research reveals interesting parallels. Timothy Lenoir and editor Jessica Riskin, in *Genesis Redux* describe

how some AI pioneers, such as Rodney Brooks, view creating intelligent systems as a process of simply building and layering structural complexity. (Riskin, 2007) In this way artificial intelligence is brought about not by design, but by emergence. In a likewise manner the healthcare simulation industry continues to develop increasingly complex and sophisticated platforms for training, whether they be mannequin or virtual reality based. This raises important questions about how healthcare providers interact with such systems and devices, questions that are not new. Automata such as those of de Vaucanson, Pierre Jaquet-Droz (1721-1990) and others stimulated much philosophical debate about what it means to be living or intelligent (e.g. is simulated breathing a sign of life). (Wood, 2002) As previously discussed, in the 1960's computer gaming researchers such as Clark asked related questions, wondering how interacting in a computer generated milieu would affect a student's later interactions with real patients. (Clark, 1977) To date these questions remain unanswered, with little discourse on the subject in the current literature. There are also questions emerging about the need to address environmental and psychological fidelity in simulation. (Alinier, 2011)

Fidelity is also a potential concern with regard to our assessment of simulation outcomes. The issue of what, if anything, trainees and clinicians are taking with them to their practice when they leave a simulation training session is unclear. The one area that appears to have good support for transference, team-training simulation, has at the same time, generated conflicting evidence with regard to efficacy and outcomes. Chan et al in two publications has presented data that seem to indicate that the team-training model is not effective. Introduction of a Rapid Response Team model at a large academic medical center, Saint Luke's Hospital in Kansas City, Missouri, "was not associated with a reduction in the primary end point of hospital-wide code rates ... although lower rates of non-ICU codes were observed ... Similarly, hospital-wide mortality did

not differ between the preintervention and postintervention periods.” (Chan et al., 2008) This was followed up with a meta-analysis of the available literature on RRTs from 1950 until 2008. Chan’s conclusion was that “Although RRTs have broad appeal, robust evidence to support their effectiveness in reducing hospital mortality is lacking.” (Chan, Jain, Nallmothu, Berg, & Sasson, 2010) From an epistemic point of view, even the positive results have some methodological weaknesses in terms of being clear about causation. Most of the cited studies of team training do not take measurements from direct observations of the clinical practitioners. Secondary markers of outcome are instead evaluated, and conclusions about the impact of simulation are drawn from what one would assume to be a simple inductive process as described by Steel: “Assume that the causal generalization true of the base population also holds approximately in related populations, unless there is some specific reason to think otherwise.” (Steel, 2007)

This is highly problematic as the outcomes measured can potentially be impacted by multiple points of input into the process. To begin with, the types of events that team-training simulation is designed to address are complex and often multifactorial in their underlying causes. The clinical providers themselves also present complexities. Staffing issues severely limit what types of clinical outcome studies can be undertaken in attempts to measure the impact of simulation interventions. If we look just at the nursing staff that could be involved in, say, an obstetric crisis at a major teaching hospital this becomes obvious. Nursing personnel often work varying shifts. Making them available for intensive simulation training is often difficult due to staffing needs. In large tertiary care settings both trained and untrained providers could come in contact with the patient population in question. If the simulation intervention is based on team interactions, and complete capture of all personnel in the training is not achieved, it becomes difficult to determine what the impact will be on measured outcome markers. Is there a minimum threshold capture rate

that is acceptable? To date this has not been specifically addressed in the team training literature. There are other ways in which simulation training appears to exhibit issue related to extrapolation as outlined by Steel. As previously discussed, the issue of measuring transference has yet to be completely addressed. There are still circumstances where an arbitrary simple inductive leap from the simulation setting to clinical is made. That is, if the trainee has improved in the simulation setting, then we can assume that performance will likely improve in clinical practice. This assumption implies that the baseline conditions, or as Steel would say the base population, holds across other referential populations and therefore that

“capacities or causal powers that exert a characteristic influence independently of context are a basis for extrapolation. However, this proposal does not adequately explain how one is to know that one is dealing with a capacity rather than a context-sensitive causal relationship, aside from already having found that the causal relationship obtains in all of the contexts in question.” (Steel, 2007)

The issue of context-sensitive causal relationships is clearly important. Colyvan tries to give an illustration of the problem in his example of walking across a narrow beam under two sets of circumstances.

“To illustrate how utilities can make a difference to a decision, suppose we are asked to walk across a narrow beam placed a few inches above the ground. If we successfully cross the beam, we win a small sum of money; if we fall off, then we miss out on the cash. In this case the utilities are such that we would elect to cross the beam provided we have no specific objection to beam walking. Now suppose the beam is placed high across a ravine, below which are crocodile-infested waters. Even though the probability of successfully crossing the beam is identical to that of the previous scenario (our beam walking skills have not altered) the utilities have changed to

such an extent that we are no longer inclined to accept the challenge.” (Colyvan, Regan, & Ferson, 2001)

If we were attempting to use a method of maximization of expected utility (MEU), this would have to be considered. Weirich would claim that we would be anticipated to change the weighting of perceived risk in our utility calculation, that is, we would actually be factoring in aversion or attraction to the “risk involved in an option when the probabilities of the states of the world determining its consequences are based on slender evidence.” (Weirich, 1986) It would seem reasonable to conclude that this risk calculation alone would change the playing field from the low-risk beam walking in Colyvan’s example to the high-risk version. But something more fundamental might be at work as well. We wouldn’t react the same in this circumstance because *the conditions themselves* are not the same –the baseline conditions have been altered, and it also seems reasonable to conclude that this also holds true in the field of healthcare simulation. No matter what degree of fidelity we create in the simulated world, the risks are clearly different when a patient’s health, and possibly life, are at stake. Because of this, the simple inductive leap from performance in simulation to performance in the real world cannot be valid, at least not under high stakes conditions.³

With this in mind, some researchers have begun to explore other innovative methods to measure transference. Randomized control trials are problematic for many of the reasons previously alluded to. In a large health center with constantly changing compliments of personnel, isolating a ‘control’ group of providers from an ‘intervention’ group may only be possible on a large scale. As mentioned earlier, O’Donnell et al attempted this approach in the implementation

³ This idea was discussed in several personal conversations with John Worrall in 2011.

of a team-based training model intended to reduce the incidence of musculoskeletal injury associated with the physical movement or repositioning of patients. (O' Donnell et al., 2011)

Two physically separate hospitals, each with similar neurologic rehabilitation units, and no movement of personnel between them, were utilized for the control and intervention cohorts. All personnel on the intervention units were trained with the simulation intervention. The outcome measure was adherence to a prescribed 10-step protocol for moving patients. This prospective, longitudinal observational design incorporated a unique approach to assessing the impact of the simulation intervention: the use of Hierarchical Task Analysis (HTA).

HTA methodology was used to deconstruct the process of moving a patient, allowing for the development of the idealized move protocol. This protocol then became the objective scoring tool by which trained observers rated personnel on both the control and intervention units as they carried out actual patient moves. This allowed not only the measurement of adherence to the protocol over time, but also meaningful interpretation of reported musculoskeletal injury rates among personnel at the intervention sites. (O' Donnell, Goode, et al., 2012; O' Donnell et al., 2011)

The pilot study for this current program of research also attempts to leverage this HTA methodology.

The obvious concern with regard to the use of HTA is another aspect of the problems mentioned earlier with extrapolation, the Extrapolator's Circle. Assumptions about the target population – in this case, our healthcare providers and their clinical actions post-simulation training- must be similar to the comparative population, here being the reactions of these trainees in the simulation setting. We have seen that this is potentially problematic based upon our discussion of Colyvan's beam walking example. HTA offers the opportunity to perhaps avoid this problem from the outset. By deconstructing the target goals into component steps, we can create a

detailed model of structured, idealized responses to clinical problems. If this deconstructive process can be done in enough detail, perhaps extrapolation itself would be unnecessary. What we would be trying to do is change the target conditions –the real-world clinical responses- to match the simulated world, and not worrying that the simulated environment is a replica of the current, often problematic real-world conditions. The extrapolation problem as it exists is only of concern when the real-world conditions are assumed to be ideal, which is exactly the opposite of the assumption of conditions in the drive to utilize healthcare simulation to enhance patient safety. In a way, this might be considered reverse modeling.

3.0 Theoretical Model: Moving Towards a New Model of Healthcare Simulation

3.1 Introduction

As we have seen, simulation training has become ubiquitous in healthcare education, and over the last decade evidence has begun to emerge that supports a long-held contention that this educational modality has the potential to change how providers practice with resultant improvement in patient outcomes. However, there are still few studies that delineate the precise relationships between the use of various simulation educational elements and student (or patient) outcomes. As a result, answering the question of which methodology to use for a particular circumstance remains elusive. Cook et al.'s (2011) systematic review and meta-analysis of the simulation literature was among the first to illustrate this problem. (Cook et al., 2011) Identifying 609 eligible studies from a pool of 10,903 articles these authors analyzed eligible papers and categorized six areas of simulation study outcomes: knowledge gain, task completion time, measures of process skills, measures of product skills, measures of behaviors related to patient care and effects on patient care. The first five of the six areas demonstrated pooled effect sizes which were characterized as 'large'. In the sixth area, 'effects on patient care', the effect size was considered 'moderate'. All of these areas demonstrated wide ranges of effect size, indicating high levels of inconsistency across these six outcomes .⁴

⁴ The I^2 statistic was used to quantify inconsistency (heterogeneity) across studies, with values greater than 50% indicating high inconsistency. The I^2 values for each of the areas was: Knowledge 96%, Time Skills 84%, Process Skills 89%, Product Skills 87%, Behavior Skills 66% and Effects on Patient Care 67%.

Given the near universal adoption of simulation in healthcare training and the recognition of value by stakeholders ranging from administrators to students, we are well beyond the point of arguing if simulation should be used. However, there are still many examples of published papers demonstrating poorly designed research methods. Possibly this is a result of something missing in the design, implementation or evaluative processes in the healthcare simulation field. This heterogeneity in quality of the published literature affects the science of simulation in two ways. First, it weakens our ability to identify the gaps in our understanding of what does and does not work and under which circumstance. Second, it creates difficulties in the development and testing of meaningful theory around the science of simulation in a way that would meet Karl Popper's famous validating Test of Falsifiability. Although the proverbial horse may be already out of the barn, it would be prudent to take a step back and examine the state of simulation science with respect to its grounding theoretical constructs. The purpose of this paper is to explain why theoretical grounding is important, examine existing models of simulation and to posit a new *functional* model of healthcare simulation that accounts for both theory and process.

3.2 The Importance of Theoretical Grounding

Cook et al.'s series of meta analyses and systematic reviews provided other meaningful insights into the current state of simulation science, most notably that the measurable impact of simulation education was large enough for the authors to recommend that there was little value in continuing to do studies comparing simulation with no simulation. (Cook et al., 2012; Cook, Brydges, et al., 2013a, 2013b; Cook, Erwin, et al., 2010; Cook, Hamstra, et al., 2013; Cook et al., 2011; Cook, Levinson, et al., 2010) However, most of the studies available for review had serious

methodological limitations, clearly identifying significant theoretical and structural weaknesses for *the science* of simulation. As Cook (2011, pp. 987-988) stated, “The important questions for this field are those that clarify when and how to use simulation most effectively and cost-efficiently. Unfortunately, the evidence synthesized herein largely *fails to inform the design of future simulation activities*. [emphasis added]”

Earlier we reviewed areas such as simulation team-training and highly specific task training, and the promising and compelling evidence for a positive impact on both educational and patient outcomes they have generated. (Barsuk et al., 2015; Barsuk, McGaghie, Cohen, O’Leary, et al., 2009; Crofts et al., 2008; Draycott et al., 2008) However, from an epistemic point of view, even these positive results do not inform us of the predictive or causal relationships between simulation design and outcomes. (Chan et al., 2010; Chan et al., 2008; R. L. Kneebone, 2016; McGaghie, Issenberg, Petrusa, & Scalese, 2010; Pusic, Boutis, & McGaghie, 2018; Stayt, 2012)

As alluded to in section one, critical, core questions need to be answered in order to help to determine the direction of future research in healthcare simulation. To date only a handful of authors have engaged in the types of foundational work that any emerging science needs.⁵ It is important to restate that while *pieces* of the simulation process have been examined in detail, no description of what the *process* of simulation entails at a macro level currently exists. The same is

⁵ An example of a structured search for articles relating to models of the simulation process: (("Models, Nursing"[Mesh] OR "Models, Educational"[Mesh]) OR "Models, Psychological"[Mesh]) AND (("Patient Simulation"[Mesh] OR "Computer Simulation"[Mesh:noexp]) OR "Manikins"[Mesh]). This search was also run separately looking for these terms as components of the title or abstract, and run again in combined form with the mesh terms. These efforts yielded potential 1010 articles of which after review only 23 related to either proposing or critiquing an existing model of the healthcare simulation process.

true of causal or predictive relationships. Are they present? If so, what do these associations look like structurally and what evidence exists in support of them?

Some aspects of simulation design and implementation have been clearly described. First, participants enjoy taking part in simulation exercises and report improved self-confidence and perceived competence. (Beyea, von Reyn, & Slattery, 2007; Harder, 2010) Second, knowledge gains have been extensively demonstrated with the use of a variety of simulation training modalities from standardized patients to human patient simulators. (Cook et al., 2011; Crofts et al., 2007; Hoffmann et al., 2007; Seibert et al., 2004) Third, psychomotor skills have been demonstrated to be good targets for the use of simulation training. (Barsuk, Ahya, Cohen, McGaghie, & Wayne, 2009; Barsuk et al., 2015) In nursing education alone there exists well over a century of examples of this. (Barsuk et al., 2015; Chamberlain, 2008; Goode et al., 2007; Herrmann, 1981, 2008; Nickerson, Morrison, & Pollard, 2011; O' Donnell J et al., 2007; JM O' Donnell & Goode, 2008; Schiavenato, 2009) These demonstrated successes support the hypothesis that there is *at least* a predictive relationship: simulation does seem to positively impact certain components of the learning process. Based on known outcomes, several best practice papers have been published which have helped to identify important simulation design features. (Issenberg et al., 2005; Mc Gaghie et al., 2009; Motola, Devine, Chung, Sullivan, & Issenberg, 2013) One simulation professional organization has even promulgated 'standards'. (INACSL, 2016) These best practices represent the most prevalent design themes identified in structured reviews of the literature while the standards represent the consensus of a group of nursing and medical educators. While these are helpful, what remains unclear are the what, when, why and how questions: *what works best, when is the timing optimal, how does it work and why do certain approaches work better than others* (Schiavenato, 2009). Our contention here is that a sound global model of the

healthcare simulation continuum is needed and that this should be a *functional* model, helping to guide efforts to answer these important questions.

In its simplest form, a causal or predictive model of healthcare simulation can be represented by a straight-line continuum from the simulation intervention to the practice patterns of individuals in the clinical setting (see **Figure 2**). In this model, the relationship between simulation educational interventions and positive changes in clinical practice is not clearly defined; it is represented by the ‘black box’ referred to in the literature as transference or translation. (Britt, Reed, & Britt, 2007; Haque & Srinivasan, 2006; Hunt, Shilkofski, Stavroudis, & Nelson, 2007; Knight, Moule, & Desbottes, 2000; JM. O’ Donnell et al., 2007; Ogan et al., 2004; Robertson, 2006; Wong et al., 2008) Much discussion has occurred over the last decade in regard to the nature of transference/translation, although not much of this discussion has focused on exactly what these terms mean, or if they even mean the same thing.⁶ Both are used as representational terms for a process by which the experiences gained during a simulation training exercise are internalized by the participants and then, hopefully, drawn upon to change clinical practice in a positive manner. Educators assume that ‘something’ is happening along this continuum, but despite significant investments of resources (financial, personnel, material and equipment) important questions remain regarding what that something else is. (Cook et al., 2011; R. Kneebone, 2003) This is

⁶ The healthcare simulation community essentially uses these terms interchangeably. While either term might suffice in the context of its use in the relevant literature, I have chosen to use transference for etymologic reasons. Translation implies that something is examined and reinterpreted, with the possible loss of important components of its meaning (context for example). It is often said of some literary works that they cannot be translated to English as subtlety and context are lost. I prefer transference because no loss of meaning is implied, and in terms of healthcare simulation targets, that should be our goal, to transfer best practice information as whole and intact as possible, allowing the learner to incorporate that information into their personal perceptive and practice framework.

because the steps along the continuum and their relationships to each other have not been completely defined, nor has the relative impact of each step been quantified. Our contention is that the development of a model of the entire healthcare simulation continuum would help to consolidate terminology surrounding the simulation process, identify potential causal or predictive relationship, describe what these associations may look like structurally and provide a guide for targeting research in empirically weak areas.

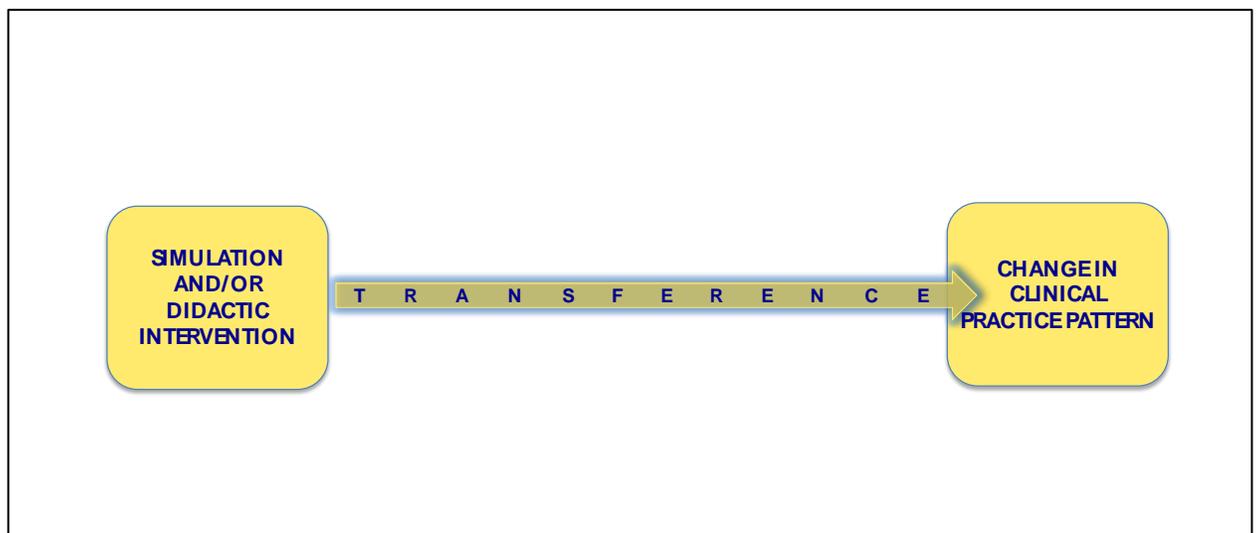


Figure 2. Simple Continuum From Simulation to Change in Clinical Practice Pattern.

Other authors agree, having identified the need for this type of foundational work. In 2005, Roger Kneebone noted with regard to simulations and simulation technologies that “a coherent theoretical structure for their use is seldom presented. Making wise judgments about the usefulness or otherwise of simulations can therefore be difficult”. (R. Kneebone, p. 549) While not offering a specific theoretical model of his own, Kneebone identified four key areas that could constitute a framework for *evaluating* simulation interventions: 1) Gaining and retaining technical proficiency; 2) the place of expert assistance in task-based learning; 3) learning within a professional context;

and 4) the affective component of learning. (R. Kneebone, 2005, p. 549) Eleven years later, he asks similar structural and theoretical questions, attempting to frame simulation “not as a static array of educational procedures but as an active principle which can transmute experience from one context into another.” (R. L. Kneebone, 2016, p. 2) In the field of nursing education, Schiavenato explicitly calls for theory to answer the ‘why question’ in regard to simulation: “A cohesive ideology is lacking for the very existence of simulation in nursing education. The missing theoretical imperative has led to a lack of direction or empirical expectation – that is *why* nursing simulation?” (Schiavenato, 2009, p. 390)

Based on reports from a broad spectrum of healthcare disciplines, it seems clear that this claim is true outside of nursing as well. (Harder, 2010; Harris, Eccles, Ward, & Whyte, 2013; R. Kneebone, 2005; Stayt, 2012) In focusing on the what, where, why and how Schiavenato echoes the conclusions of Cook that the evidence is unclear as to which simulation method is best in a particular setting. (Cook et al., 2011; Schiavenato, 2009) Harris et al. reach similar conclusions, pointing out that a sound theoretical framework would provide description and explanation of the phenomena and have the ability for prediction in the system that it describes. (Harris et al., 2013) Harder’s systematic review noted inconsistencies in the types of evaluations being done in simulation and called for the development of standardized evaluation tools specific to simulation taking into account “a type of leveling of the simulation as well as of the evaluation.” (Harder, 2010) While Harder was focusing on so-called ‘high fidelity simulation’ this argument can clearly be made for any type of simulation endeavor based on her systematic review and the Cook meta-analyses. More recently, Cheng et al. attempted to address these reporting inconsistencies by developing extensions to the Consolidated Standards of Reporting Trials (CONSORT) and the Strengthening the Reporting of Observational Studies in Epidemiology Statements (STROBE).

(Cheng et al., 2016) Stayt not only identified the weaknesses in the simulation literature, but recognized the need to examine the underlying ontology and epistemology. She utilized Carper's Patterns of Knowing in Nursing as a lens to 'reverse engineer the introduction of clinical simulation into nurse education curricula,' however Stayt does not address the fundamental issue of the need for a theoretical underpinning for the simulation process itself. (Stayt, 2012) An excellent review of constructivist, behaviorist and reflective approaches to learning is given, as well as how they might conflict or support each other within the confines of the design of a specific simulation exemplar, but Stayt does not address the fundamental issue of the need for a model describing the processes and theoretical underpinning of simulation. Finally, Parker and Myrick also identified "a need to engage in further modes of inquiry to contribute to a knowledge base from which nurse educators can draw to inform fundamentally sound pedagogical decisions regarding the development of simulation-based curriculum." (Parker & Myrick, 2012)

All of these authors recognize the need for a model elucidating the components, process and theoretical foundations to guide healthcare simulation educational efforts. But examining the variety of 'models' discussed in the simulation literature reveals that most address only isolated components of the simulation process or the leveraging of particular educational approaches. For example, Pusic, Boutis and McGaghie outline the "continuous process of hypothesis generation and model building" that is a part of any successful scientific discipline, describing as an example how learning curve theory provides "a meaningful foundation for simulation-based education research programs." (Pusic et al., 2018) They fail, however, to address the macro-level healthcare simulation continuum and instead seem to assume that healthcare simulation processes are merely composed of one or more knowledge acquisition theories. The larger view is essential to enable

simulation researchers to address the issues of *what, where, when, why and how* identified by Cook, Schiavenato and others.

3.3 Current Models and Their Role in Simulation Based Research

Having identified the need for both theoretical grounding and a model of the healthcare simulation continuum, it is important to review the literature regarding existing attempts to describe healthcare simulation processes either in terms of its constituent components, its grounding educational theory or some larger theoretical construct.

Nickerson et al. attempted to “define clinical simulation and its attributes and to add to the body of knowledge surrounding this concept”. (Nickerson et al., 2011) They utilized Walker and Avant’s concept analysis framework for theory development. (Walker & Avant, 2005) These authors provide a good presentation of the definitions of the attributes and characteristics of simulation as a concept, stressing the importance of debriefing and self-reflection. As referenced earlier, Roger Kneebone attempted to provide “a framework for considering simulation-based learning” based on four areas “that underpin simulation-based learning” – gaining technical proficiency, expert assistance, situated learning and the affective component of learning. (R. Kneebone, 2005) The roles and theoretical foundations of these areas are discussed in detail. Parker et al. present a qualitative study using Glaserian grounded theory to develop a model of empowering through fading support. (Parker & Myrick, 2012) Similarly, Harris et al. describe a method of basing simulation training on the theory of deliberate practice (EsPerT). (Harris et al., 2013) This is an extremely important concept in simulation education as evidenced by other efforts to deconstruct a target goal into component parts based on the literature and expert consultation,

then using that task analysis to drive a simulation based intervention. (Mc Gaghie et al., 2009; O' Donnell, Goode, et al., 2012; O' Donnell et al., 2011) Harris states that “improvements in cognitive skills require prolonged engagement within a domain, including deliberate practice.” and implies that the deliberate practice model will work for high-fidelity, contextually complex simulation scenarios. There may be some argument for this, but there are two confounding issues. First, repetitively running highly sophisticated, highly contextual scenarios (often done in a team training session) in the manner of the construct of deliberate practice is not logistically or operationally feasible- even if sound from an educational standpoint. Secondly, the studies cited to support the claim for cognitive and/or critical thinking improvements often don't utilize deliberate practice, begging the question of its absolute necessity in targeting these outcomes. The biggest shortcoming is that this conceptual approach is highly skewed toward psychomotor task training, and thus the effort to frame this as a potential global approach for all of simulation education falls short.

Probably the best-known attempt to devise a theoretical construct of healthcare simulation has been that of Pamela Jeffries. (Jeffries, 2005) Developed and tested for the National League of Nursing/Laerdal Simulation Study (now known as the Nursing Education Simulation Framework) Jeffries offered a “proposed framework to guide the process of designing, implementing and evaluating simulations in nursing...”. (Jeffries, 2005) Identifying the need to specify “...relevant variables and their relationships...to conduct research in an organized, systemic fashion.”, the model posits five major interacting components: teacher factors, student factors, educational practices, the simulation intervention and outcomes. This model provides a good foundation from an *educational theory perspective*, but gives few clues as to how the components fit together in a functionally relevant whole, how outcomes derive from the order and form of implementation and how they might work together in real-world processes. (Schiavenato, 2009; Wilson & Hagler,

2012; Young & Shellenbarger, 2012) The Jeffries model appears to be accurate in that instructor, student and educational factors interact, but the way in which these and other factors interact and their nexus of interaction, including the role of self-reflection and structured debriefing is not apparent. The model also does not describe the relationships between these factors in a functional way that allows easy identification and targeting for researchers or educators. In 2010 by Ravert et. al. assembled teams of simulation experts to conduct focused literature reviews in order to determine to what extent the NLN/Jeffries model was being used in the five years after its publication. The resulting publications were focused on the construct areas of participant, teacher, simulation design, educational practices and learning outcomes. The conclusion was that despite support by the Laerdal™ corporation and the National League of Nursing, the model has been incorporated in only a few isolated educational and research applications (Durham, Cato, & Lasater, 2014; Groom, Henderson, & Sittner, 2014; Hallmark, Thomas, & Gantt, 2014; Jones, Reese, & Shelton, 2014; O' Donnell, Levett-Jones, Decker, & Howard, 2012; Ravert & McAfooes, 2014). Helping in our understanding of these conclusions is the work of Wilson and Hagler who attempted to use the Jeffries model to design simulations for post-graduate nurses in the acute care setting (Wilson & Hagler, 2012). They reported that the model lacked functional utility. Specific weaknesses identified included lack of guidance in scenario design, pre-course preparation for learners, guided reflections and curricular integration. As a result of these weakness among others, the Jeffries model has not gained widespread utilization in the field.

These studies present a wide variety of attempts to identify and describe the processes and theoretical grounding associated with healthcare simulation-based education and research. What appears to be missing is the consideration of the process as a whole. There are clear descriptions of the utility of specific educational approaches for specific types of simulation events and at

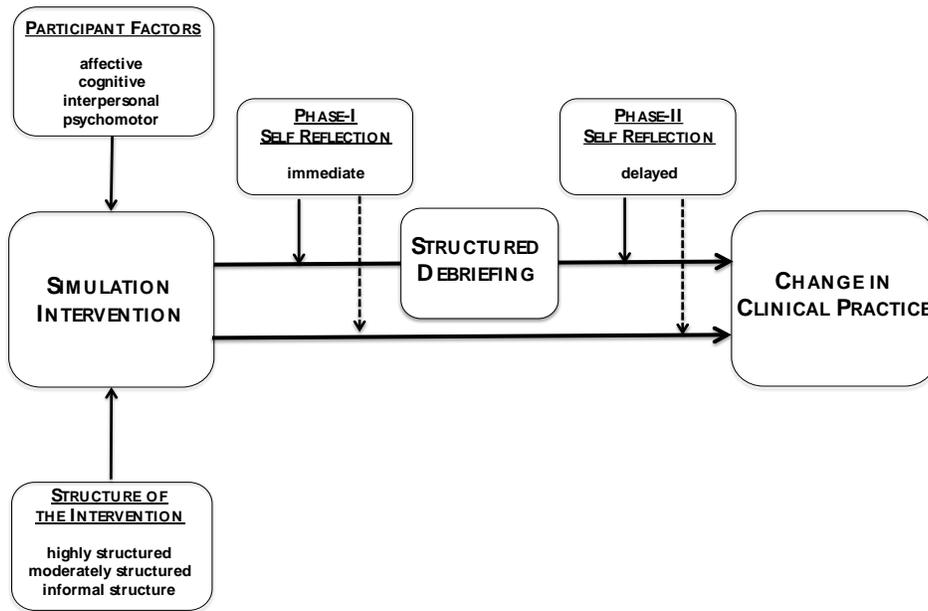
proscribed times in the simulation educational continuum. There are also vague or partial references toward providing philosophical grounding for healthcare simulation educational initiatives. A common example is the (explicit or implicit) referencing to John Dewey's conception of 'learning by doing', often in regard to psychomotor skills or repetitive practice. However, there seems to be little fundamental understanding of his work. Dewey is quite clear that task repetition, separated from the contextual application of the task, is detrimental to reflective thought and true learning (J. Goode Jr, 2016). He viewed non-contextual repetitive drilling as a "mechanical analogy of driving, by unremitting blows, a foreign substance into a resistant material." (Dewey, 1910). There is also a nearly complete disregard for how all of the components of the simulation education process fit together, in a way that would allow educators and researchers to examine individual components in their temporal and contextual setting. What is missing in the science of healthcare simulation is a 'functional' model of healthcare simulation that can help to identify areas of deficit in the science of simulation and provide guidance as to potential targets for intervention across the spectrum of the simulation events.

3.4 A New Model of Healthcare Simulation

Because of some recognized inconsistencies in the literature with regard to simulation nomenclature, it is important that adequate descriptions and definitions be provided for the terms and labels used in the subsequent discussion of the proposed model. This includes both general terms and specific operational definitions. Use of the words simulation or simulator are not always consistent in the literature. Here "Simulation Intervention" refers to the use of a 'simulator' (a physical or virtual object that is used to represent a component of a task that one wishes to

represent). (Cooper & Taqueti, 2004a) Simulators are to be utilized during “simulation”; that is, simulators are tools and simulation is curriculum. Previously the distinction was made between the words transference and translation, and here we will continue to use the word transference. More specifically, our conception of transference is that of *a cognitive process wherein the learner ideally assimilates all of the potential components of the simulation experience, forms a new working mental model of the training target and then applies this to clinical practice*. This process may, or may not, result in a *positive change* in clinical practice patterns. The work of Chi, et al is supportive of this conception. They examined the self-generated explanations of students studying pre-solved mechanics problems. A principle finding was “that ‘good’ students learn with understanding: they generate many explanations” that “result in example-independent knowledge”. (Chi, Bassok, Lewis, Reimann, & Glaser, 1989) Transference then is not a singular, indefinable process that happens at a discrete point after the simulation experience. Transference represents a process or series of processes that are posited to result in a new conceptualization of the targeted task. This new conceptualization subsequently allows the learner to change clinical practice patterns in a positive way.

Our proposed model of the healthcare simulation educational process is illustrated in **Figure 3**. The hypothesis is that the relationship is a predictive one, with direct, indirect (mediated) and moderated aspects. In this model both debriefing, and self-reflection play crucial roles. Structured debriefing is seen as an indirect (mediated pathway) that exists when a structured reflection process (debriefing) is used as a part of the simulation education intervention. While debriefing does not have to occur, it is considered to be a core component of best practices in simulation education. (Issenberg et al., 2005; Jeffries, 2005; McGaghie et al., 2010)



© Joseph S. Goode, Jr. 2014

Figure 3. Proposed Model of the Impact of Healthcare Simulation Interventions on Clinical Practice.

In considering whether structured debriefing should be classified as a mediating or a moderating variable, definitions as provided by Baron and Kenny as well as by Jaccard and Jacoby provided some guidance. (Baron & Kenny, 1986; Jaccard & Jacoby, 2010) Jaccard says of mediating variables that “other variables “work through” it to influence the outcome”. (Jaccard & Jacoby, 2010) Baron’s definition is more nuanced. The paths referred to are shown in **Figure 7**, taken from Baron and Kenny’s 1986 publication:

“A variable functions as a mediator when it meets the following conditions: (a) variations in levels of the independent variable significantly account for variations in the presumed mediator (i.e., Path a), (b) variations in the mediator significantly account for variations in the dependent variable (i.e.,

Path b), and (c) when Paths a and b are controlled, a previously significant relation between the independent and dependent variables is no longer significant, with the strongest demonstration of mediation occurring when Path c is zero. In regard to the last condition we may envisage a continuum. When Path c is reduced to zero, we have strong evidence for a single, dominant mediator.” (Baron & Kenny, 1986)

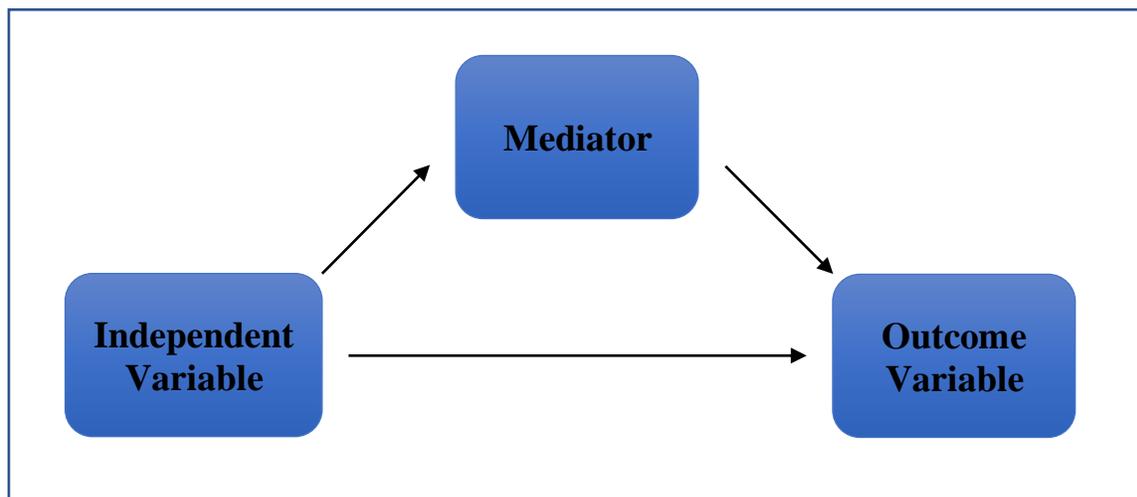


Figure 4. Causal Chain Involved in Mediation as Proposed by Baron and Kenny (1986)

Issenberg et al. undertook a structured review of the medical education literature to determine what practices worked best in simulation education. (Issenberg et al., 2005) Feedback was identified as the most frequently cited variable with structured debriefing one commonly recommended approach. This was reiterated in McGaghie et al.’s 2010 review of simulation-based medical education (SBME) and further supported by the 2016 INACSL standards. (INACSL, 2016; McGaghie et al., 2010) For this reason, structured debriefing plays an anchoring role in this proposed model. While immediate face-to-face structured debriefing protocols appear to be superior to other methods, we make no distinction among the various types of structured debriefing; even simple written feedback in conjunction with a screen-based trainer has been

demonstrated to improve performance on follow up standardized evaluative simulation scenarios. (Sawyer, Eppich, Brett-Fleegler, Grant, & Cheng, 2016; Schwid, Rooke, Michalowski, & Ross, 2001) Given these considerations, structured debriefing as proposed in this model fulfills the requirements of being a mediator, including that of variations in the independent variable causing variations in the mediator.⁷ It should also be noted that a direct effect of the simulation itself, without the mediating effects of debriefing, cannot be ruled out, and this as well is reflected in the proposed model.⁸

Other factors may have an impact on this causal relationship at various points. All of these have been classified as moderators, again based on accepted definitions. (Baron & Kenny, 1986; Jaccard & Jacoby, 2010; Mackinnon, 2011)⁹ Moderators that impact the simulation intervention itself fall into two broad categories. The first are factors that are native to each participant in the simulation education intervention. These would include baseline knowledge of the topic of the intervention, affective factors (feelings about simulation in general, prior exposure to simulation educational interventions), intrinsic psychomotor skills that may be leveraged in the simulation and interpersonal/communication skills. This is not an exhaustive list, but it does identify factors frequently cited in the existing literature. The second broad category is how the intervention itself

⁷ The structure of the simulation intervention can have a significant impact on the way that the structured debriefing can be carried out. A simulation intervention with a highly informal structure does not lend itself to a highly structured debriefing process. The optimal nature of the structure of debriefings is still under investigation, but I believe that a more formalized structure is able to better focus on the intended goals of the intervention.

⁸ Baron and Kenny indicate that mediators are “not both a necessary and a sufficient condition for an effect to occur.” (Baron & Kenny, 1986) p 1176

⁹ Baron’s conceptualization of a moderator is “a qualitative (e.g. , sex race, class) or quantitative (e.g., level of reward) variable that affects the direction and/or strength of the relation between an independent or predictor variable and a dependent or criterion variable.”(Baron & Kenny, 1986) p 1174

is structured, which can range from informal to very highly structured (e.g. hierarchical task analysis). Moderators that exert their influence after the simulation intervention are identified as **Phase-I Self-Reflection (P1-SR)** and **Phase-II Self-Reflection (P2-SR)**. Whether a structured debriefing process is used or not, it is assumed that participants in simulation reflect on their experiences after completion of the event itself. While this most likely is a semi-continuous process, it is conceptually useful to define two moderating periods of self-reflection ‘sandwiching’ the formal debriefing. The first (P1-SR) most often occurs in the period immediately following the simulation intervention, but perhaps could begin even before the entire educational intervention is completed. The second (P2-SR) occurs at a later time. Possibly the later reflection encompasses both the intervention experience itself and the review of these experiences during the structured debriefing, again supporting the position of structured debriefing as a central, anchoring mediator. P1-SR and P2-SR are somewhat analogous to Schoen’s *reflection-in-action* and *reflection-on-action* conception of reflective learning. P1-SR and P2-SR are classified as moderators primarily because these reflective processes generally occur in an unstructured or informal manner when each participant of a simulation educational event reflects on and examines the experiences of the simulation exercise. At present, we should consider these to be non-manipulatable variables.¹⁰

The operational definitions for the components of the proposed model are as follows:

1. **Participant factors** are affective, cognitive, interpersonal or psychomotor qualities that are intrinsic to each participant.

¹⁰ This is not to say that these reflective periods could not be leveraged as independent variables with targeted interventions. The problem at this time is that these reflective periods have not been extensively examined; qualitative analysis of both periods would certainly help to inform how one might effectively leverage this self-reflective process.

2. **Structure of the Intervention** is considered to be the formalized process by which the simulation educational event happens. There are many ways that this can occur. The literature supports the contention that a highly structured intervention has multiple methodological advantages. (Harris et al., 2013; Knight et al., 2000; Parker & Myrick, 2012; Tjiam et al., 2012) It helps to clearly define targets for the intervention and in turn can help to guide the structured debriefing to those targets. One example of how to achieve such a structure is with the use of hierarchical task analysis (HTA) which lists each step in a particular task and analyzes them to determine a preferential sequencing of task steps. HTA theory is “based on goal-directed behavior comprising a sub-goal hierarchy linked by plans”. (Annett et al., 2000; Shepherd, 1998; Stanton, 2006) Annett has demonstrated the use of HTA to describe the complex task management sets encountered in the military nuclear submarine setting. (Annett et al., 2000; Shepherd, 1998) HTA may be a viable tool for evaluating the impact of some components of this proposed model, being that it is scalable from the level of the individual to that of whole, complex systems. However, we will discuss later why specific details such as the structure of the intervention may be irrelevant in terms of the explanatory power of the model. (Robert W. Batterman & Rice, 2014)
3. **Phase-I Self Reflection (P1-SR) and Phase-II Self-Reflection (P2-SR)** are considered to be any reflective process done by the participant post-simulation intervention. These can be defined and described temporally if a structured debriefing is utilized, with Phase-I occurring in the period between the simulation intervention and the structured debriefing and Phase-II occurring at any time after the structured debriefing.
4. **Structured Debriefing** is, operationally, a debriefing that has a defined process, ideally drawing upon the structure of the simulation intervention itself. One example is the structured

and supported debriefing model adopted by the American Heart Association, but there are other viable debriefing approaches. (O'Donnell et al., 2009)

3.5 Discussion and Support for the New Model

The simulation literature has focused on the learning theories that underpin simulation. While important, this focus has been at the expense of an examination of the simulation process as a whole. A global, macro perspective is useful so that educators and researchers might be able to envision all the interacting parts of the simulation education process. Additionally, no one educational approach is capable of describing all that occurs in the totality of the simulation experience for the learner. Indeed, the simulation literature often touts the ability to leverage several educational approaches to target several educational domains simultaneously as a huge advantage. What gets lost is a serious consideration of the temporal aspect of this process. *These events do happen in time, and there are gaps in time where the learner is self-directing.* While obliquely addressed in the body of simulation literature, to date no real effort has been made to describe these processes and their temporal relationship to each other so that they might be targeted for supportive intervention. For example, P2-SR is a potential key target point along the proposed continuum. As of now, this is a 'black box'; we have almost no understanding of what learners take with them immediately, how they process it over time, and how this links to later clinical performance. A properly designed model can allow us to see the functional component parts of the process as a whole and across time. It also can support the examination of the mediating and moderating effects of its constituent parts. All of the major points of contention in the discussion of healthcare simulation processes were considered in the design of this proposed functional

model; the best evidence in the literature was leveraged to support the structure. Four of the key areas that were considered are addressed below.

3.5.1 Falsifiability

Kneebone raised the issue that “the assumption that such (task) learning is directly transferable to a clinical context often goes untested.”(R. Kneebone, 2005) Paradoxically, many authors make note of Karl Popper’s contention that the falsifiability of a scientific theory is crucial to its acceptance. (Popper, 1989) Contemporary authors in the field of healthcare simulation have reached the same conclusion, arguing for stronger theoretical grounding of simulation-based education and research. (R. Kneebone, 2005; McGaghie & Harris, 2018; Pusic et al., 2018) Inconsistency in this regard across the body of healthcare simulation research has contributed to the difficulties in objectively evaluating outcomes attributed to healthcare simulation. (Cheng et al., 2016; Cook et al., 2011) This proposed model, unlike others described in the literature, considers the whole of the process and clearly identifies process steps, their posited associations and temporal relationships. While this current paper initiates the process, the hope is that our model will guide further systematic reviews of the literature that clearly identify the level and quality of empirical evidence at each stage of the continuum.

3.5.2 Curricular Integration

Multiple authors have made the argument that curricular integration of simulation is inconsistent and haphazard. (R. Kneebone, 2005; McGovern, Lapum, Clune, & Martin, 2013; Stayt, 2012) A key attribute of our proposed model is its functionality: being able to define phases

of the simulation process in relation to each provides both guidance and opportunities for further investigation as to what types of simulation interventions should be used at specific stages of a defined curriculum. This model is not intended to replace the use of foundational educational theories that help in the design and development of particular simulation interventions, but recognizes that differing educational theoretical approaches may be required at different times. A structured global view of simulation will help with targeting *what* and *when*.

McGovern et al expressed concern about a predominant focus in nursing simulation on psychomotor skills and the biomedical model with the use of High Fidelity Simulation. They tried to incorporate Carper's Fundamental Patterns of Knowing into the Jeffries framework of 2007. (McGovern et al., 2013) As McGovern points out, it is important to recognize that simulation needs to address these 'other ways of knowing'. Our proposed model allows for the opportunity to reflect on where and when these non-psychomotor conditionals can be leveraged. At first glance, this model may appear to be designed in the context of a singular simulation educational intervention; however, its structure is flexible enough to be expanded. We envision that this will allow for the description of both horizontal and vertical integration of simulation exercises and to achieve an array of curricular objectives.

3.5.3 The central role of debriefing and reflection

It is now widely accepted that debriefing should be a component of all scenario-based simulations, simple or complex. (Issenberg et al., 2005; Jeffries, 2005, 2007; McGaghie et al.,

2010)¹¹ The proposed model acknowledges the importance of debriefing by identifying it as a key mediating component.

Kneebone and others cite Vygotsky and his description of the Zone of Proximal Development (ZPD) as foundational to the concept of expert guided reflection. (R. Kneebone, 2005; Parker & Myrick, 2012). Kneebone cites Tharp's (1991) view that the learner first receives external help, then works toward a process of internalization. This is not unlike Chi's concept of learning with understanding that results in "example-independent knowledge." (Chi et al., 1989) Earlier we used Structured and Supported Debriefing (SSD) as an example of an approach to the guided debriefing process. There are, of course, other viable approaches to conducting a debriefing. The merits of each is currently a topic of much discussion. Viewing these approaches not as separate from, but as part of a whole simulation continuum - through the lens of a good model - may give insight as to the particular situational strengths and weaknesses of each. We have identified three distinct phases where reflection can impact the internalization of concepts learned in the simulation continuum. These can be either structured (guided and facilitator-led, screen based and team-led) or unstructured (self-reflection). The identified unstructured reflections are the least described phases of the continuum but hold much promise for targeted description and possible intervention. In this sense as well, proposed micro-process theories such as Parker et al.'s Empowering Through Fading Support fit nicely into the proposed functional model because of the essential implied temporal component. (Parker & Myrick, 2012) In many ways, P2-SR is the time point where the concept of fading support lives or dies; the antecedents determine the effectiveness of this empowerment.

¹¹ Even the continuous feedback given during psychomotor task training can be considered a form of real-time debriefing.

3.5.4 Social and Clinical Contextualization

Burke, et al have championed leveraging Social Cognitive Theory (SCT) in designing simulation interventions. (Burke & Mancuso, 2012) They identify the need for a strong structured debriefing process, but there is little said about how SCT continues to work beyond the confines of a given simulation intervention. Self-regulation and modeling are mentioned but no evidence is given as to how these occur. Likewise, Kneebone cites Lave and Wenger's view (1991) of learning being inseparable from the social practice context and also cites Guile and Young's 2001 expansion of this concept in regard to Vygotsky's ZPD. Given its global view of simulation processes, our model can easily incorporate these concepts as foundational for most of the described components.

Many of the theoretical approaches used in nursing education and nursing simulation attempt to leverage Benner's Novice to Expert framework, but there is scant empirical evidence to support its utility. While not a simulation article per se, Nicol et al. describe a means of operationalizing Benner's model across the domains of the didactic, simulation (né skills laboratory) and clinical settings. (Nicol & Freeth, 1998) Our functional model is robust enough to encompass qualitative markers such as this and more highly structured approaches such as Harris's EsPerT. (Harris et al., 2013) More importantly, the model may help to describe the *when* and the *where* of the application of these types of approaches.

3.5.5 Explanatory Structure

In philosophy of science there are a wide range of specific definitions as to what exactly a model is. Most sources agree that models, by definition, are intended to be able to explain empirical

events or phenomena. (Bailer-Jones, 2002; Rosenberg, 2005) One possible definition proposed by Daniela Bailer-Jones is that “A model is an interpretive description of a phenomenon that facilitates access to that phenomenon.” (Bailer-Jones, 2002) These phenomena can be objects, processes or even sets of definitions, but models are usually only a shorthand for the phenomenon in question, providing partial descriptions, focusing on details of interest and often disregarding other factors (causal and non-causal) seen as less important. (Bailer-Jones, 2002; Robert W. Batterman & Rice, 2014; Rosenberg, 2005) It has often been accepted that models are the component parts of or at least support larger theories. However given the widespread use of models in actual scientific practice, many times without a clearly defined larger theoretical connection, some authors reject “the received opinion that theories prevail over models”. (Bailer-Jones, 2002; Rice, 2015) In many ways the proposed model concurs with this approach as we posit that it can incorporate other theories (e.g. learning curve theory) into its component parts without the need of explaining those particular theories. At the same time this model may lay the groundwork of a larger theory of explanation of healthcare simulation processes.

Any proposed model of healthcare simulation must describe and justify its explanatory nature as a measure of its usefulness. How and why a given model may (or may not be) explanatory – indeed the nature of scientific explanation itself – continues to be vigorously debated in philosophy of science. One approach is to claim that a model has explanatory power because it contains components that ‘represent’, ‘mirror’ or ‘map onto’ the process or system that is to be modeled, often with strict accuracy or mirroring requirements. Batterman and Rice refer to these as *common features accounts*. (Robert W. Batterman & Rice, 2014) The claim is that such a model “explains just when it has certain relevant features in common with the actual systems and that having these features in common is exactly what does the explaining.” (Robert W. Batterman &

Rice, 2014) While this may work in some circumstances, Batterman and Rice believe “that a model can meet certain extremely minimal accuracy conditions...and be explanatory. What makes such models explanatory has nothing to do with representational accuracy to any degree. Instead the models are explanatory in virtue of there being a story about why large classes of features are irrelevant to the explanandum phenomenon.” (2014, p. 356)

These types of models are called minimal models and they can “explain patterns of macroscopic behavior across systems that are heterogeneous at smaller scales.” (Robert W. Batterman & Rice, 2014) The connection between minimal models and real world systems is that they may look nothing like each other at the micro level, but that they share the same universality class. “Universality” simply means that many different systems...exhibit the same patterns of behavior at much higher scales.” (Robert W. Batterman & Rice, 2014) Batterman and Rice point to Nigel Goldenfeld, speaking from the field of physics, who asserted that “it is only important to start with the correct minimal model, i.e. that model which most economically caricatures the essential physics”. (Robert W. Batterman & Rice, 2014; Goldenfeld, 1992)

This concept that absolute representation is not always advantageous was important in the development of our proposed model. A key goal for a healthcare simulation model would be to explain over and across differing approaches at the component level of the simulation educational process (e.g. different debriefing approaches, the fidelity of the simulation event). For example, a model that mapped or mirrored in great detail to any one approach to debriefing may not be explanatory for other debriefing styles, yet globally we consider debriefing to be a best practice. Our model allows for representation of global processes and patterns, and does not rely on strict representational similarity for any of its component parts. There are common components which are placed in a recognizable temporal matrix, and while baseline operational definitions have been

assigned (mediator, moderator), the intent is to provide a process map for identifying knowledge deficits and stimulating further research to refine or revise component attributes. Rice identifies two features of what Batterman terms “universal behavior”:

1. The details of the system (those that would feature in a complete causal mechanical explanation of the system’s behavior) are largely irrelevant for describing the behavior of interest.
2. Many different systems with completely different “micro” details will exhibit the identical behavior. (Robert W. Batterman, 2002, p. 73; Rice, 2015)

For minimal models, while some common features are necessary for the phenomenon to occur, this is not the same as claiming that including these common features is what makes the model explanatory. (Robert W. Batterman & Rice, 2014) From Batterman’s perspective, for our proposed model to meet the requirements of a minimal model, three questions need to be answered (adapted from Batterman and Rice 2014, p361):

1. Why are the identified common macro-level features of this proposed healthcare simulation model (the universality class) necessary for the phenomenon to occur?
2. Why are the remaining heterogeneous details (left out or not fully described, e.g. particular debriefing style level of fidelity) irrelevant for the occurrence of the phenomenon?
3. Why do very different healthcare simulation educational interventions have these macro-level features in common?

In answer to question one, we have identified six common features that lead to an outcome resultant from any healthcare simulation educational event. Five of these features (participant factors, structure of the simulation, the simulation event itself, P1-SR and P2-SR) are there of necessity. The simulation event cannot occur without them, though the particulars of each can be

quite variable. The sixth, debriefing does not necessarily have to occur but is considered a best practice and at this point in time is likely almost universally a component of the simulation educational process in some form. The answer to question two is more complex. The heterogeneous details are the particular ways that each of the global features are implemented. Two useful examples can be given in relation to the structure of the simulation event itself and the structured debriefing. There are many ways to structure the simulation event, based upon the intended educational objectives. A good example of the potential for variability of this structure is the perceived level of fidelity needed to meet the educational objectives. A simple way to define fidelity is: “The degree to which the simulation replicates the real event and/or workplace; this includes physical, psychological, and environmental elements.” (Group., 2016) In the early days of healthcare simulation there was an emphasis on replicating the real world to the maximal level of detail possible. This was driven in part by the emergence of an entire industry around the production of simulators and associated supplies, and it still remains an issue. From ‘high-fidelity’ human patient simulators to extensive moulage, the early mantra was the more ‘real’ the better. More recently this approach has come into question as simply attempting to operationalize fidelity as a construct has proven to be difficult. (Schoenherr & Hamstra, 2017; Stokes-Parish, Duvivier, & Jolly, 2017) Indeed, attempting to consistently maintain a standardized level of fidelity across all simulations may be logistically and financially untenable at best and educationally unsound at worst. As one healthcare simulation expert has put it, “If an orange works sufficiently well for practicing subcutaneous injections, use an orange!” Similarly with debriefing there are a wide variety of approaches (very formally structured to relatively unstructured). To date, no single approach has been shown to be empirically superior. (Sawyer et al., 2016) These are just two examples of the wide degree of heterogeneity of equipment, curricular structure and educational

theoretical approaches that are leveraged at a micro-level across the macro-expanse of the simulation education continuum. Finally in answer to question three, the relevant features are the ones that must exist for the educational process to occur. Referring back to Figure 1, our simple straight-line continuum model, the simulation educational event is both necessary and sufficient; the *specific form* of the event is not (e.g. use of standardized patient versus a human patient simulator). While we posit that a functional model of healthcare simulation processes is more complex than a straight-line, the same holds true of the other core components. Based on Cook's et al.'s series of reviews, similar positive results have been achieved (at a minimum in terms of knowledge gains) regardless of the specific type of simulation approaches. (Cook et al., 2012; Cook, Brydges, et al., 2013a, 2013b; Cook, Erwin, et al., 2010; Cook, Hamstra, et al., 2013; Cook et al., 2011; Cook, Levinson, et al., 2010) Debriefing style, level of fidelity, type of simulator, underlying educational theoretical approach – these are the irrelevant factors. That there is a simulation event, that a debriefing occurs, that there are broadly speaking educational theoretical approaches that can be leveraged and participant affective factors to be considered - these constitute the universality class of this minimal model.

3.5.6 Summary

Because the science of simulation, at least in healthcare, is still emerging, work remains to be done in regard to construction of theoretical models and constructs of simulation and the component parts of these models. In comparison to other models of healthcare simulation in the literature, our proposed model identifies components of a simulation educational encounter that are amenable to empirical testing. This functional model of simulation in healthcare, is justified with calls in the published literature for just such a guiding framework.

This work is incomplete; there is an obvious need for a stringent process of validation be carried out. Future research in regard to this model would include: 1) confirmation of the role of the identified components (e.g. do the assigned roles in the model of mediator versus moderator hold; are the temporal placements of components valid); 2) clarification of the role of currently undefined components such as unguided self-reflection (P1-SR and P2-SR) given that there is little data to describe these phenomena. These are likely rich targets for qualitative analysis; 3) further exploring the role of Hierarchical Task Analysis (HTA) as a tool to both guide the simulation intervention and measure performance in relation to the model. Answering specific research questions that target components of our model will help to define the relationships between those components and may help to determine whether any global theoretical statements can be made. At the very least, a fuller understanding of the processes involved throughout the continuum of a simulation educational experience may allow the refinement of the delivery of these experiences (e.g. designed for specific types of learners or designed for specific types of clinical targets like psychomotor tasks or group interactions). Finally, a structured historical analysis of the role of simulation in healthcare could provide insight into philosophical and ethical issues related to this model and may inform our understanding of component parts.

In regard to utility of the model, we understand that our proposed structure may not be the only way to describe or explain healthcare simulation processes. Rosenberg supplies a fine example of the Nobel laureate Richard Feynman proposing an alternative formulation of a Newtonian system that focuses on how gravitational forces operate (classically over long distances instantaneously, in Feynman's version in points as close as one would like). (Rosenberg, 2005) Both work as explanations of a Newtonian gravitational system, but in some circumstances one

model may provide a better description of phenomenon. Likewise, our hope is that this model is a starting point for further exploration and discussion.

4.0 Pilot Study

Adherence to and Retention of an HTA Optimal Task Set for Standard Induction of General Anesthesia (OTS-SIGA): Measuring Transference of Skills Learned in a Simulation-Based Educational Intervention for First Term Student Registered Nurse Anesthetists

4.1 Methods and Development of the Study Protocol

4.1.1 Introduction

The use of simulation is widespread in healthcare education, and the potential impact of its use large. This is especially true for nursing education as we look to address problems with obtaining clinical experiences, develop critical thinking skills and create methods to measure the impact of simulation interventions. There is substantial empirical evidence in support of a predictive relationship between simulation training interventions and knowledge acquisition. This has been extensively demonstrated with the use of a variety of simulation training modalities from standardized patients to human patient simulators. (Cook et al., 2011; Crofts et al., 2007; Hoffmann et al., 2007; Seibert et al., 2004) Data to support changes in clinical practice and improved patient outcomes are quite limited. Among the first to document improved clinical outcomes was DeVita et al. in a retrospective analysis which reported a reduction in code-related mortality after implementing highly structured simulation team training for an in-hospital Medical Emergency

Team (MET). (DeVita & Minnini, 2004) Since then other authors have pointed to areas such as simulation team-training and highly specific task training, which are generating promising and compelling evidence for a positive impact on both educational and patient outcomes. (Barsuk et al., 2015; Barsuk, McGaghie, Cohen, O'Leary, et al., 2009; Crofts et al., 2008; Draycott et al., 2008) One example is the simulation-based mastery learning (SBML) approach to teach central venous catheter (CVC) insertion skills. Barsuk et al. in a series of studies regarding the use of mastery learning techniques in central venous catheter insertion training have demonstrated retention and transference of simulation acquired skills, improved patient outcomes and healthcare system cost savings. (Barsuk et al., 2010; Barsuk, McGaghie, Cohen, Balachandran, et al., 2009; Cohen et al., 2010) They reported immediate post-intervention internal jugular CVC insertion skill retention of 100% with impressive 6- and 12-months post-training skill retention rates (82.4% and 87.1%, respectively). (Barsuk et al., 2010) What is missing from the literature is an attempt to measure the impact of simulation education on retention and transference of knowledge and skill for a more complex healthcare process. The overall objective of this research project was to build on and begin to expand that body of evidence by assessing the impact of a simulation intervention on just such a complex task, an induction of general anesthesia.

In the scope of this project we utilized a Hierarchical Task Analysis (HTA) derived protocol for the standard induction of general anesthesia (OTS-SIGA) to train new Student Registered Nurse Anesthetists (SRNAs), then measured their adherence to the process steps in the clinical setting. We also attempted to measure whether repeating the HTA-derived OTS-SIGA simulation training would have an impact on knowledge and transference of simulation-developed skills to the clinical environment. Specific outcome markers of this study included: knowledge retention regarding the process of a standard induction of general anesthesia, ability to complete

OTS-SIGA steps in the clinical environment, and measurement of specific outcome markers related to that process. The self-reported confidence of SRNAs in their ability to complete the OTS-SIGA steps was also measured to assess the impact of this simulation training intervention on self-efficacy. Outcomes are reported in the results section. Here we describe the development of our OTS-SIGA and its associated data collection tools, processes for rater training and how these impacted the design and implementation of the study.

4.1.2 Background

As a part of the standard Nurse Anesthesia BSN to MSN curriculum at the University of Pittsburgh, all SRNAs receive didactic training and a series of preparatory workshops (approximately 60 hours of classroom content and 40 hours of targeted simulation training) prior to a Mock Induction of Anesthesia simulation session. Mock Induction simulation educational exercises have been conducted in some form at the University of Pittsburgh for more than 15 years and have evolved over time. It is a core component of the preparation process for new SRNA entry into the clinical setting for the University of Pittsburgh Nurse Anesthesia Program. Prior to this study, only anecdotal evidence was collected on the impact of this training on the performance of SRNAs in their first clinical rotation. At present, the Mock Induction simulation is structured around an HTA derived optimal task set for a standard induction of general anesthesia (OTS-SIGA) (see Appendix A). The process of developing such a task set through HTA methodologies has been previously described by O'Donnell and Goode (O' Donnell, Goode, et al., 2012; JM O' Donnell & Goode, 2008) and is also discussed in section 1.3.1.

The primary objective of this pilot study was to assess the impact over time of the Mock Induction simulation for first term SRNAs. Comparisons were made between two groups

randomized to group assignment: first term SRNAs who received initial OTS-SIGA training only (SIGA-O group) and first term SRNAs who received both the initial OTS-SIGA training *plus* refresher training (SIGA+R group). The specific aims were to: 1) examine whether there were differences between these two groups in retention of knowledge about the process of a standard induction of general anesthesia during their first clinical rotation; and 2) examine whether there were differences in adherence to the process steps of the OTS-SIGA in the clinical setting between the SIGA-O and the SIGA+R groups during their first clinical rotation. Outcomes included knowledge retention as measured by a 20-item multiple-choice Standard Induction of General Anesthesia Knowledge Assessment Instrument (SIGA-KAI), adherence to the optimal task set steps in both the simulation and clinical settings, and clinical outcome markers relevant to the induction process.

4.1.3 Experimental Design, Setting and Sampling

The study utilized a randomized control trial design with repeated measures. Students were randomized to two study arms: initial SIGA training only (SIGA-O, n=12) and initial SIGA training *plus* refresher training (SIGA+R, n=12). The setting for the study was the University of Pittsburgh School of Nursing, Nurse Anesthesia Program classrooms and labs as well as affiliated clinical sites appropriate for first-year student registered nurse anesthetists.

A convenience sample of SRNAs was recruited from the cohort of incoming students beginning the Masters Nurse Anesthesia Program in Fall Term (August) 2016 (total available n=36). This 28-month full-time graduate program prepares professional nurses for the role of certified registered nurse anesthetist (CRNA). The requirements for admission to this program are a BSN degree and a minimum of 1 year of full time experience working as a professional nurse in

a critical care setting. All graduate students in the Nurse Anesthesia Program participate in a Mock Induction workshop as a standard part of the curriculum in the laboratory portion of the Basic Principles of Anesthesia course. All students are subsequently supervised in the induction of general anesthesia in the clinical setting by both credentialed CRNA and anesthesiologist instructors. Potential participants, by virtue of meeting minimum acceptance criteria for the Nurse Anesthesia Program, met inclusion criteria; there were no exclusion criteria. A recruitment script describing this University of Pittsburgh Human Resource Protection Office-approved study was distributed to these students in class and via e-mail and informed consent was obtained.

An *a priori* power analysis was conducted for two groups (t-tests for two independent means) assuming equal sample sizes. This analysis was performed with G*Power (version 3.1) using an estimate of a large effect size of 0.95 ('large' defined by Cohen, 1992, as 0.80 of a population standard deviation difference between the means), significance level (α) of 0.05 and power ($1 - \beta$) of 0.80. (Jacob Cohen, 1988; J. Cohen, 1992) The effect size of 0.95 was based on the average of effect sizes across six simulation measurement domains as identified in Cook's 2011 meta-analysis. (Cook et al., 2011) To achieve this level of predictive power, it was calculated that a total of $n = 38$ participants (19 per group) would be required (see **Figure 5**). The total available participant pool was 36; after recruitment the final number of actual participants in the study was 24. Given this sample size, the level of power that could be achieved with $n = 12$ for each arm of the study was calculated (see **Figure 6**). The analysis indicates a 60% chance of detecting a large effect size (set at 0.95 based on the Cook meta-analysis) with significance set at the 0.05 level (two tailed).

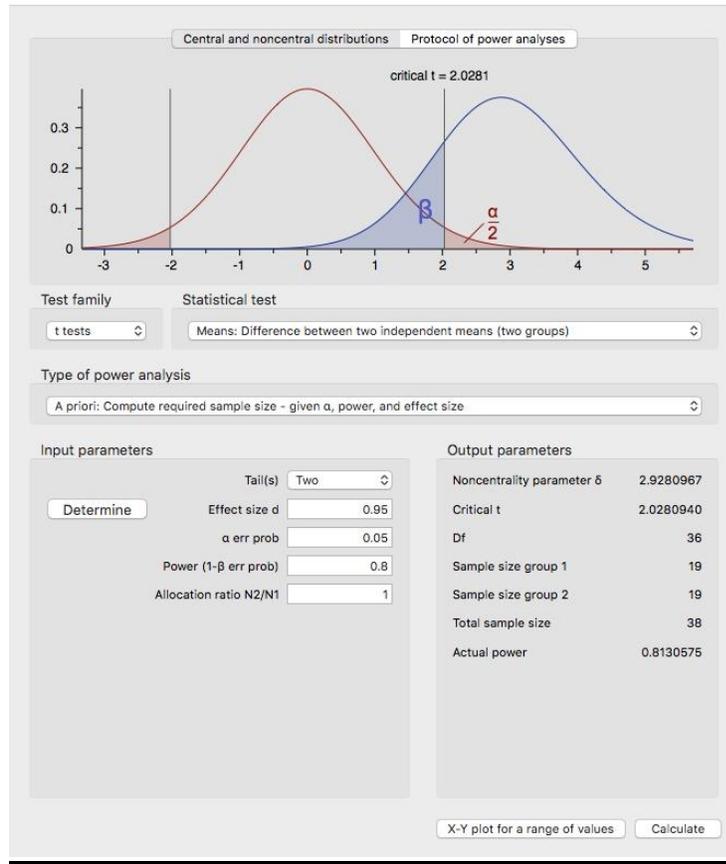


Figure 5. A priori power analysis for two groups (t-tests for two independent means) assuming equal sample sizes.

Note: A large effect size (as defined by Cohen, 1992) of 0.95 was assumed at a significance level of .05 and Power ($1 - \beta$) = 0.8 (based on effect sizes for simulation identified by Cook, 2011). To achieve this level of predictive power, it was calculated that a total of $n = 38$ participants (19 per group) would be required.

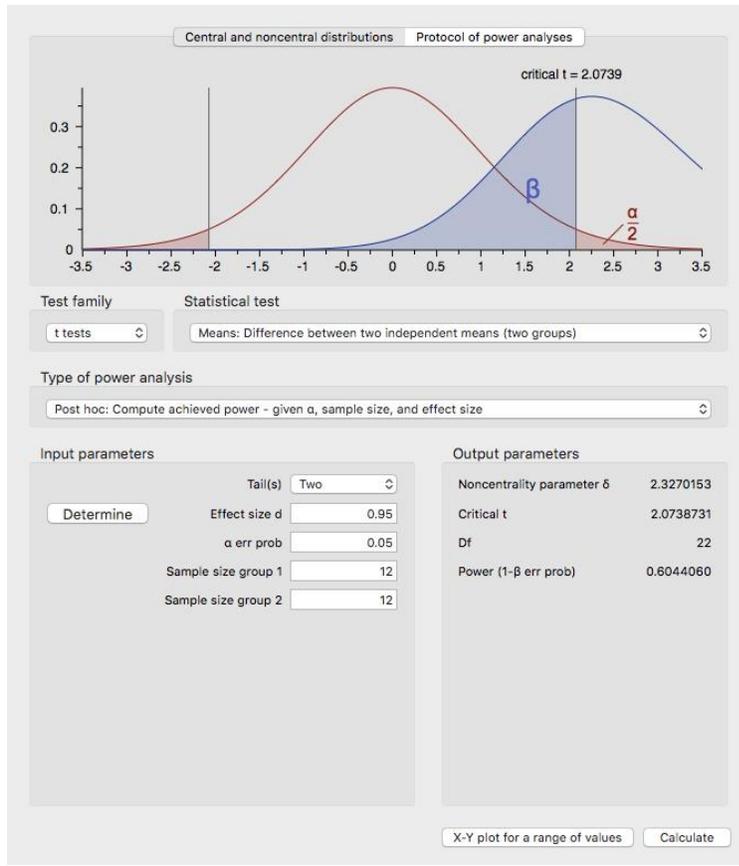


Figure 6. Power analysis for $n = 12$ for each arm of the study indicated a 60% chance of detecting a large effect size (defined by Cohen, 1992, as .95 of a population standard deviation between the means) with significance set at the 0.5 level(two-tailed).

4.1.4 Study Timeline and Methodology

A graphic timeline of the protocol and the data collection points is attached (**Figure 7**). To place the specific study data collection points in context, this timeline includes events that were part of the normal educational process for this cohort of first term SRNAs in the University of Pittsburgh Nurse Anesthesia Program. The data collected, and the timing of the data collection are reviewed below.

Clinical Training Intervention Timeline

Activities and Measurements Standard Training (SIGA-O) and Training PLUS Refresher (SIGA+R) Cohorts	Pre-Study Standard Curricular Events			OTS-SIGA Study Events		
	7 Weeks Pre-Mock Induction	4 Weeks Pre-Mock Induction	Mock Induction Training	Clinical Weeks 1 – 3	Clinical Week 9	Clinical Week 17
	Didactic Content (60 hours)	Mock Induction Simulation Exercise Introduced	Recruitment Period		Mock Induction Refresher Training <i>SIGA+R Cohort Only</i>	
	Simulation Exercises (40 hours) ^a	Mock Induction Practice Sessions ^b				
			Demographic Survey			
			SIGA-KAI <i>pre/post simulation</i>	SIGA-KAI	SIGA-KAI	SIGA-KAI
			OTS-SIGA <i>live scoring: sim</i>	OTS-SIGA <i>live scoring: clinical</i>	OTS-SIGA <i>live scoring: sim</i>	OTS-SIGA <i>live scoring: clinical</i>
			OTS-SIGA <i>video scoring: sim</i>		OTS-SIGA <i>video scoring: sim</i>	
				Interviews ^d	Interviews ^d	Interviews ^d
				Clinical Outcome Markers ^c		

^aSimulation workshops/exercises included: assessment, monitoring and positioning, intravenous catheter placement, arterial line placement, spinal administration, suturing, anesthesia gas machine, basic airway management, central venous catheter placement, and medication preparation.

^bEach SRNA received two 2-hour sessions for practicing the Mock Induction process. These sessions include 1:1 instruction with a simulation teaching assistant.

^cClinical outcome markers include: airway assessment skills, success rate in securing the airway for general anesthesia, competence in performing the OTS-SIGA steps (preceptor assessment), confidence in performing the OTS-SIGA steps (SRNA self-report) and the number of adverse events such as oral soft tissue or dental injury. The goal was for measures for these outcome markers be recorded each day that study participants were assigned to clinical.

^dInterviews will be conducted with a subset of up to 10 members of each cohort at each time point.

Figure 7. Study Timeline

4.1.4.1 Standard Preparation and Familiarization with the OTS-SIGA Protocol

In the seven weeks prior to Mock Induction, all SRNAs, regardless of whether they eventually participated in the pilot study or not, underwent the same education and training processes. This included approximately 60 hours of didactic content and 40 hours of targeted simulation training. Approximately four weeks prior to the Mock Induction event, all SRNAs were introduced to the OTS-SIGA and the expectations of the Mock Induction workshop were explained. SRNAs were then assigned to a date for the Mock Induction simulation exercise (7 or 8 students per session) and were instructed to familiarize themselves with the OTS-SIGA. Each student was given background information for an assigned ‘patient’. This included basic history and physical information the history of the present illness and the proposed surgical procedure. These ‘patients’ were created to represent a variety of common co-morbidities and common surgical procedures. Each SRNA was given the assignment of developing an anesthesia care plan that included performing an induction of general anesthesia. As a part of the preparation process, all SRNAs were assigned to two 2-hour structured practice sessions in the simulation lab. These sessions provided students with the opportunity to discuss and practice the steps of the OTS-SIGA. Practice sessions were supervised by second year anesthesia graduate student teaching assistants who had undergone training to become familiar with the OTS-SIGA and the educational goals of the Mock Induction (this process is described in the rater training section). All practice sessions and the Mock Induction simulation training intervention were conducted at the University of Pittsburgh School of Nursing Simulation Laboratory.

During the four-week Mock Induction preparation period recruitment for the pilot study was accomplished. Informed consent was obtained by trained research study assistants from those

students willing to participate in the research protocol. Regardless of study participation, all SRNAs in the potential study population participated in Mock Induction as this is considered to be standard training. The principal investigator did not participate in the Mock Induction sessions, and the course instructors did not know which SRNAs had agreed to participate in the study.

4.1.4.2 Baseline Data Collection

Demographic data were collected via a questionnaire accessed through the password protected on-line site. All SRNAs (study participants and non-participants) accessed the course through the Winter Institute for Simulation Education and Research (WISER) web portal and were directed to the demographic survey. On the day of Mock Induction, all SRNAs were asked to complete the 20-item multiple-choice Standard Induction of General Anesthesia-Knowledge Assessment Instrument (SIGA-KAI) as a standard part of the simulation education event (**Appendix B**). This assessment was administered both pre- and post-simulation. The SIGA-KAI was also administered at several other points to study participants as described below.

4.1.4.3 Baseline Mock Induction Event

The OTS-SIGA simulation training and performance evaluation was conducted at the University of Pittsburgh School of Nursing Simulation Laboratory. SRNAs presented their prepared anesthetic management plan to a second year anesthesia graduate student teaching assistant playing the role of CRNA. Adjustments were made to the management plans verbally as appropriate and the SRNA proceeded to implement their induction plan. A faculty member served as the anesthesiologist for the induction. Other SRNAs present were given assigned observational tasks to discuss in the later debriefing. The goal was for each SRNA to complete all of the steps of the OTS-SIGA, either independently or after prompting from the scenario CRNA or

anesthesiologist. The scoring tool used to assess subject performance was based on the HTA-derived induction process (see **Figure 1 and Appendix A**). Scoring was done by a trained rater observer in communication with the faculty member directing the Mock Induction from the control room. The live scoring was used in a formative manner during the debriefing portion of the simulation event, but all training sessions were also recorded using digital video, as is the usual educational practice for Mock Induction. Post-hoc video review and scoring by trained raters were accomplished for all study participants' sessions.

After the completion of the Mock Induction sessions, the study participants were randomized to either the SIGA-O or SIGA+R groups. Computer generated randomization was performed by the WISER information technology and data management staff.

4.1.4.4 Clinical Observation Period

All SRNAs began their first clinical rotation within two weeks after the Mock Induction exercise. SRNAs were assigned to clinical sites that have existing clinical affiliation agreements with the University of Pittsburgh School of Nursing. This initial clinical rotation lasted for 17 weeks. SRNAs were in the clinical setting two days a week for the first five weeks and then for three days a week thereafter. Various measures were used throughout this first clinical rotation including:

- i. The 20-item multiple choice Standard Induction of General Anesthesia-Knowledge Assessment Instrument (SIGA-KAI) was administered at three time points: clinical weeks one, nine and seventeen.
- ii. Live observations in the clinical setting measured adherence to the OTS-SIGA process steps. These observations were scheduled to be made at the beginning of the clinical rotation (weeks one through three), the middle of the clinical rotation

(week nine), and the end of the clinical rotation (weeks sixteen and seventeen). The OTS-SIGA Scoring Application (see **Figure 1**) was used for data collection by trained observational raters. There was an extended time range for the initial observation period as the types of clinical experiences first term SRNAs have in the initial weeks of their first clinical rotation vary and will include orientation day(s) where limited or no patient care is performed.

- iii. To supplement live observations, a clinical assessment tool was used by the precepting CRNA and/or anesthesiologist to evaluate SRNA competence in completing the OTS-SIGA steps. This Standard Induction of General Anesthesia Clinical Assessment Application tool included five dichotomously scored questions (Yes or No) regarding five key steps of the induction process and a 10-point anchored global rating scale of the overall performance of the SRNA during induction and intubation (**Figure 8**). SRNAs participating in the pilot study were asked to identify themselves as such to their daily clinical preceptors. Clinical preceptors were asked to complete an assessment for every induction of general anesthesia that they performed with a study SRNA. They were also asked to record a standardized 7-point airway assessment on these patients.
- iv. A self-reported confidence scale of the SRNA's ability to complete the OTS-SIGA. This was measured by a 10-point anchored rating scale completed for each induction of general anesthesia event the study SRNAs participated in during the clinical rotation. SRNAs were also asked to record a standardized 7-point airway assessment on their patients. These data were also recorded via the Standard Induction of General Anesthesia Clinical Assessment Application tool (**Figure 9**).

- v. SRNA self-reports of the incidence of oral soft tissue or dental injuries was extracted from the Nurse Anesthesia Program's clinical event tracking system.

ISER
Education and Simulation
Improving Healthcare

Standard Induction of General Anesthesia Clinical Assessment Application

Pick One!

Student survey

Facilitator survey

Preceptor Assessment Tool [back to main page](#)

Student :

Case Date : 2016

Case # :

Preceptor Last Name :

Below is a list of tasks associated with performing a standard induction of general anesthesia that first rotation SRNAs need to become competent doing. The criteria to complete each task are listed. All of the criteria for a task must be completed independently by the SRNA with very little or no prompting for you to select "Yes".

Yes **No** **NA**

Equipment Verification :

Criteria:

- Complete gas machine checkout (first case) or at minimum system pressure check
- Appropriate induction/intubation equipment organized and available
- Backup equipment identified or available (e.g. manual resuscitator, independent light source, location of code button)

Yes **No** **NA**

Apply Monitors :

Criteria:

- Confirm presence of safety strap on patient
- Placement of
 - Pulse oximeter
 - BP cuff
 - EKG leads
 - Other case specific monitors

Yes **No** **NA**

Induction Steps :

Criteria:

- Ability to state induction medications, dosage and sequence of administration, when asked
- Assesses patient and monitors during induction
- Assesses for apnea and unconsciousness
- Responds to patient eyes at appropriate time

Yes **No** **NA**

Mask Ventilation :

Criteria:

- C and E finger positions
 - 10" on mask, 12" fingers on bony structures
- Places mask from nose to chin, pull-in, clear lips and tongue of compression
- Use one-handed technique for mask ventilation
- Avoid PEEP > 20 cm H₂O
- If respirations ineffective- reposition head, correctly place oral airway

Yes **No** **NA**

Laryngoscopy and Intubation :

Criteria:

- Open mouth, clear lips, inserts blade on right side, sweep and lift tongue, jousting motion
- Keep eyes on vocal cords area in view
- Insert ETT under direct vision, stylette removed with approximation
- Insert ETT to appropriate depth for patient
- Inflate cuff

Yes **No** **NA**

I would rate the overall performance of this SRNA during this induction and intubation as:

Continuous Prompting and/or Correction Needed Minimally Prompting and/or Correction Needed Able to function Independently

Patient Information:

Oral Opening

Mallampati Classification

Thyromental Distance (cm)

Cervical Range of Motion

Mandibular Length (cm)

Cricothyroid Membrane Palpable

Submit !!!

Figure 8. SIGA Clinical Assessment Application Tool (SIGA-CAA) Preceptor Portal View.

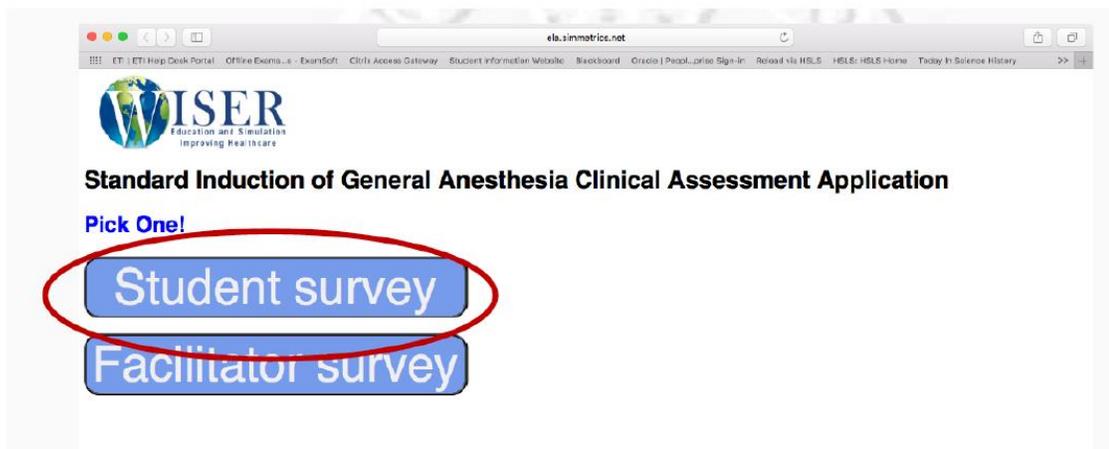


Figure 9. SIGA Clinical Assessment Application Tool (SIGA-CAA) SRNA Portal View.

4.1.4.5 Refresher Simulation Event

During the ninth week of clinical training, the SIGA+R group returned to the simulation lab for a refresher Mock Induction session and evaluation. These sessions were conducted in the same manner as the initial Mock Induction, the difference being that study participants went through the refresher session alone (initial standard Mock Inductions are done with other class members present to take advantage of potential peripheral learning). The principal investigator was again not present for these sessions. The OTS-SIGA Scoring Application was used to assess

participant performance in the same manner as with the baseline Mock Induction training. Post-hoc scoring of the digital video of the refresher simulations was also performed to ensure scoring accuracy. The SIGA-O group did not receive this additional training.

4.1.5 Measures

4.1.5.1 Demographic data

Data collected included age, gender, handedness, educational background, prior simulation experience, years of professional experience as a critical care nurse and total years of professional experience in any kind of nursing. The demographic survey took less than 15 minutes to complete.

4.1.5.2 Knowledge assessment

The 20-item Standard Induction of General Anesthesia Knowledge Assessment Instrument (SIGA-KAI) was developed from the didactic content presented to SRNAs regarding basic airway management and induction of general anesthesia. This content is a part of the University of Pittsburgh Nurse Anesthesia Program curriculum and referenced to best evidence in the literature. Questions were drawn from a pool of items developed by the research team. Each administration of the SIGA-KAI included two questions directly related to each of the 10 major process steps of the OTS-SIGA. Questions were repeated randomly throughout the course of the study. The SIGA-KAI was administered at five time points: immediately pre-and immediately post- the Initial Mock Induction, then during the first, ninth and seventeenth weeks of each participant's subsequent clinical rotation. The Refresher simulation event also occurred during the ninth week. Scores were expressed as the percentage correct. Study participants accessed the SIGA-KAI through the

password-protected on-line course for Mock Induction (NURSAN-2700). The SIGA-KAI took approximately 19 minutes to complete.

4.1.5.3 Adherence to the OTS-SIGA Process steps

Adherence was measured in two ways. First, direct observation by trained raters using a scoring tool developed from the HTA-derived induction process map (OTS-SIGA). The scoring tool was deployed as a web-based application (developed by WISER information technology staff) that was able to be accessed from any tablet or smartphone platform (**Figure 1**). This is the same observational tool that was used for assessment during the Mock Induction simulation training educational exercise. Each step of the induction process was scored in real-time. Possible scores for each step included ‘completed’, ‘completed after prompting’, or ‘not available’. In Hierarchical Task Analysis nomenclature, the Overall task – here induction of general anesthesia – is considered the superordinate goal. The major steps of the superordinate goal are termed subgoals which themselves are then broken down into operational steps. The ten primary subgoals of the OTS-SIGA were composed of these types of operational steps. The primary mode of scoring was to determine if the primary subgoals were achieved. If a single operational step of a subgoal was not completed independently by the SRNA that subgoal was considered not successfully completed. These scoring definitions and guidelines were used for both live clinical scoring and post-hoc digital video scoring of simulation events. Historically during the Mock Induction sessions, the standard of training is to have each participant complete every step, regardless of how much prompting is required. The ‘not available’ scoring option was intended to standardize the use of this tool between the simulation lab and the later clinical setting where it was possible that the clinical preceptors would not allow the SRNA to do a process step or may have completed it themselves for purposes of clinical safety and efficiency. A comments field allowed raters to enter

information regarding the submitted scoring (e.g., explanations of why a specific score was entered). After collection, observational data were submitted and stored in a secure database managed by WISER information technology personnel. These observations did not require any additional time commitment from the study participants, as clinical training is a part of the standard educational process for nurse anesthesia students. The principal investigator as well as observational raters trained in the use of the OTS-SIGA scoring tool performed the data collection. The raters (including the PI) were blinded to the randomization assignment of each study participant.

Logistically, it was impossible for trained observers to be present at every induction of general anesthesia in which study SRNAs participated. The second way adherence was measured took advantage of the fact that a CRNA is always assigned with an SRNA in the clinical setting, allowing capture of data from induction events when trained observers were not present. Clinical preceptors were asked to assess the SRNA study participant's performance during a standard induction of general anesthesia by answering five questions based on five key steps of the OTS-SIGA induction process. Preceptors did not have the option to score a step as prompted but they did have the option to score a subgoal step as NA. Additionally, the preceptor was asked to use a 10-point anchored global rating scale to rate the overall performance of the SRNA during induction and intubation. Preceptor scoring was accomplished via the SIGA Clinical Assessment Application tool (**Figure 8**). This secure cloud-based tool and its associated database were also developed and administered by WISER information technology staff. Links to the application tool were available to CRNA preceptors on an internally accessed anesthesia department website or from any smartphone platform. Familiarization with the tool was accomplished by email distribution of an explanatory PowerPoint presentation to all CRNAs at the clinical sites to which

study SRNAs were assigned. The principal investigator also visited each participating clinical site to provide in-person training and explanation of the tool. A resource person for each site, generally the site Chief CRNA or CRNA Educator, received a more in-depth orientation. In addition, contact information for the principal investigator was made available for questions or to report technical difficulties with the on-line application. CRNA preceptors were incentivized with complimentary Starbucks gift cards (total amount determined by how many evaluations were completed by each preceptor by the end of the study period).

4.1.5.4 Self-report of confidence in the ability to complete the OTS-SIGA

The SRNA study participant was asked to provide a self-rating of their confidence in performing the steps of the OTS-SIGA using a 10-point anchored rating scale that was also accessed via the SIGA-CAA tool (**Figure 13**). They were asked to do this for each induction event for which they participated. All of the clinical observation data were recorded without using any patient identifiers. The completion of the SIGA Clinical Assessment Application tool information was estimated to take no more than 5 minutes per case.

4.1.5.5 Agreement of SRNA airway assessment with expert airway assessment

As a standard part of a patient pre-operative evaluation, anesthesia providers routinely perform an airway assessment. Ideally seven anatomic factors are assessed: oral opening and dentition, Mallampati score, mandibular length, thyromental distance, cervical range of motion, ability to palpate the cricothyroid membrane and anterior mandibular displacement. SRNAs are also required to perform an interview and assessment of each patient as a part of their training. We asked both the clinical preceptor and the SRNA study participants to record their airway assessment findings (without recording patient identifiers) for each general anesthesia case that

they were involved in together. These assessments were also done via the SIGA-CAA. The clinical preceptor and the SRNA accessed separate portions of the application independently for each clinical encounter so that the student could not see the preceptor's airway assessment data. The SRNA airway assessments were later matched in the database with those of the CRNA so that comparisons could be made.

4.1.5.6 Self-report and clinical preceptor report of the ability to secure the airway (successful placement of an endotracheal tube or a supraglottic airway device)

While we did not anticipate 100 percent compliance with entering induction events into the SIGA-CAA, all University of Pittsburgh SRNAs are required to log intubation attempts and other clinical experience data in the password-protected, HIPPA compliant, secure, on-line Typhon case entry system. These data are entered in a de-identified manner (e.g., total number of tracheal intubation attempts, number of successful intubation attempts) so that no connection with particular patients can be made. These data allowed for comparisons of success rates reported by the CRNA preceptors and the self-reported success rates of the SRNA study participants.

4.1.5.7 Self-report of the incidence of oral soft tissue or dental injuries

All University of Pittsburgh Nurse Anesthesia Program students are required to report any adverse clinical event to the program via a Clinical Event tool. The Nurse Anesthesia Program Director reviews all of these submissions. Adverse events directly related to study participant performance of SIGA (e.g., oral soft tissue or dental injury during laryngoscopy and intubation, corneal abrasion during mask ventilation) were collated by the Program Director and provided to the PI without identifiers. Comparisons were made between the SIGA-O groups and the SIGA+R groups as well as to historical rates of similar injuries.

4.1.6 Observational Rater Training

Observational raters included graduate student research assistants, practicing CRNAs and research study faculty. All raters underwent training that included familiarization with the OTS-SIGA and use of the OTS-SIGA Scoring Application. The training process included:

- i. Review of the operational definitions of each process step of the OTS-SIGA (**Appendix C**)
- ii. Real-time scoring of three previously recorded ‘exemplar’ Mock Induction simulation events (permission to use these videos was obtained from the participants in the videos).
- iii. Comparison of rater video scores with expert-level scoring was done to assure accuracy. The principal investigator reviewed these for differences between the rater training scores and the idealized scores and provided written feedback to raters in training.
- iv. A final exemplar (test) video was viewed and scored by the raters in training. A threshold score of 90% or above agreement with the idealized score was required for each rater prior to collecting live data in the clinical setting or reviewing study digital videos. If this was not achieved additional training and a second testing were done.

Idealized (expert) rater scores were generated for all of the rater training videos. As described earlier, possible scores for each step of the OTS-SIGA included: completed (Y), not completed (N), completed after prompting (P) and either not applicable to the case or not available to the SRNA (NA). Inter-rater reliability was determined by comparing the scores for the final

exemplar training video for each observational rater with those of the idealized score generated by the principal investigator.

4.1.7 Statistical Analysis Methods

Idealized (expert) rater scores were generated for all of the rater training videos. As described earlier, possible scores for each step of the OTS-SIGA included: completed (Y), not completed (N), completed after prompting (P) and either not applicable to the case or not available to the SRNA (NA). Inter-rater reliability was determined by comparing the scores for the final exemplar training video for each observational rater with those of the idealized score generated by the principal investigator. Thirteen observational raters completed the training. Each scored step of the OTS-SIGA was matched with that of the idealized score. Cohen's kappa (κ) was used to calculate agreement for each process step between the idealized score and each of the raters. Two agreement comparisons were made, first for the overall score of ten major subgoals as described above, and then for all fifty operational tasks that made up those subgoals.

Means and standard deviations were calculated for all demographic data (age, gender, handedness, educational background, prior simulation experience, years of professional experience as a critical care nurse and total years of professional experience in any kind of nursing). Independent samples t-tests were run between the SIGA+R and SIGA-O groups for age, ICU experience and total nursing experience. The SIGA-KAI results were analyzed using two-way mixed ANOVA.

As discussed earlier, live observations were made by trained raters throughout the duration of the SRNAs first clinical rotation. Data were combined and examined in time blocks of four week intervals (Time Block 1 = weeks 1-4, Time Block 2 = weeks 5-8, Time Block 3 = weeks 10-

13 and Time Block 4 = weeks 14-17). Difficulties with training and deploying raters (as discussed earlier in the methods section) prevented the collection of any useful data during weeks five through eight so no data from Time Block 3 were included in the analysis. The data were aggregated by Time Block as total proportions of primary subgoal steps scored as yes, no, prompted or NA. Because classical repeated measures analysis methods (RM-ANOVA and RM-MANOVA) omits participants with incomplete data across the repeated measurements, a linear mixed model approach was undertaken, and the results were analyzed. Several models were tested for each set of response data with the final model being selected based on either Akaike's Information Criterion (AIC) score or Schwarz's Bayesian Criterion (BIC) score (whichever was the lowest).

For the previously described preceptor ratings of the SRNA's performance on five key steps derived from the OTS-SIGA HTA analysis, and for the preceptor overall rating, we asked that these assessments be submitted throughout the duration of the SRNAs first clinical rotation (17 weeks) for each induction of general anesthesia that the preceptor performed with an SRNA in the study cohort. Data were aggregated and examined in the four time blocks described above as total proportions of primary subgoal steps scored as Yes, No or NA. As previously described, linear mixed modeling was again used for the analysis.

4.1.8 Summary

We have described here the methods and measures used for a pilot study utilizing a Hierarchical Task Analysis (HTA) derived protocol for the standard induction of general anesthesia (OTS-SIGA) to train new SRNAs, then measure their adherence to these process steps in the clinical setting. We also attempted to measure whether repeating the HTA-derived OTS-

SIGA simulation training would have an impact on knowledge and transference of simulation-developed skills to the clinical environment. As a part of this process we developed powerful and flexible web-based applications for measuring performance in both the simulation and clinical settings. We also created a rigorous process for training observational raters to use those tools in combination with the OTS-SIGA.

4.2 Results of the Pilot Study

4.2.1 Inter-rater Reliability

Thirteen observational raters completed the training. Each scored step of the OTS-SIGA was matched with that of the idealized score. Cohen's kappa (κ) was used to calculate agreement for each process step between the idealized score and each of the raters. Two agreement comparisons were made, first for the overall score of ten major subgoals as described above, and then for all fifty operational tasks that made up those subgoals. The results of these two analyses are summarized in **Tables 1 and 2** below. Agreement for the ten subgoals ranged from $\kappa = .189$ to 1.000, with strength of agreement ranging from slight to perfect. Agreement for the fifty individual operational tasks ranged from $\kappa = .486$ to .768, with strength of agreement ranging from moderate to substantial.

Table 1 Cohen's kappa for each rater, comparing their ten subgoal operational task scores for a post-training test video with those of a benchmark score. Strength of agreement assessment from Landis and Koch (1977).

Rater	Cohen's κ	Strength of Agreement	Std Error	CI lower 95%	CI upper 95%	Approximate Sig.
1	.620	Substantial	.169	.289	.951	<.005
2	.610	Substantial	.179	.259	.961	<.005
3	.315	Fair	.172	-.022	.652	.029
4	.605	Moderate	.178	.256	.954	<.005
5	.737	Substantial	.164	.416	1.059	<.005
6	.189	Slight	.157	-.121	.499	.233
7	.467	Moderate	.169	.136	.798	.003
8	.595	Moderate	.183	.236	.954	<.005
9	.600	Moderate	.178	.251	.949	<.005
10	.872	Almost perfect	.119	.639	1.105	<.005
11	1.000	Almost perfect	.000	-	-	<.005
12	.737	Substantial	.157	.429	1.045	<.005
13	.737	Substantial	.164	.416	1.058	<.005
Mean	.622	Substantial				

Table 2 Cohen’s kappa for each rater, comparing their fifty subgoal scores for a post-training test video with those of a benchmark score. Strength of agreement assessment from Landis and Koch (1977).

Rater	Cohen’s κ	Strength of Agreement	Std Error	CI lower 95%	CI upper 95%	Approximate Sig.
1	.668	Substantial	.104	.464	.872	<.005
2	.693	Substantial	.098	.501	.885	<.005
3	.560	Moderate	.116	.333	.787	<.005
4	.589	Moderate	.111	.371	.807	<.005
5	.768	Substantial	.090	.592	.944	<.005
6	.486	Moderate	.112	.266	.706	<.005
7	.514	Moderate	.112	.294	.734	<.005
8	.499	Moderate	.111	.717	.281	<.005
9	.571	Moderate	.111	.353	.789	<.005
10	.728	Substantial	.096	.540	.916	<.005
11	.702	Substantial	.099	.508	.896	<.005
12	.666	Substantial	.099	.472	.860	<.005
13	.725	Substantial	.097	.535	.915	<.005
Mean	.627	Substantial				

4.2.2 Demographic Characteristics of the Sample

The demographic data for the cohort as a whole and for the SIGA-O and SIGA+R groups are summarized in **Table 3**. The mean age of all participants (expressed as mean \pm standard deviation) was 27.25 ± 3.23 years. The mean number of years of intensive care unit experience was 2.56 ± 1.22 and the mean total years of experience was 3.21 ± 4.11 . These findings are similar to recent Nurse Anesthesia Program Admissions data for age (mean 26.52 ± 3.68), ICU experience (2.74 ± 1.27) and total nursing experience (3.44 ± 1.82). Of the total participant cohort, 62.5% were female and 37.5% male. Recent Nurse Anesthesia Program Admissions data show that on average each admissions class is 33.5% male. Two participants appeared as outliers in terms of age, contributing to positive skewness of the total cohort. These two participants were randomized

to the SIGA+R group. ICU experience was not normally distributed as assessed by Shapiro Wilk's test ($p = .033$). Independent samples t-tests were run between the SIGA+R and SIGA-O groups for age, ICU experience and total nursing experience. The assumption of homogeneity of variances was violated, as assessed by Levene's test for equality of variances, for age ($p = .027$). There was a statistically significant difference in mean age between the SIGA-O and SIGA+R groups, with the SIGA+R group being older, 3.50 ± 3.90 years [mean difference \pm standard deviation], $t(12.844) = -3.112$, $p = .027$. There were no statistically significant differences between the two groups in terms of years of ICU nursing experience and total nursing experience.

Table 3 Demographic Characteristics for the total sample and by group.

	N	Range	Min	Max	Mean	Std. Deviation	Variance	Skewness Statistic	Std. Error	Kurtosis Statistic	Std. Error
Total Cohort											
Age (years)	24	13.00	24	37	27.25	3.234	10.457	1.766	.472	3.346	.918
ICU Experience (years)	24	4.92	1.08	6.00	2.563	1.221	1.491	1.238	.472	1.364	.918
Total Experience (years)	24	5.42	1.25	6.67	3.213	1.329	1.767	.914	.472	1.015	.918
SIGA-O Group											
Age (years)	12	3.00	24	27	25.50	1.087	1.182	-.255	.637	-1.128	1.232
ICU Experience (years)	12	3.00	1.25	4.25	2.293	1.001	1.002	1.128	.637	.295	1.232
Total Experience (years)	12	3.00	1.25	4.25	2.753	1.060	1.124	.071	.637	-1.316	1.232
SIGA+R Group											
Age (years)	12	13.00	24	37	29.00	3.742	14.000	1.087	.637	1.040	1.232
ICU Experience (years)	12	4.92	1.08	6.00	2.833	1.398	1.955	1.143	.637	1.208	1.232
Total Experience (years)	12	4.59	2.08	6.67	3.673	1.452	2.109	1.052	.637	.400	1.232

4.2.3 Standard Induction of General Anesthesia-Knowledge Assessment Instrument

There were outliers in the collected data as assessed by boxplots, but with one exception these could be attributed to a very truncated (high) range of scores on the assessment (not real outliers). These scores were not removed from the analysis. The SIGA-KAI scores were not normally distributed as assessed by Shapiro-Wilk's test of normality for the SIGA-O group at Post-Simulation ($p = .011$) and Clinical Week-9 ($p = .029$) and for the SIGA+R group at Pre-Simulation ($p = .019$) and Post-Simulation ($p = .002$). This was probably attributable to outliers in each of these circumstances. The decision was made to proceed with two-way mixed ANOVA without removing the outliers as these were no errors in data entry for these points. Additionally, ANOVA is fairly robust as to deviations from normality. Homogeneity of variances ($p > .05$) and the equivalence of covariance matrices between groups ($p = .977$), as assessed by Levene's test

and by Box's M test, respectively, were confirmed. Mauchly's test confirmed the assumption of sphericity, $\chi^2(9) = 14.201, p = .116$. There was not a statistically significant interaction between the refresher simulation and time on SIGA-KAI scores, $F(4, 88) = .831, p = .509$, partial $\eta^2 = .036$. The main effect of time showed a statistically significant difference in mean SIGA-KAI scores at the different time points, $F(4, 88) = 19.607, p < .001$, partial $\eta^2 = .471$. There was not a statistically significant difference in mean SIGA-KAI scores between the SIGA-O and SIGA+R groups, $F(1, 22) = .004, p = .953$, partial $\eta^2 = <.001$. The statistically significant main effect of time changes (all $p < .001$) occurred between Pre-Simulation scores (mean = 92.29 ± 6.08) and Clinical Week-2 (mean = 80.63 ± 6.81), and Pre-Simulation scores and Week-17, the end of the clinical rotation (mean = 84.58 ± 5.50). The same pattern held for the Post-Simulation (mean = 93.54 ± 4.03) scores and Clinical Weeks 2 and 17 (**Figure 10**). The means for both groups and the total study cohort at each measurement point are summarized in Table 4.

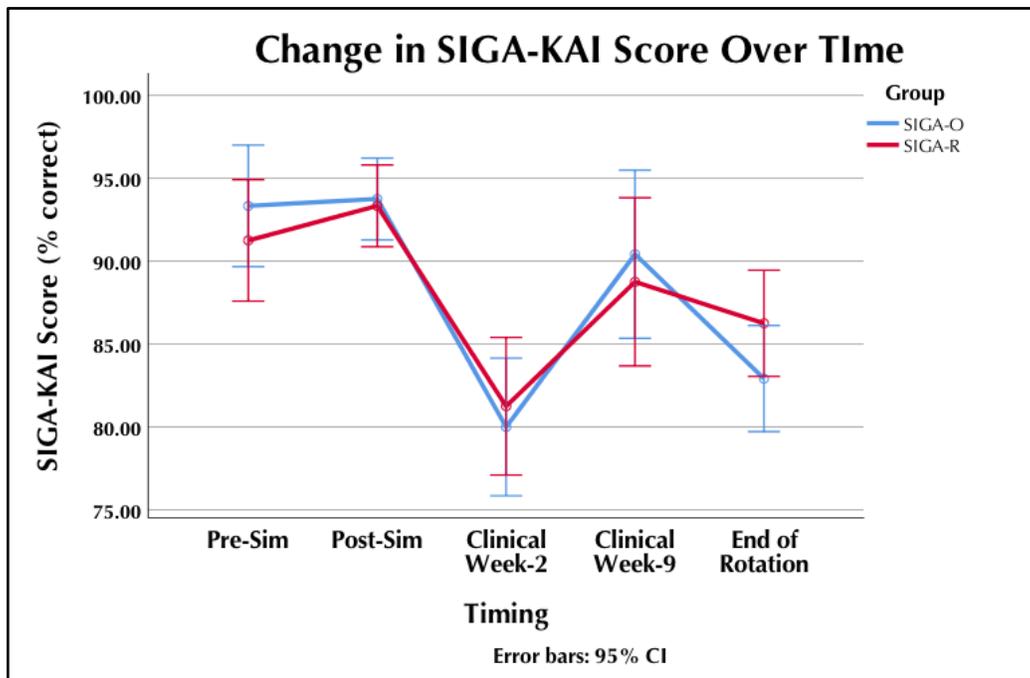


Figure 10 SIGA-KAI Scores Over Time

Table 4 SIGA-KAI Score Means and Standard Deviations

Mean SIGA-KAI Scores Over Time				
	Group	Mean	Std. Deviation	n
Pre-Simulation	SIGA-O	93.33	4.44	12
	SIGA+R	91.25	7.42	12
	Total	92.29	6.076	24
Post-Simulation	SIGA-O	93.75	3.77	12
	SIGA+R	93.33	4.44	12
	Total	93.54	4.03	24
Clinical Week 2	SIGA-O	80.00	7.39	12
	SIGA+R	81.25	6.44	12
	Total	80.62	6.81	24
Clinical Week 9	SIGA-O	90.42	10.10	12
	SIGA+R	88.75	6.44	12
	Total	89.58	8.33	24
Clinical Week-19	SIGA-O	82.92	4.98	12
	SIGA+R	86.25	5.69	12
	Total	84.58	5.50	24

4.2.4 OTS-SIGA Live Observations

The results of the linear mixed modeling analysis for the OTS-SIGA live observations are summarized in **Table 5**. Tests of Fixed Effects indicated that there were no significant group main effects or group-by-time interactions, suggesting that the refresher Mock Induction did not demonstrate any significant effect. For the SIGA+R group, there were two participants whose ages were identified in data screening as potential outliers. Analyses were rerun accounting for age as a covariate; this did not significantly impact the results. Additional analyses were run using clinical site and both clinical site and age as covariates, again with no significant impact on the results. When examining the total cohort, there was a significant effect of time. Estimated mean proportions from the model for each possible score at each time block demonstrate this and are summarized with model test statistics in **Table 6**. These trends are graphically represented in

Figure 11. Statistically significant increases in the proportion of subgoals scored as ‘yes’ were demonstrated between Time Blocks 1 and 3 and Time Blocks 1 and 4. The proportion of subgoal steps prompted and the proportion scored as ‘not available’ to the SRNA decreased for those same time intervals. Examined separately, the SIGA-O group and the SIGA+R group both demonstrated similar trends over time, but there were not significant differences between the two groups for any Time Block for any of the possible subgoal scores.

Table 5 Linear Mixed Modeling Results for OTS-SIGA Live Scoring: Proportion of Subgoal Steps Scored Yes, No, Prompted or Not Available

Step Score Proportion	TB-1 (weeks 1-4) Mean (SE)	TB-3 (weeks 1-13) Mean (SE)	TB-4 (weeks 14-17) Mean (SE)	Test Statistics
Proportion 'Yes'	0.073 (0.023)	0.208 (0.033)	0.156 (0.018)	$F(\text{group}) = 0.738, p = .414$ $F(\text{time}) = 6.349, p = .006$ $F(\text{GxT}) = 0.961, p = .436$
Proportion 'No'	0.177 (0.036)	0.223 (0.061)	0.294 (0.028)	$F(\text{group}) = 0.013, p = .910$ $F(\text{time}) = 3.120, p = .060$ $F(\text{GxT}) = 0.286, p = .754$
Proportion 'Prompt'	0.133 (0.022)	-0.021 (0.032)	0.021 (0.017)	$F(\text{group}) = 0.510, p = .481$ $F(\text{time}) = 10.234, p = .002$ $F(\text{GxT}) = 2.741, p = .122$
Proportion 'NA'	0.633 (0.045)	0.516 (0.071)	0.531 (0.036)	$F(\text{group}) = 0.290, p = .603$ $F(\text{time}) = 1.849, p = .177$ $F(\text{GxT}) = 0.480, p = .624$

Note: Significance indicated for Group effects (differences between SIGA-O and SIGA+R groups); Time effects (differences between time periods (1, 2, 3 or 4) for the entire study group); and Group x Time Effects (for the time block indicated: 1, 2, 3 or 4). CS = compound symmetric; Autoregressive; AR1 = TP = Toeplitz; AIC = Akaike's Information Criterion; BIC = Schwarz's Bayesian Criterion

Table 6 OTS-SIGA Live Scoring Effect of Time: Proportion of all Subgoal Scores by Time Block for the Entire Cohort

	Repeated Variance-Covariance Type					
	CS		AR1		TP	
	AIC	BIC	AIC	BIC	AIC	BIC
Proportion Yes	-51.612	-48.877	-52.399	-49.665	-51.127	-47.025
group effects	none		none		none	
time effects	1 vs 4		1 vs 3, 1 vs 4		TBs 1-3, 1-4	
group*time effects	none		none		none	
Proportion No	-25.933	-23.199	-25.398	-22.664	-24.003	-19.901
group effects	none		none		none	
time effects	1 vs 4		1 vs 4		TBs 1-4	
group*time effects	none		none		none	
Proportion Prompted	-54.724	-51.990	-55.426	-52.692	-55.063	-50.961
group effects	none		none		none	
time effects	1 vs 3, 1 vs 4		1 vs 3, 1 vs 4		TBs 1-3, 1-4, 3-4	
group*time effects	none		none		TB 3	
Proportion NA	-12.445	-9.710	-12.786	-10.052	-11.534	-7.433
group effects	none		none		none	
time effects	none		1 vs 4		none	
group*time effects	none		none		none	

Note: ‘Yes’ = step completed independently by the SRNA; ‘No’ = step not completed by the SRNA; ‘Prompt’ = SRNA required prompting to complete the step; ‘NA’ = Step completed by clinical preceptor or not applicable to clinical situation.

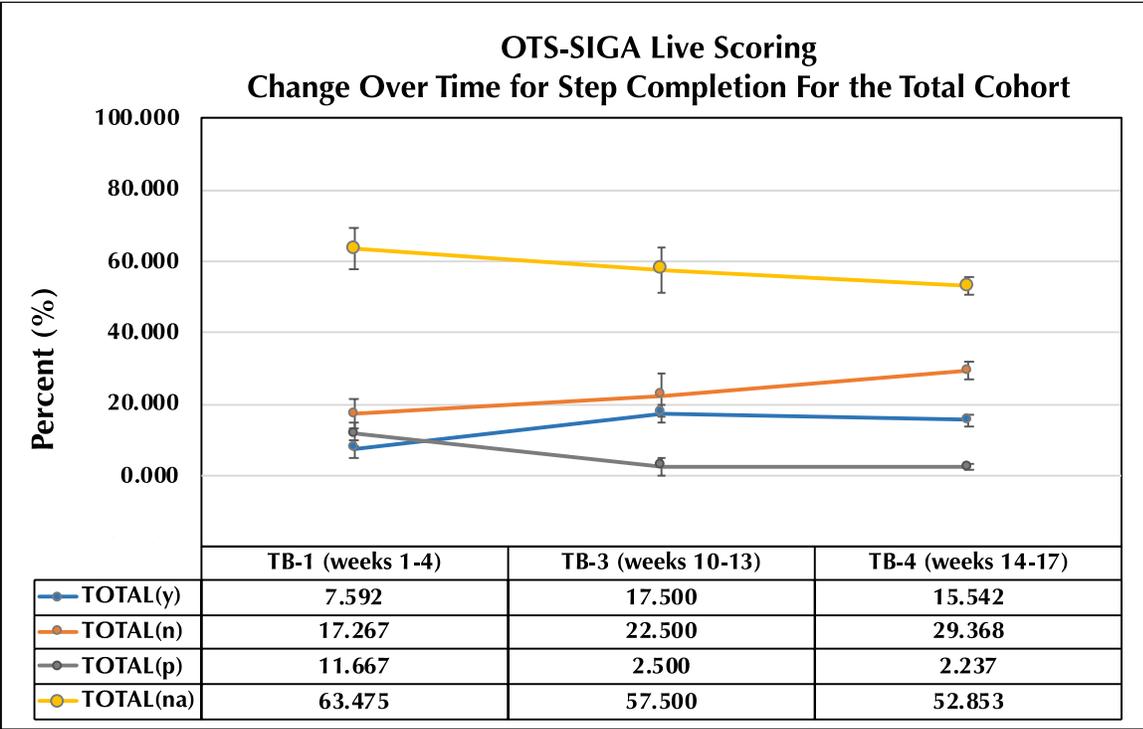


Figure 11 OTS-SIGA Live Scoring Change in Proportion of all Subgoal Scores by Time Block for the entire cohort (mean \pm SE)

4.2.5 SIGA-CAA Preceptor Scoring

As previously described, clinical preceptors who worked with SRNAs in the study cohort were requested to fill out the preceptor portion of the SIGA Clinical Assessment Application. This was composed of five of the ten primary subgoals taken from the OTS-SIGA: Step 1, Equipment Verification; Step 3, Apply Monitors; Step 5, Induction Steps; Step 6, Mask Ventilation; and Step-7, Laryngoscopy and Intubation.

Tests of Fixed Effects for ‘group’ indicated significance for Question 5 ‘laryngoscopy and intubation’. Tests of Fixed Effects for ‘time’ indicated significance for Question 2 ‘apply monitors’, Question 3 ‘induction steps’, Question 4 ‘mask ventilation’ and Question 5 ‘laryngoscopy and intubation’. The refresher Mock Induction did demonstrate significant Fixed

Effects (Group x Time) for ‘apply monitors’, ‘mask ventilation’ and ‘laryngoscopy & intubation’. Estimated mean proportions from the model for each possible score at each time block are summarized with model test statistics in **Tables 7 through 11**. These trends are graphically represented in **Figures 14 through 18**.

For the total cohort, effects of time were noted at several points. Question 2 (apply monitors) had statistically significant increases in the proportion of ‘yes’ ratings between Time Blocks 1 and 3 and Time Blocks 1 and 4 (**Figure 15**). The ‘no’ responses demonstrated statistically significant decreases for these same Time Blocks. For Question 3 (induction steps) there were significant increases in the proportion of ‘yes’ ratings between Time Blocks 1 and 4 (**Figure 16**). For Question 4 (mask ventilation) significant increases in the proportion of ‘yes’ ratings between Time Blocks 1 and 4 were noted along with significant decreases in no responses for the same time frame. (**Figure 17**). NA responses increased between Time Blocks 1 and 2. For Question 5 (laryngoscopy and intubation) there was a significant increase in ‘yes’ ratings at Time Block 4 while ‘no’ ratings decreased in a similar manner (**Figure 18**).¹²

¹² For *Tables 7 through 11*: Significance indicated for Group effects (differences between SIGA-O and SIGA+R groups); Time effects (differences between time periods (1, 2, 3 or 4) for the entire study group); and Group x Time Effects (for the time block indicated). CS = compound symmetric; Autoregressive = AR1; TP = Toeplitz; AIC = Akaike’s Information Criterion; BIC = Schwarz’s Bayesian Criteria.

Table 7 Linear Mixed Modeling Results for Preceptor Scoring of Equipment Verification: Proportion of Subgoal Steps Scored Yes, No or Not Available

Preceptor Question-1 (Equipment Verification)						
Repeated Variance-Covariance Type						
	CS		AR1		TP	
	AIC	BIC	AIC	BIC	AIC	BIC
Proportion Yes	-6.503	-2.803	-6.474	-2.774	-2.789	4.611
group effects	none		none		$(p = .002)$	
time effects	none		none		none	
group*time effects	none		none		none	
Proportion No	-9.609	-5.909	-9.319	-5.619	-7.295	0.106
group effects	none		none		none	
time effects	none		none		none	
group*time effects	none		none		none	
Proportion NA	-136.896	-133.196	-136.896	-133.196	-132.896	-125.496
group effects	none		none		none	
time effects	none		none		none	
group*time effects	none		none		none	

Table 8 Linear Mixed Modeling Results for Preceptor Scoring of Apply Monitors: Proportion of Subgoal Steps Scored Yes, No or Not Available

Preceptor Question-2 (Apply Monitors)						
Repeated Variance-Covariance Type						
	CS		AR1		TP	
	AIC	BIC	AIC	BIC	AIC	BIC
Proportion Yes	-10.331	-6.631	-11.775	-8.075	-8.522	-1.121
group effects	none		none		none	
time effects	1-3 ($p = .042$) 1-4 ($p = .022$)		1-3 ($p = .019$) 1-4 ($p = .022$)		1-3 ($p = .016$) 1-4 ($p = .021$)	
group*time effects	1 ($p = .021$)		1 ($p = .023$)		1 ($p = .028$)	
Proportion No	-15.970	-12.270	-18.208	-14.507	-15.578	-8.177
group effects	none		none		none	
time effects	1-3 ($p = .050$)		1-3 ($p = .018$)		1-3 ($p = .015$)	
group*time effects	1-4 ($p = .026$)		1-4 ($p = .029$)		1-4 ($p = .026$)	
	1 ($p = .006$)		1 ($p = .007$)		1 ($p = .009$)	
Proportion NA	-164.522	-160.822	-164.522	-160.822	-160.522	-153.122
group effects	none		none		none	
time effects	none		none		none	
group*time effects	none		none		none	

Table 9 Linear Mixed Modeling Results for Preceptor Scoring of Induction Steps: Proportion of Subgoal

Preceptor Question-3 (Induction Steps)						
Repeated Variance-Covariance Type						
	CS		AR1		TP	
	AIC	BIC	AIC	BIC	AIC	BIC
Proportion Yes	38.841	42.541	38.817	42.517	42.797	50.198
group effects	none		none		none	
time effects	1-4 ($p = .032$)		1-4 ($p = .026$)		1-4 ($p = .033$)	
group*time effects	none		none		none	
Proportion No	34.829	38.529	34.793	38.493	38.662	46.062
group effects	none		none		none	
time effects	none		none		none	
group*time effects	none		none		none	
Proportion NA	-85.947	-82.247	-90.764	-87.064	-101.311	-93.911
group effects	none		none		none	
time effects	none		none		none	
group*time effects	none		none		none	

Steps Scored Yes, No or Not Available

Table 10 Linear Mixed Modeling Results for Preceptor Scoring of Mask Ventilation: Proportion of Subgoal

Preceptor Question-4 (Mask Ventilation)						
Repeated Variance-Covariance Type						
	CS		AR1		TP	
	AIC	BIC	AIC	BIC	AIC	BIC
Proportion Yes	47.514	51.215	44.794	48.494	45.384	52.785
group effects	none		none		none	
time effects	1-4 ($p = .012$)		1-4 ($p = .010$)		1-4 ($p = .030$)	
group*time effects	none		none		none	
Proportion No	46.373	50.073	44.886	48.586	47.256	54.656
group effects	none		none		none	
time effects	1-4 ($p = .013$)		1-4 ($p = .010$)		1-4 ($p = .023$)	
group*time effects	none		none		none	
Proportion NA	-151.321	-147.621	-151.321	-147.621	-147.321	-139.921
group effects	none		none		none	
time effects	1-2 ($p = .043$)		1-2 ($p = .043$)		1-2 ($p = .043$)	
group*time effects	2 ($p = .013$)		2 ($p = .013$)		2 ($p = .013$)	

Steps Scored Yes, No or Not Available

Table 11 Linear Mixed Modeling Results for Preceptor Scoring of Laryngoscopy & Intubation: Proportion of Subgoal Steps Scored Yes, No or Not Available

Preceptor Question-5 (Laryngoscopy & Intubation)						
Repeated Variance-Covariance Type						
	CS		ARI		TP	
	AIC	BIC	AIC	BIC	AIC	BIC
Proportion Yes	50.248	53.948	50.474	54.175	53.777	61.178
group effects	$(p = .026)$		$(p = .017)$		$(p = .029)$	
time effects	1-4 ($p = .011$) 2-4 ($p = .005$)		1-4 ($p = .016$) 2-4 ($p = .007$)		1-4 ($p = .012$) 2-4 ($p = .003$)	
group*time effects	3-4 ($p = .025$)		3-4 ($p = .031$)		3-4 ($p = .035$)	
	2 ($p = .004$)		2 ($p = .005$)		1 ($p = .028$)	
Proportion No	41.368	45.068	39.873	43.573	43.673	51.073
group effects	$(p = .023)$		$(p = .030)$		$(p = .026)$	
time effects	1-4 ($p = .006$) 2-4 ($p = .008$)		1-4 ($p = .011$) 2-4 ($p = .007$)		1-4 ($p = .021$) 2-4 ($p = .009$)	
group*time effects	2 ($p = .0011$)		2 ($p = .006$)		2 ($p = .006$)	
Proportion NA	-7.392	-3.692	-7.219	-3.519	-187.964	-180.564
group effects	none		none		none	
time effects	none		none		none	
group*time effects	none		none		none	

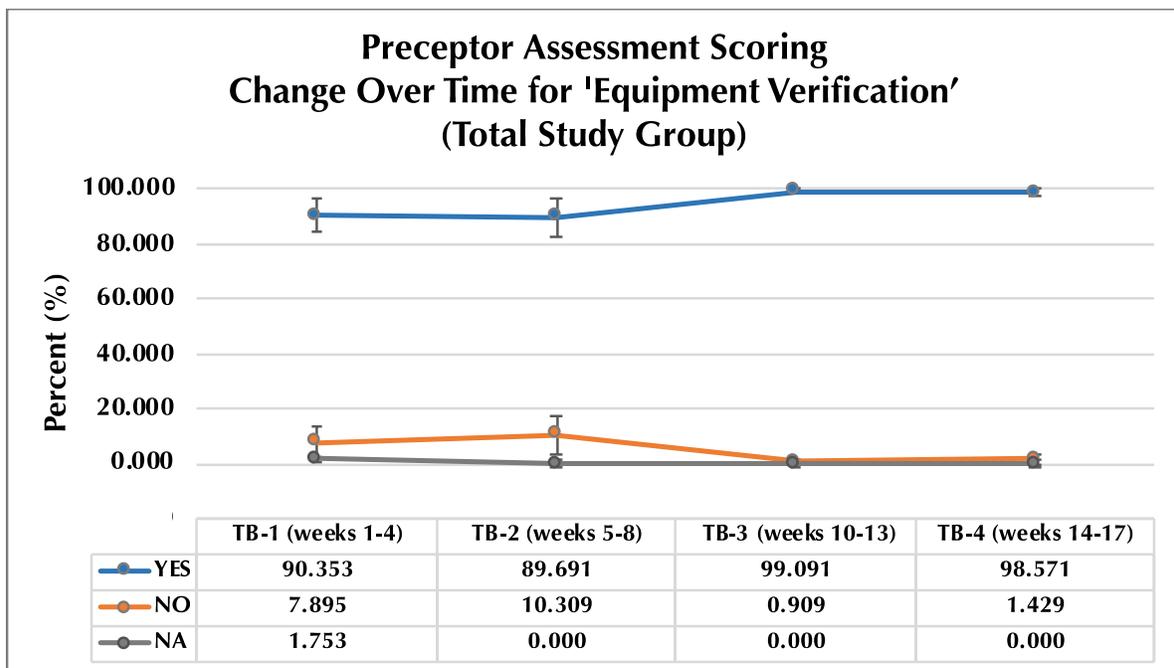


Figure 12 Change in Preceptor Scores Over Time: 'Equipment Verification' for the Total Study Group (mean ± SE)

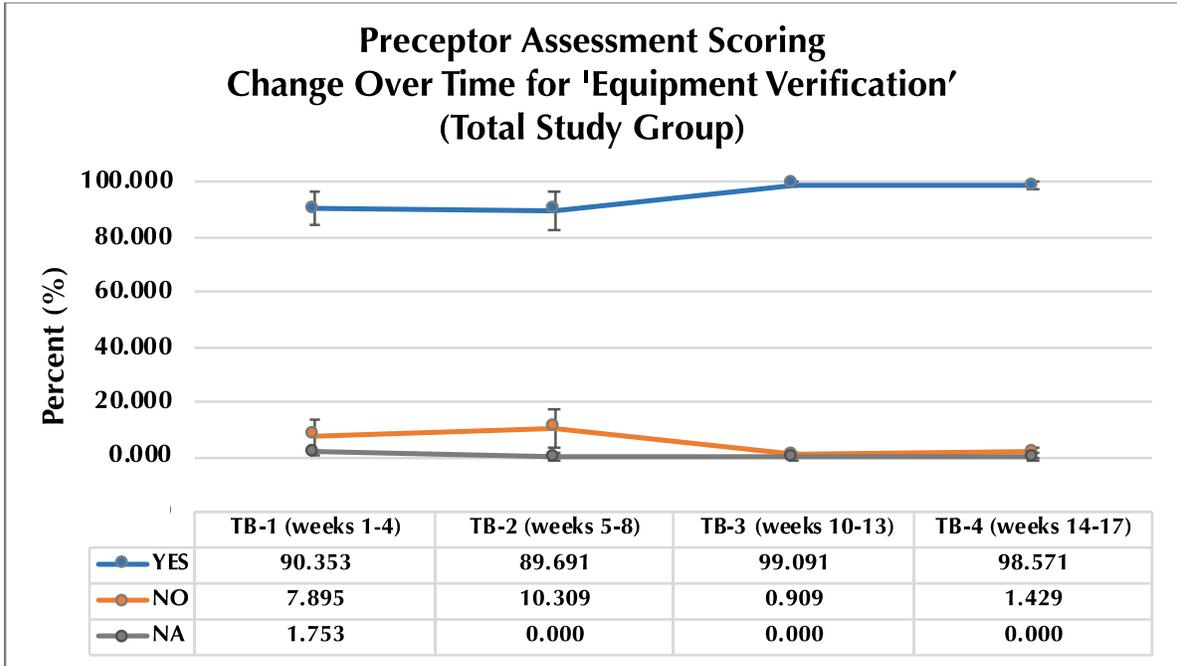


Figure 13 Change in Preceptor Scores Over Time: 'Apply Monitors' for the Total Study Group (mean ± SE)

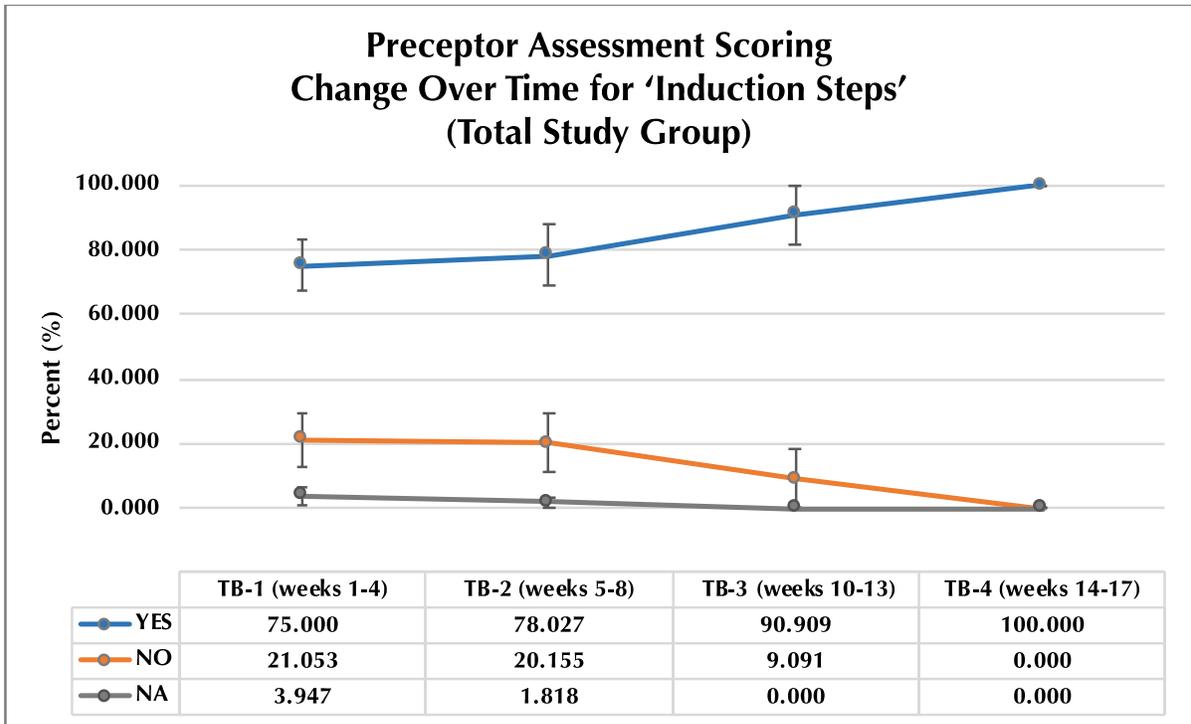


Figure 14 Change in Preceptor Scores Over Time: 'Induction Steps' for the Total Study Group (mean ± SE)

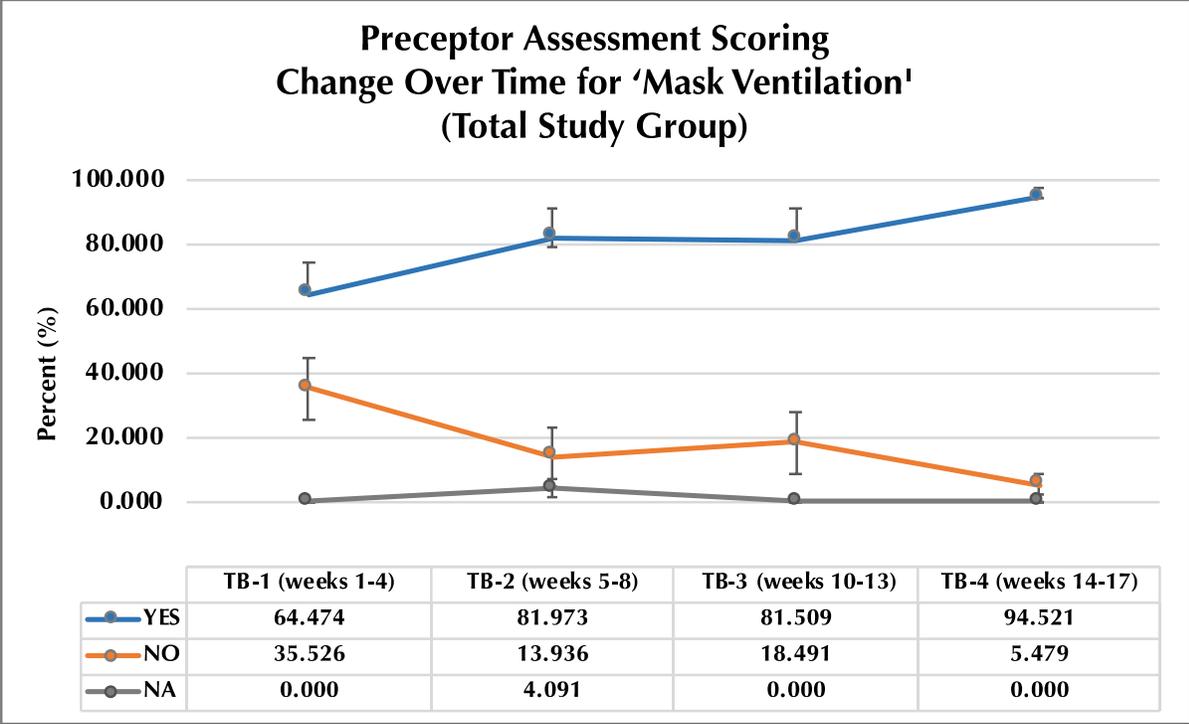


Figure 15 Change in Preceptor Scores Over Time: 'Mask Ventilation' for the Total Study Group (mean ± SE)

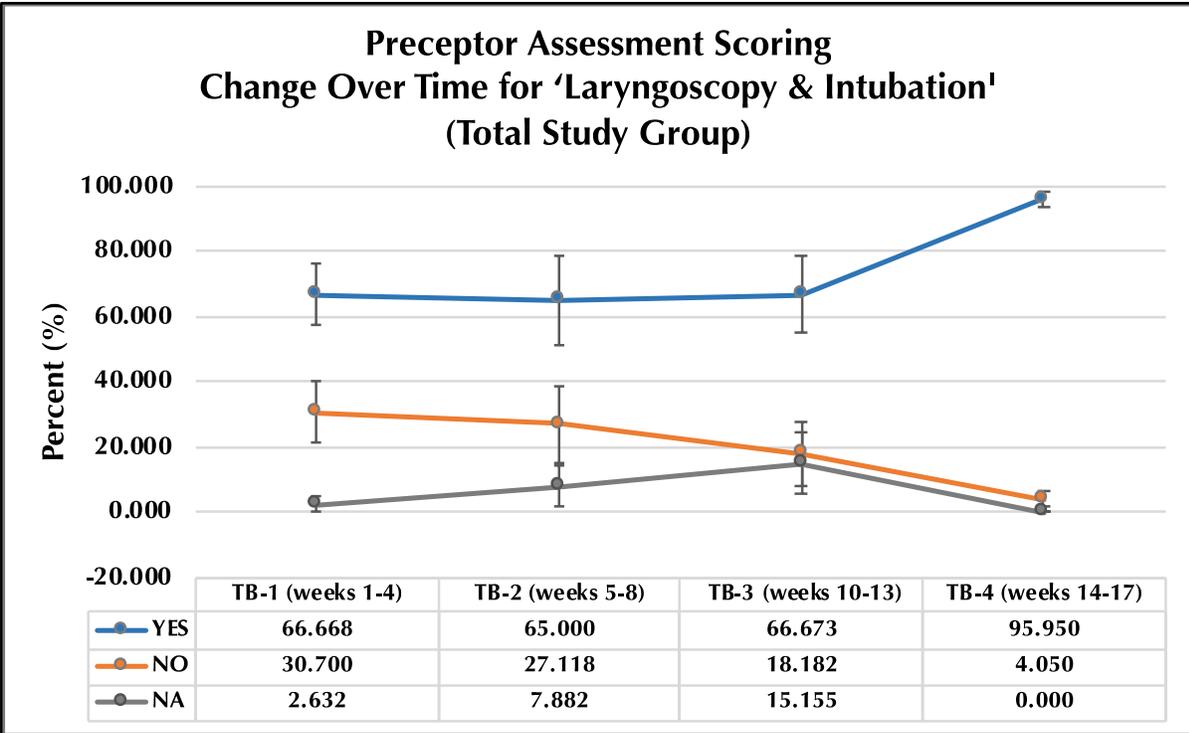


Figure 16 Change in Preceptor Scores Over Time: 'Laryngoscopy & Intubation' for the Total Study Group (mean ± SE)

While the Fixed Effects for ‘Group x Time’ was not statistically significant, there were some significant differences between the SIGA-O group and the SIGA+R groups for some preceptor scores at specific Time Blocks. For Question-2 ‘apply monitors’ the SIGA+R group had a higher proportion of ‘yes’ scores and a lower proportion of ‘no’ scores for Time Block-1 (**Figure 20**). For Question-4 ‘mask ventilation’ the SIGA+R group had a significantly lower proportion of ‘NA’ scores at Time Block-2 (**Figure 22**). For Question-5 ‘laryngoscopy and intubation’ the SIGA-O group had a significantly higher proportion of ‘yes’ scores and a significantly lower proportion of ‘no’ scores for Time Block-2 (**Figure 23**).

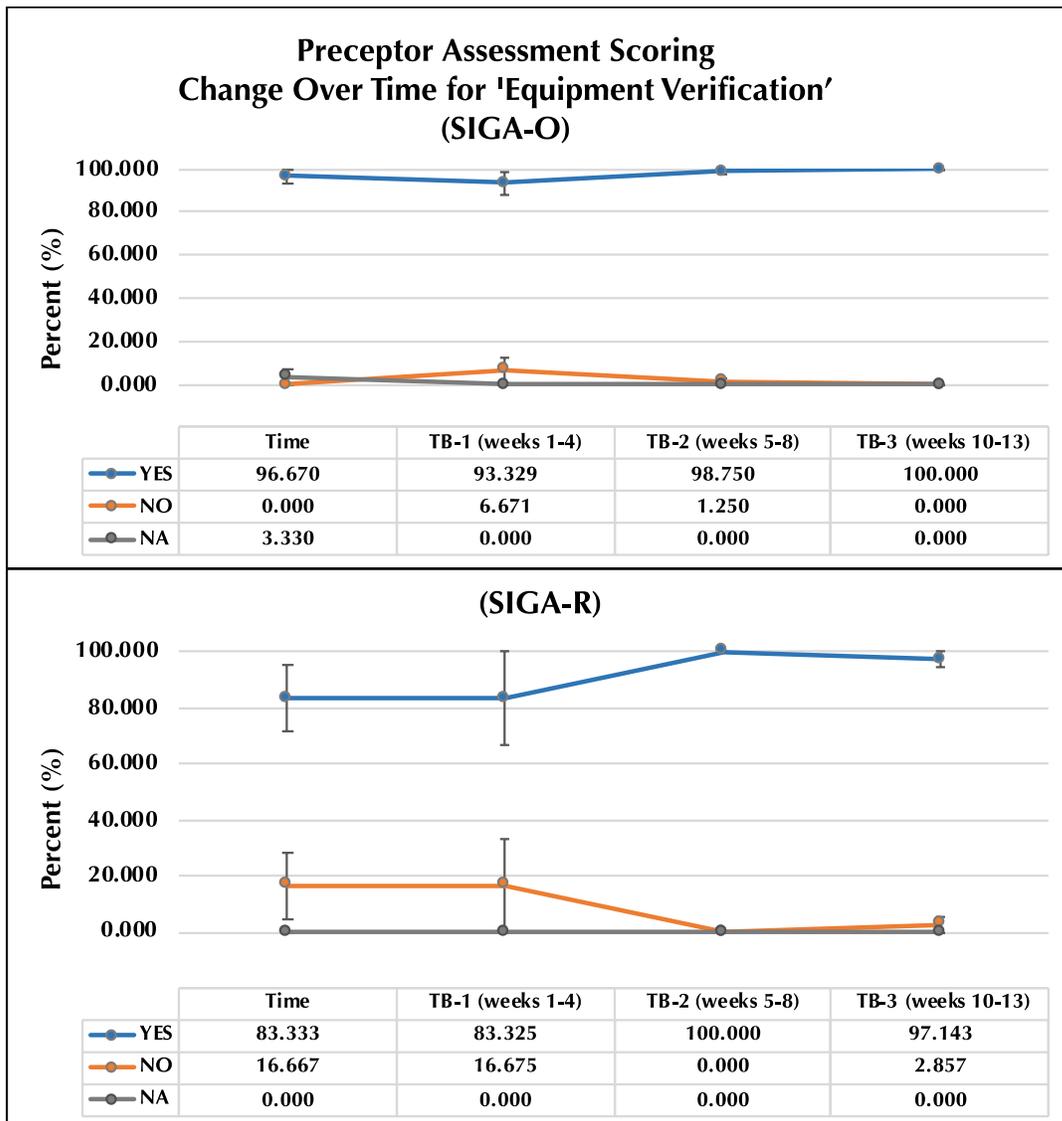


Figure 17 Change in Preceptor Scores: 'Equipment Verification' SIGA-O and SIGA+R groups (mean ± SE)

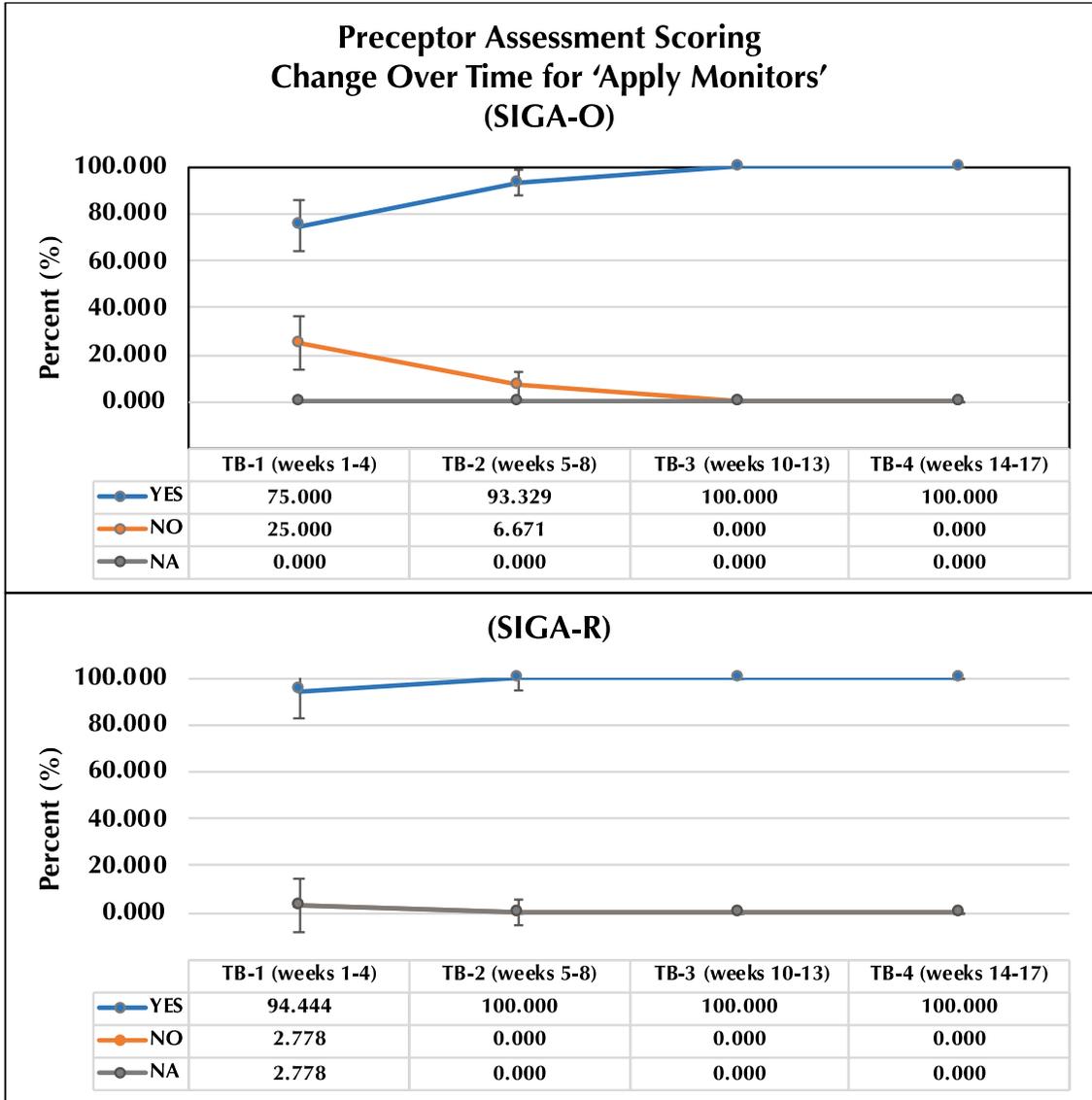


Figure 18 Change in Preceptor Scores: 'Apply Monitors' SIGA-O and SIGA+R groups (mean ± SE)

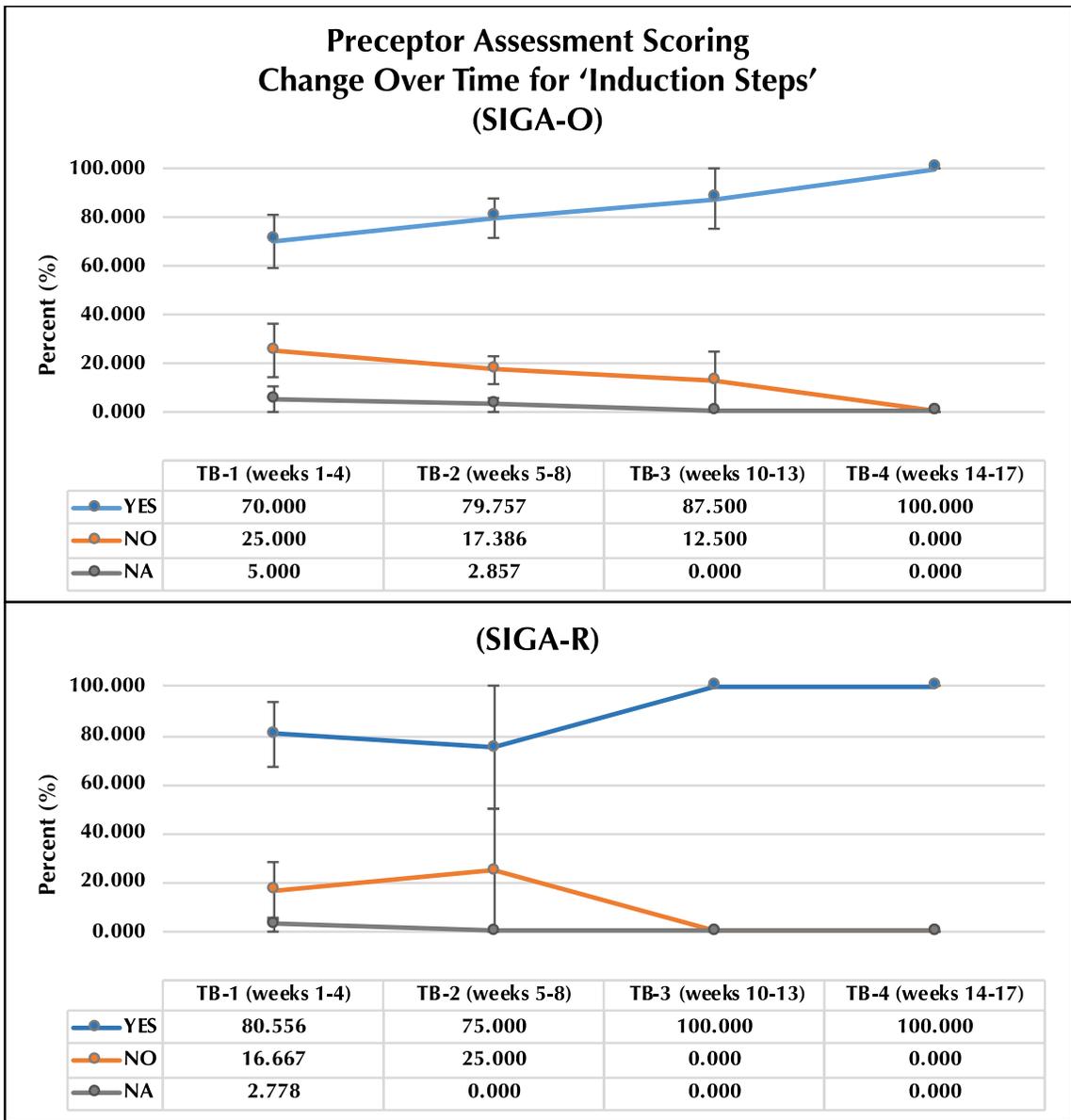


Figure 19 Change in Preceptor Scores: 'Induction Steps' SIGA-O and SIGA+R groups (mean ± SE)

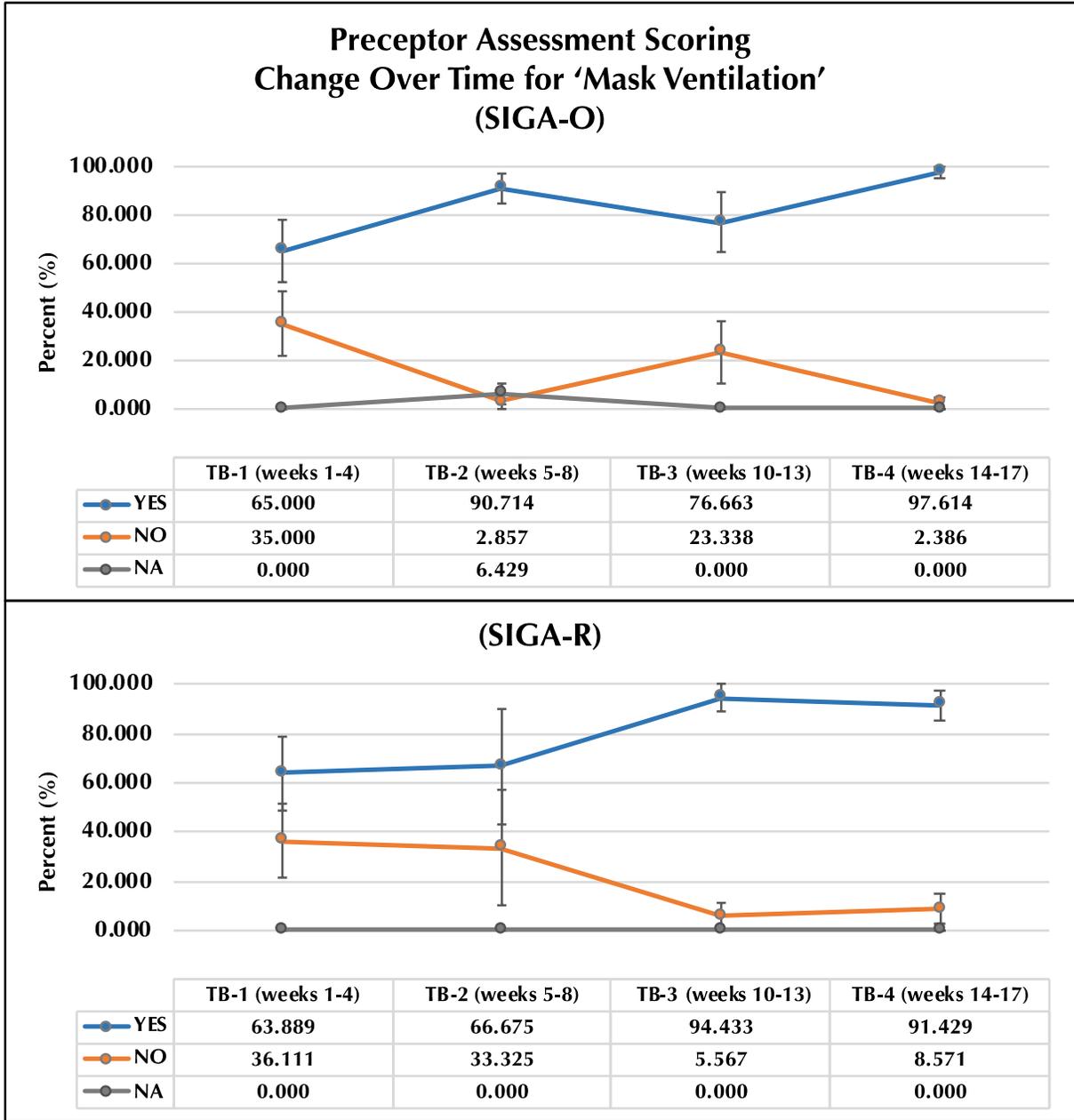


Figure 20 Change in Preceptor Scores: 'Mask Ventilation' SIGA-O and SIGA+R groups (mean ± SE)

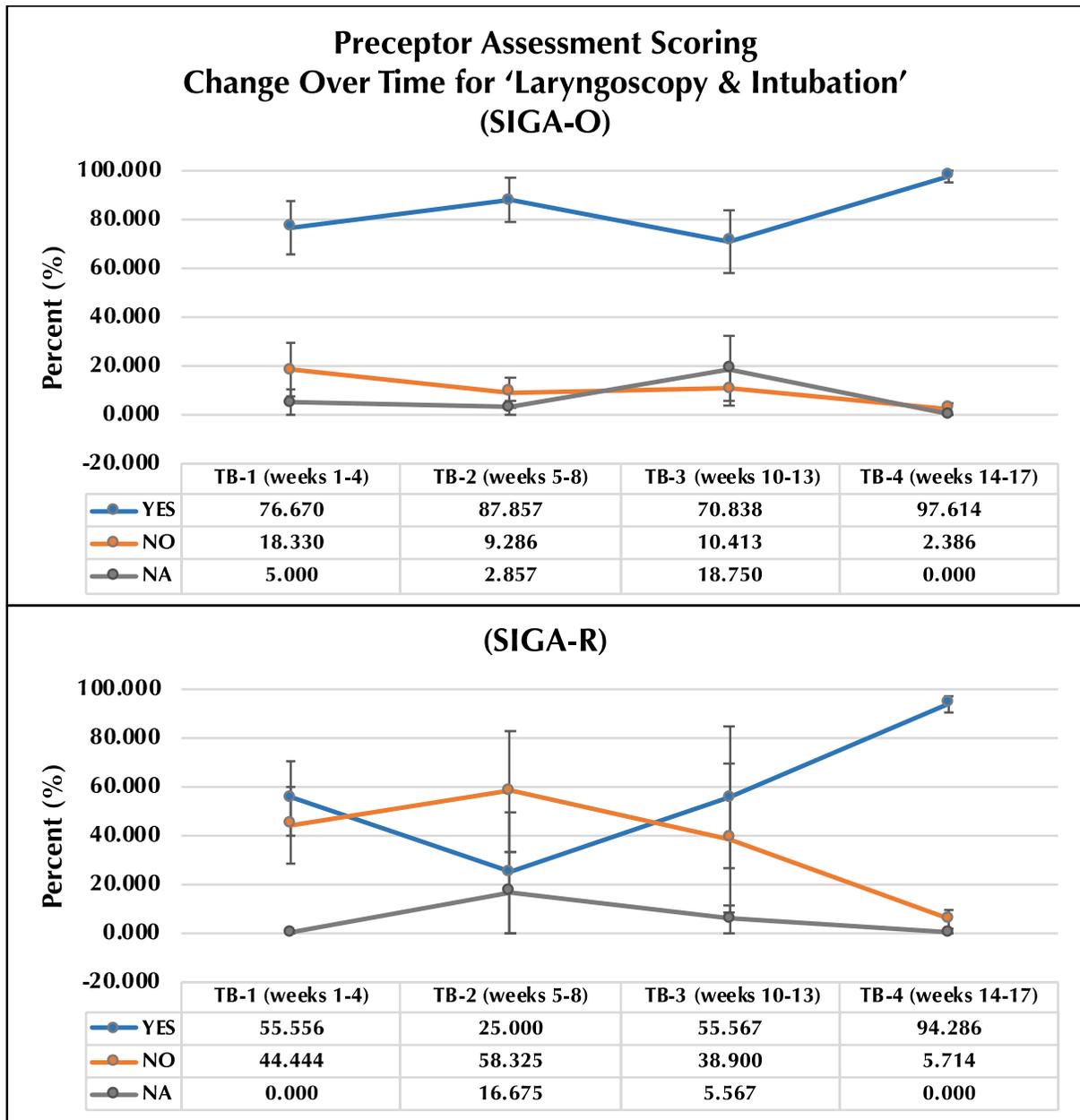


Figure 21 Change in Preceptor Scores: 'Laryngoscopy & Intubation' SIGA-O and SIGA+R groups (mean \pm SE)

4.2.6 SIGA-CAA Preceptor Overall Rating and SRNA Self-Evaluation

For the preceptor rating scale there were no significant group effects, but tests of Fixed Effects for ‘Group x Time’ were significant at Time Block 2 with the scores for the SIGA+R group being lower (mean ± SE: SIGA-O = 7.747 ± 0.261; SIGA+R = 5.500 ± 0.569). For the SRNA self-rating scale there were also no statistically significant group effects, but there were statistically significant ‘Group x Time’ effects at Time Block 3 with the SIGA+R scores again being lower (mean ± SE: SIGA-O = 7.354 ± 0.616; SIGA+R = 5.752 ± 1.095) (Table 12 and Figure 24). Of note, the correlations between the preceptor ratings of actual performance and the SRNA self-confidence ratings across the study demonstrated strong correlation.

Table 12 Linear Mixed Modeling Results for Preceptor Overall Performance Rating for Performing OTS-SIGA Steps and for SRNA Self-Rating of Confidence in Performing OTS-SIGA Steps

<u>O6 Repeated Variance-Covariance Type</u>						
	CS		AR1		TP	
	AIC	BIC	AIC	BIC	AIC	BIC
Preceptor Mean Score	174.486	178.186	174.673	178.373	177.656	185.056
group effects	none		none		none	
time effects	1-2 (<i>p</i> = .043)		1-2 (<i>p</i> = .044)		1-2 (<i>p</i> = .037)	
	1-3 (<i>p</i> = .000)		1-3 (<i>p</i> = .000)		1-3 (<i>p</i> = .000)	
	1-4 (<i>p</i> = .000)		1-4 (<i>p</i> = .000)		1-4 (<i>p</i> = .000)	
group*time effects	2-3 (<i>p</i> = .014)		2-3 (<i>p</i> = .015)		2-3 (<i>p</i> = .014)	
	2-4 (<i>p</i> = .000)		2-4 (<i>p</i> = .000)		2-4 (<i>p</i> = .000)	
	2 (<i>p</i> = .005)		2 (<i>p</i> = .005)		2 (<i>p</i> = .006)	
SRNA Mean Score	243.947	248.201	237.814	242.069	237.224	245.732
group effects	none		none		none	
time effects	1-3 (<i>p</i> = .001)		1-3 (<i>p</i> = .001)		1-3 (<i>p</i> = .002)	
	1-4 (<i>p</i> = .000)		1-4 (<i>p</i> = .000)		1-4 (<i>p</i> = .000)	
group*time effects	2-3 (<i>p</i> = .012)		2-3 (<i>p</i> = .003)		2-3 (<i>p</i> = .002)	
	2-4 (<i>p</i> = .000)		2-4 (<i>p</i> = .000)		2-4 (<i>p</i> = .000)	
	3-4 (<i>p</i> = .008)		3-4 (<i>p</i> = .006)		3-4 (<i>p</i> = .013)	
	3 (<i>p</i> = .046)		3 (<i>p</i> = .035)		3 (<i>p</i> = .033)	

Note: Significance indicated for Group effects (differences between SIGA-O and SIGA+R groups); Time effects (differences between time periods (1, 2, 3 or 4) for the entire study group); and Group x Time Effects (for the time block indicated). CS = compound symmetric; Autoregressive = AR1; TP = Toeplitz; AIC = Akaike’s Information Criterion; BIC = Schwarz’s Bayesian Criteria.

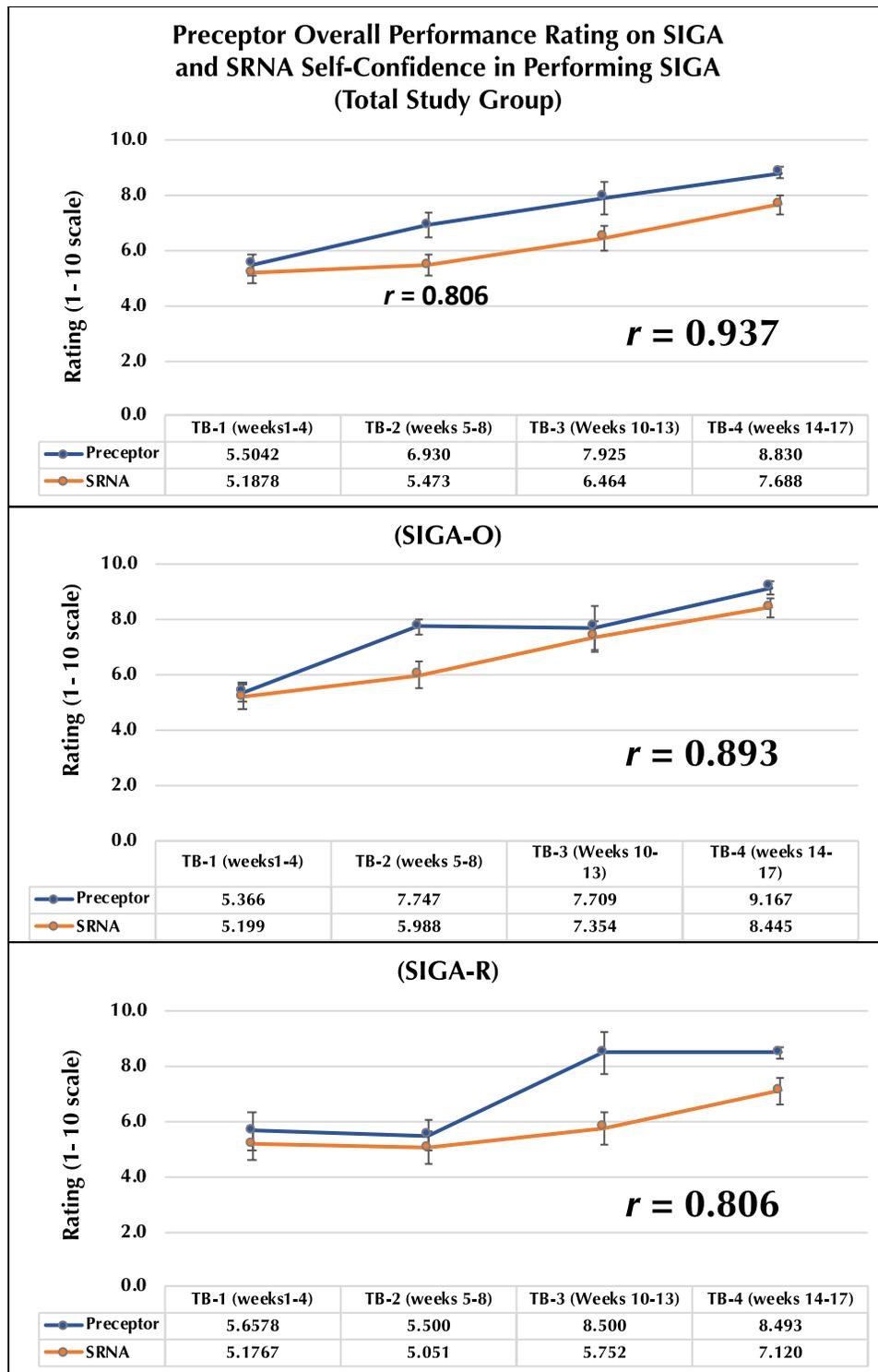


Figure 22 Change Over Time in Preceptor Overall Rating of SRNA Ability to Perform OTS-SIGA Steps Compared to SRNA Self-Rate Confidence in Performing OTS-SIGA Steps for the SIGA-O and SIGA+R Groups and the Total Study Group.

4.2.7 Preceptor to SRNA Airway Assessment Comparisons

For each patient under their care, anesthesia providers perform an airway assessment to help predict potential difficulties with mask ventilation and/or endotracheal intubation. Commonly, seven assessments are performed (oral opening and dentition, Mallampati classification, cervical range of motion, thyromental distance, mandibular length, palpability of the cricothyroid membrane and mobility of the mandible). SRNAs are expected to perform these assessments as well. Both SRNAs their clinical preceptors were asked to record their assessment data via the SIGA-CAA. These separate assessments were matched in the application database for later comparison. A total of 33 matched assessments were obtained. The SRNA assessments were coded as either in agreement or not in agreement with preceptor assessment. These data were aggregated and examined independently using linear mixed modeling in the same manner as previously described.

Tests of Fixed Effects for ‘Time’ and ‘Group x Time’ were significant for only one assessment, mandibular length. For the entire cohort, thyromental distance agreement had statistically significant changes between Time Blocks 1 and 3 (a decrease) and 3 and 4 (an increase). Mandibular length had significant changes between Time Blocks 1 and 2 (decrease), 2 and 3 (increase), and 2 and 4 (increase). Mandibular mobility assessment agreement showed a significant increase between Time Blocks 1 and 4 (**Figure 25**). Differences between the study cohorts (Group x Time effects) were present for thyromental distance at Time Block-3 with the SIGA+R cohort having a significantly lower success rate. Similar significant results were recorded for mandibular length assessment at Time Block-2 with the SIGA+R group again having less agreement with preceptors (**Figures 26 and 27**).

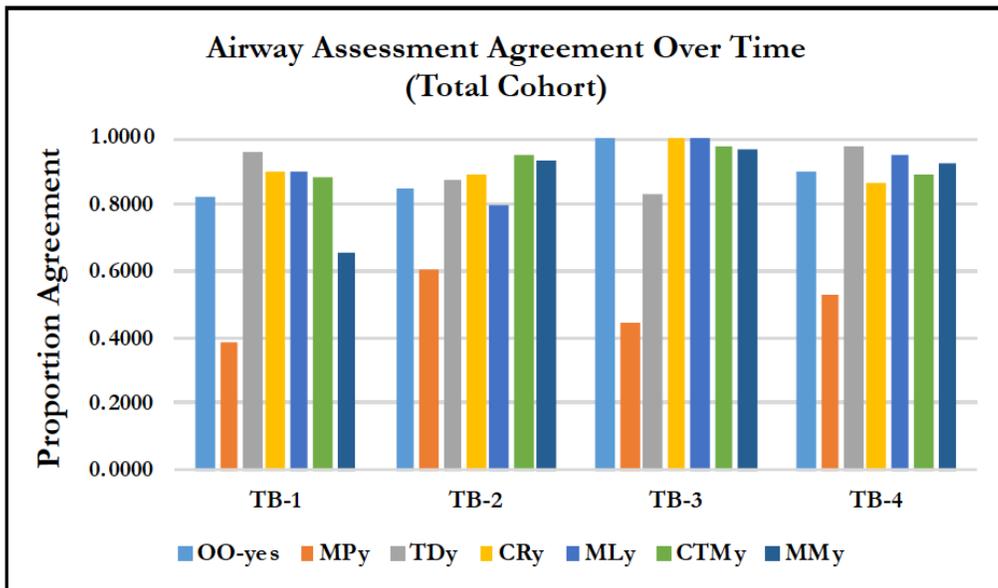


Figure 23 Airway Assessment Agreement over time for the Total Cohort.

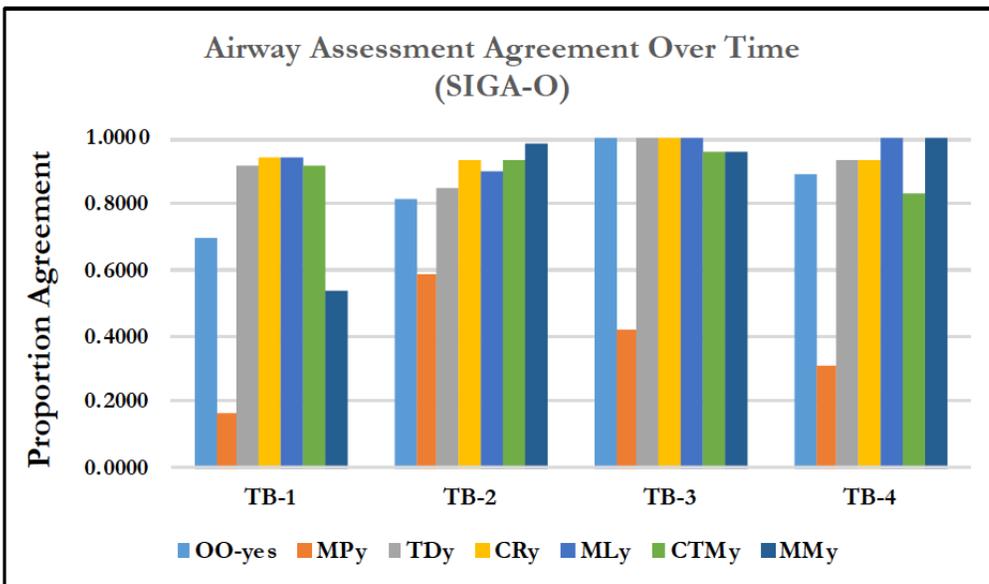


Figure 24 Airway Assessment Agreement Over Time for the SIGA-O Cohort.

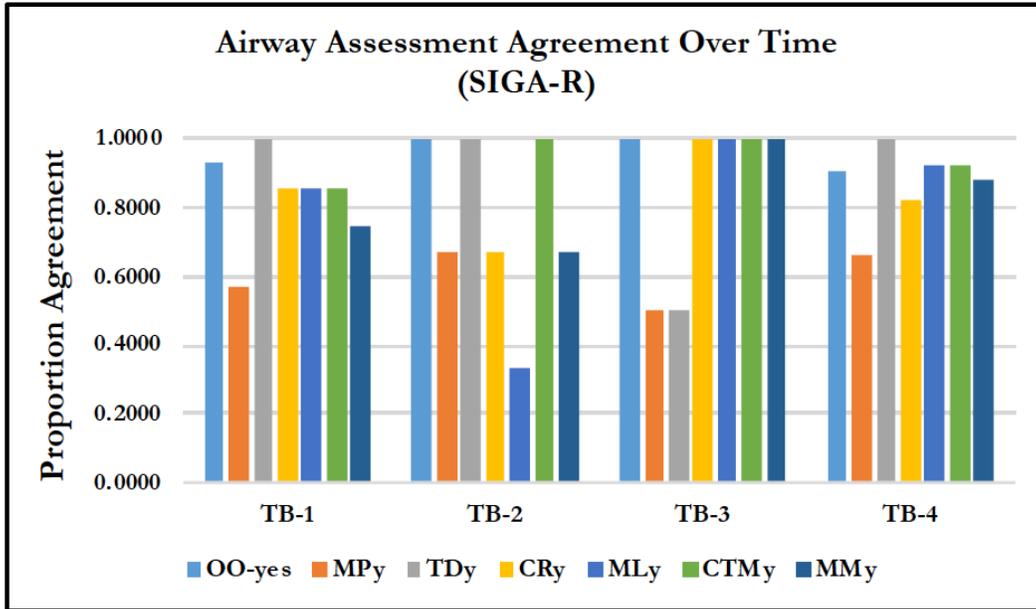


Figure 25 Airway Assessment Agreement Over Time for the SIGA+R Cohort.

4.3 Discussion

In this pilot study we attempted to measure the impact of a simulation intervention on first-term SRNA performance of the complex task of an induction of general anesthesia. Observations and measurements took place during their initial simulation event and then over the course of their 17-week initial clinical rotation. This included introducing a refresher Mock Induction simulation exercise at the mid-point of the rotation (nine weeks). Our general hypothesis was that all measured outcome markers would demonstrate improvement over the course of this first rotation. We also hypothesized that the cohort receiving the refresher Mock Induction (SIGA+R cohort) would demonstrate significantly better performance than the control cohort (SIGA-O). Our findings indicate that overall there were not any identifiable statistically significant differences between the study cohorts attributable to the refresher Mock Induction. Through linear mixed modeling we

attempted to assess the impact of age and clinical site as covariates, but this did not result in any meaningful differences in the analysis. Where there were isolated time points with statistically significant differences between the cohorts, these usually indicated better performance by the control group. Of note, however, this study revealed the overall readiness of the total study cohort for their first clinical rotation, with continued improved importance and increasing confidence. Prior to this, we had only anecdotal evidence to support the effectiveness of our simulation training in preparing SRNAs for their first clinical rotation. The specific outcome markers are discussed below.

4.3.1 Rater Training

We developed a rigorous rater training process. All raters and simulation teaching assistants were graduate students in the nurse anesthesia program who had been exposed to previous versions of the OTS-SIGA in their own training. All were trained the same, whether they ultimately scored live events, recorded events or both or acted as simulation teaching assistants. When determining whether a rater was able to move on to independent scoring of video events or live events, we required an overall agreement of 90% with the idealized score for a post-training test video. The post-training test video scores showed substantial agreement using Cohen's kappa for both the 10 primary subgoals (slight to perfect) and all 50 process steps (moderate to substantial). The Mock Induction event teaching assistants were drawn from this pool of trained raters. A single highly trained teaching assistant was used for all of the baseline Mock Induction training sessions. That the agreement for the fifty operational tasks was higher than for the 10 primary subgoals alone is not unexpected. Our operational definitions for scoring the primary subgoals required that if even a single operational task for a subgoal was rated anything other than

a 'yes' (completed), this could result in a different subgoal score (not completed, prompted, not available). The results shown here are for the first attempt at scoring a post-training test video by these raters. It should be noted that during training, raters were asked to score the videos they watched in 'real time' without pausing the playback. This was true for the post-training test videos as well. This rigorous threshold was necessary as raters would need to be able to score events live in the clinical setting where the process steps are often completed within less than five minutes. While rigorous and thorough, the training process took a substantial amount of time to complete. As discussed below, this impacted our ability to collect live observational data at the beginning of the study groups clinical rotation.

4.3.2 Demographic Profile of the Study Group

Overall, the total pool of potential participants and the final study group generally matched historical University of Pittsburgh Nurse Anesthesia Program mean demographic data for the five years prior to the study implementation. The one exception was age. There were two outliers in the total study group (older participants). Randomization placed these two participants in the SIGA+R group and this was the only substantial differences in demographic variables between the two study groups. We performed Linear Mixed Modeling taking age into account as a co-factor and found no significant differences in our results. Additional LMM analyses were run using clinical site as a covariate, again with no significant impact on the results.

4.3.3 Knowledge Retention: SIGA-KAI

The Standard Induction of General Anesthesia Knowledge Assessment Instrument (SIGA-KAI) was administered at multiple time points over the course of the study: pre- and post-simulation, during the second and ninth weeks of the clinical rotation and during the last (seventeenth) week of the rotation. All participants completed the assessment at all time points. In general the distribution of scores was truncated as the scores were relatively high (80 – 93 percent on a 0-100% scale). There were no significant differences between the study groups in performance on the knowledge assessment at any time point. One unexpected result was the global decrease in scores from Post-Simulation (93.54 ± 4.03) to Clinical Week-2 (80.63 ± 6.81). Scores did increase at Clinical Week-9 but decreased again by the end of the rotation. This was surprising because Clinical Week-2 occurred approximately three weeks after the initial Mock Induction exercise. Immediate post-simulation scores were the highest, and never returned to this level at any subsequent time point. It is possible that instruction by clinical preceptors could have created confusion for SRNAs about what the appropriate answer for some questions should be. As noted in the preliminary studies section, our HTA analysis revealed that some steps occurred at variable points during the induction process. This variability was related to several factors. For example, some process steps did not have been supporting evidence in the literature for a definitive time to do the task, which led to CRNAs and anesthesiologists drawing on personal experience as their guide. Some questions in the SIGA-KAI were related to these types of steps. However, for standardization of training we needed to position process steps at consistent points in the HTA process map. This conflict between the simulation training and clinical preceptor direction of the SRNAs in the study may have resulted in the performance decreases. For example, performance on questions related to Step 4 (Pre-Induction Preparation) particularly

demonstrated this at Clinical Week-2, with all but one of the study group participants missing one of the questions related to this step. Scores at clinical week 2 were the lowest of any of the time points.

Additionally, timing may have impacted the results of the SIGA-KAI at Clinical Week-2. The study participants were just beginning their clinical rotations and were also preparing for final examinations for the academic term. Participants were also being asked to complete multiple tools and assessments over the course of the study; fatigue with regard to study tasks versus on-going clinical and classroom responsibilities may have been a factor.

4.3.4 OTS-SIGA Live Scoring and Clinical Preceptor Scoring

Obtaining the live scoring data proved to be logistically challenging. Study participants were assigned to eight different clinical sites across the Pittsburgh region and were not always at their assigned clinical sites on the same days each week, The types of operative cases they were assigned to also varied for each clinical day. This made predicting when a particular participant might be doing an induction for general anesthesia difficult. As discussed above, the rater training process took longer than expected, and data collection needed to begin before all of the raters had been fully trained. The impact was such that we were not able to collect any live observational data during Time Block-2 (clinical weeks 5 to 8).

The data presented here represent the percent of all 10 primary subgoals that were scored as yes (Completed), no (not completed), prompted or not available. While there was no significant impact of introducing a refresher Mock Induction at the mid-point of the clinical rotation (no significant differences between the SIGA-O and SIGA+R groups), there were still important findings. To begin with, to our knowledge no one has previously reported data related to SRNA

performance in the complex task of induction of anesthesia before. Additionally, this is the first time that we have been able to quantify our previous anecdotal assumption of steady performance improvements over the course of the first clinical rotation. There was a significant effect of time for all study participants. Statistically significant increases in the proportion of subgoals scored as ‘yes’ (the SRNA completed the step) were demonstrated between Time Blocks 1 and 3 and Time Blocks 1 and 4 for the study cohort as a whole. The proportion of subgoal steps ‘prompted’ and the proportion scored as ‘not available’ to the SRNA decreased for those same time intervals. Also of interest is the steady increase in the proportion of ‘no’ scores (meaning that these SRNAs failed to successfully complete a subgoal) across the clinical rotation. These rose from approximately 18% to 29%. These data suggest that over the course of the rotation clinical preceptors felt increasingly comfortable allowing the SRNAs to *attempt* process steps. With more opportunities, SRNAs completed more process steps and did so successfully and independently (fewer prompts). This data, while limited, begins to provide a more complete picture of how clinical learning may be taking place. As we might intuit, learning may be directly related to the number of opportunities presented to perform a task independently. At first, many of the SRNA attempts at a subgoal step might not be completed or not completed correctly, but the proportion of correctly completed steps increased along with the unsuccessful ones.

The clinical preceptors did not score all ten HTA process primary subgoals, but were asked to score five of the HTA derived subgoals considered most critical by the investigators. The refresher Mock Induction did not demonstrate any significant Group x Time effects but there were significant effects of time for the whole study group. The scoring done by clinical preceptors generally followed the same pattern as the live scoring done by trained raters in regard to the proportion of steps completed successfully. There was a clear general trend for all

five preceptor scored subgoals with ‘yes’ (completed) scores increasing over the course of the clinical rotation and ‘no’ or ‘not available’ scores decreasing.

There were notable differences in preceptor scoring and live rater scoring. Clinical preceptors generally scored SRNA performance much more favorably than the trained raters. While the clinical preceptors did not have as rigorous of an orientation process as the trained raters, it was clear that the clinical preceptors, who work with SRNAs on a regular basis, felt that the study group participants were highly prepared prior to beginning their clinical rotation. The proportion of steps rated as successfully completed was considerably higher for the preceptors. The *highest* proportion of ‘yes’ scores at any time block from the trained rater data was 20.8%; the *lowest* proportion from the preceptor scoring was 64.5%. Time Block-1 ‘yes’ scores ranged from 64.5% to 90.4% across the five steps preceptors were asked to evaluate (**Figures 14 – 18**). There are several possible reasons for this. First, the clinical preceptors were instructed how to score using the SIGA-CAA tool and were given operational definitions, however they did not go through the intensive rating and feedback process that the trained raters did. Secondly, the clinical preceptors were not given the option to indicate whether a step was prompted. During rater training it became apparent that prompts could be subtle and difficult to identify. The study team met regularly during the rater training process to discuss the types of verbal and non-verbal prompting that were being identified in training videos. Because of this, the decision was made to keep the scoring as simple as possible for the clinical preceptors and eliminate this option. This could have contributed to the much higher rate of steps scored as successfully completed as a completed subgoal might have been considered successfully done (‘yes’ score) even if it was prompted. And finally, despite being instructed that their ratings of student performance for this study would not

be used in the clinical and academic evaluation of the SRNAs in the research protocol, it is possible that the clinical preceptors were just reluctant to give a ‘bad’ score to a student.

The more cognitively complex and more difficult psychomotor subgoals (Induction Steps, Mask Ventilation and Laryngoscopy and Intubation) overall had a lower percentage of preceptor ‘yes’ scores than the other subgoals. These are processes that first rotation SRNAs generally struggle with and it has been reported that provider experience may significantly impact the four-point Cormack rating of the view of the larynx. (Benumof, 1996) However, these steps also demonstrated statistically significant increases over time in the percentage of ‘yes’ scores from Time Block-1 to Time Block-4. Concurrently, no responses decreased, and NA responses remained relatively constant. This appears to demonstrate that preceptors consistently allowed SRNAs to perform these tasks during the first clinical rotation, and that the continued repetition of these steps allowed for improved performance.

4.3.5 Clinical Preceptor Overall Rating and SRNA Rating of Self-Confidence

There were no overall statistically detectable differences between the two study groups, indicating that the refresher Mock Induction did not have an impact on these outcome measures. There was only one time point (Time Block-2) where preceptor ratings demonstrated significant Group x Time effects. At this time point, the preceptor scores for the SIGA+R group were lower. SRNA scores of self confidence in performing the OTS-SIGA steps demonstrated significant differences between groups at Time Block-3. For the total study group, preceptor overall ratings of SRNA performance in completing the OTS-SIGA process steps increased across the entire study period. There were statistically significant increases in preceptor ratings between all Time Blocks except between Time Blocks-3 and 4. SRNA self-confidence scores increased across the entire

study period. There were statistically significant differences in self-ratings between all Time Blocks except for between 1 and 2. These trends held when the results for the SIGA-O and SIGA+R cohorts were examined separately. More interesting were the correlations between preceptor overall ratings of SRNA performance on OTS-SIGA steps and SRNA self-confidence ratings for performing those steps (**Figure 24**). There were differences in this correlation between the two study groups. The SIGA+R group in general did not feel as confident as the SIGA-O group after Time Block-1. Outside of Time Block-2, their perception of their own performance did not match that of their clinical preceptors as well as did the SIGA-O group (**Figure 29**); these differences did reach statistical significance at Time Block-3. The correlation for the total cohort was 0.937. This is exceptionally high, and we believe is perhaps the first reported evidence of self-confidence versus actual performance reported in relation to a simulation intervention.

When looking at these ratings over time, it is important to keep in mind that the preceptors and the SRNAs were being asked two different, but related, questions. Preceptors were asked to provide an overall rating of the SRNA's performance in completing the process steps for a standard induction of general anesthesia. SRNAs in the study were asked to provide a rating of their confidence in performing these steps. While both were rating on a 10-point scale, the fact that the questions were different made it inappropriate to analyze the differences in score over time. Qualitatively, however, we can infer that for the SRNAs performing the complex task of induction of anesthesia, success breeds confidence and in turn reinforces further success.

4.3.6 Agreement with Airway Assessment

Commonly, seven assessments of the patients are performed when presenting for anesthesia services (oral opening and dentition, Mallampati classification, cervical range of

motion, thyromental distance, mandibular length, palpability of the cricothyroid membrane and mobility of the mandible). Both clinical preceptors and SRNAs were asked to record their airway assessments for general anesthesia cases they participated in together. While there were statistically significant differences over time for agreement between preceptor and SRNA assessments of the airway, these were generally of an inconsistent nature (e.g. increase, then decrease then increase again across Time Blocks). There were no significant differences between the study cohorts. No readily discernable patterns were noted with the exception that one assessment had consistently lower agreement with the clinical preceptors across time – Mallampati scoring. One explanation may be that all of the other assessments, with the exception of mandibular mobility, can be quantitatively measured (though this is not routinely done in the clinical setting). Mallampati Scoring, while guided by anatomical landmarks, is a relatively subjective process that can be greatly impacted by a variety of factors (e.g. is the patient supine or upright during the assessment, are they told to phonate, do they properly follow instructions and extend the tongue as far as possible). Mallampati is known to have high specificity and much lower sensitivity in terms of identifying a potential difficult airway.

4.3.7 Oral Soft Tissue and Dental Injury

There were only two reports of dental or oral soft tissue injury during the period of the study protocol. Both of these came from the SIGA-O study group, but this number is too low to infer any significant differences between the study groups.

4.4 Limitations

4.4.1 Methodological Limitations

Limitations were primarily related to technical issues with the data collection application, the length of the rater training process and the differences between the simulation lab setting and the perioperative environment where live clinical scoring was done.

The OTS-SIGA application was designed to be accessed via any mobile platform. By the end of the training process, raters were very comfortable with interacting with the interface. However, there were occasions in the early phase of data collection where the application ‘froze’. This was almost always due to poor connectivity to the internet, and internet ‘dead spots’ varied from clinical site to clinical site. This did not prevent actual scoring of events, but did prevent data uploading after scoring was complete. This resulted in raters pressing the submit button multiple times. The usual outcome was that the results were eventually transmitted several times when the device eventually came in range of a good signal. These duplicate entries then were reviewed and deleted from the final database using a series of screening rules developed by the research team. Raters quickly adapted to this issue by delaying pressing the ‘Submit’ button until they were in an area with good connectivity.

The rater training process was thorough, but took significant time for completion. After each practice scoring session, raters in training submitted their scores through the OTS-SIGA application. Data could not be directly accessed by the investigators; data were released by an honest broker from WISER. While the turnaround time was not significant, there were lags between rater practice scoring, investigator access to the scoring and principal investigator review and feedback. Additionally, as per IRB guidelines for video source data, reviewers could only

access the video files in the project offices on a computer with two-step password encryption. The result was that it took much longer to train raters than initially anticipated, and availability of trained raters impacted the ability to collect live data in the clinical setting at some time points.

Finally, as discussed earlier, there were some differences in how and when steps were done in the simulated Mock Induction sessions versus the clinical setting. For example, in the simulation setting, equipment preparation and patient interviewing occurred in the same room. In the perioperative setting, anesthesia providers prepare their equipment in the operating room, then locate their scheduled patient in a separate preoperative holding area to conduct the interviews and assessment. This may have impacted the new SRNAs as they adjusted to the change in locations and environment for the initial steps of the process. It certainly presented logistical challenges for raters as they ideally needed to be in the operating room at the time of equipment checking and preparation and then move on to the preoperative holding area. It was far easier to predict the location and timing of these events for first cases of the day than for subsequent cases. There was also variability in the sequence of some events between the OTS-SIGA and the clinical setting. One step in particular provides a good example of this. Anesthesia providers all agree that the patient's eyes must be protected by taping them closed (or using an alternative method such as goggles), but during the HTA process it became clear that there was much disagreement about when this step should occur. Our OTS-SIGA process map placed this step after induction (Step 5) but before mask ventilation (Step 6). Many providers believe that mask seal to the patient's face should never be interrupted, and thus they insist on this variation from our process map when precepting students. We emphasized to our trainees there is variability around this timing, but that we were choosing to teach everyone in a consistent manner. SRNAs reported that these types of

process variations from preceptor to preceptor affected their performance adjustment to the real-world induction process.

4.4.2 Pilot Study Data Interpretation Limitations

There are several limitations to this pilot study that are important to discuss. First, the sample size was small, with only 12 members of each cohort, and as described above, effect sizes would need to be large to detect significant differences between cohorts. This was a complex study with multiple assessments over an extended period of time. As discussed earlier, while the rater training was robust, it took time to complete, limiting the number of trained raters that could be deployed at certain time points in the study. There were also timing and logistical barriers in attempting to observe 24 student anesthetists across multiple clinical sites. Start times for first operative cases of the day are scheduled, but subsequent case start times can be difficult to predict. To offset these data collection difficulties, our study design incorporated the use of CRNA preceptors who work daily with SRNAs in the clinical setting. As routine instructors of the induction process, they were well positioned to provide good assessments of SRNA performance. A limiting factor was production pressure in the perioperative setting. Anecdotally many of the clinical preceptors reported that they felt that they did not have time in the normal flow of a clinical day to enter the assessment data, and they would often forget to do so at the end of the day or forget the particulars of the case. Some preceptors were uncomfortable entering data through the web-based application. We attempted to offset this reluctance by providing paper versions of the on-line assessments. SRNA self-report of airway assessments and self-confidence scores was also inconsistent across study participants and across time. Submissions generally decreased during Time Blocks-3 and 4. Data such as the airway assessment required having a matched pair of

assessments from the preceptor and the SRNA. There were many more assessments without a match than matched pairs and the numbers for this outcome measure, while potentially large ending up being quite small.

Aside from these logistical considerations, there are two other factors that may have significantly affected our outcomes. First, while decay of knowledge and skills has been well established in the literature, only in the last decade have significant efforts been made to offset these decays utilizing simulation. There is a large body of literature developing regarding booster training and ‘just-in-time’ training, especially as related to ACLS and BLS skills. (Barsuk, McGaghie, Cohen, Balachandran, et al., 2009; Niles et al., 2009; Sutton et al., 2011a, 2011b) What is not clear is the optimum timing for booster training, especially for more complex task set such as the one studied here. We selected the midpoint for a variety of reasons, including considering the known progression of access to and participation in clinical experiences. At the beginning of their first rotation, SRNAs in the University of Pittsburgh Nurse anesthesia program are only in the clinical setting two days per week. The beginning of their first clinical rotation also coincides with major holidays and final exams, during which time students often opt to use vacation time. Our intent was to place the refresher at a point where these learners would have had enough clinical time to allow appropriate referencing of those experience to the repeat Mock Induction. It is possible that we did not choose this time point correctly.

The second factor that may have impacted our results is the power of clinical learning itself relative to the 2-hour Refresher Mock Induction. While our preparation of SRNAs for the clinical setting is intense, there are some things that we cannot replicate in the simulated setting that are important in the operative setting. Two examples of this are mask ventilation and laryngoscopy and intubation. The mannequins used for training very good for teaching process steps and basic

techniques. They are not very good at presenting variables that make these psychomotor tasks difficult in the operative setting (e.g. tongue size and ability to displace the tongue, variations in dentition, presence of a beard). It is worth noting mask ventilation and laryngoscopy and intubation were the process steps preceptors most frequently identified as ones SRNAs in this study cohort struggled with. Given all of this, it may be that for this particular complex task set, the impact of the clinical learning was so large as to make any effect of the simulation intervention non-detectable.

5.0 Conclusions and Summary

5.1 Conclusions

A key goal for the emerging science of healthcare simulation is to demonstrate efficacy, primarily in terms of positive changes to clinical practice and improvements in patient outcomes. Strong evidence has been accumulating with regard to achieving these goals for specific cognitive and/or psychomotor tasks. (Barsuk et al., 2010; Barsuk et al., 2015; JM. O' Donnell et al., 2007) What has yet to be demonstrated is the ability to generate similar evidence for complex tasks taking place in a dynamic setting. This has been tangentially addressed by measurement of secondary outcome markers for situations such as obstetric emergencies. Use of crisis response teams have demonstrated improvements in multiple clinical markers. (Crofts et al., 2008; Draycott et al., 2008; Siassakos et al., 2009) However, few if any attempts have been made to establish a direct measurable linkage between performance in the simulation setting and performance or self-confidence in the clinical setting. That is, are these teams doing the same tasks at the same level of proficiency in the real-world setting as in the simulation setting, are they confident in their clinical performance and can these be linked to improved outcomes? Part of the difficulty in generating such data is the complexity of the processes which are the target of interest. This complexity has two components. First, understanding all of the component parts to a task; and second, being able to capture performance data for healthcare providers performing those tasks (either in the simulation setting or the patient care setting). Through previous work, we have demonstrated that Hierarchical Task Analysis can be a powerful tool for answering such questions. HTA has a long established utility in describing complex systems across a variety of domains.

(Annett et al., 2000; Shepherd, 1998) We believe that we have been among the first investigators (if not the first) to leverage HTA in the design and evaluation of simulation educational interventions. (O' Donnell J et al., 2007; O' Donnell, Goode, et al., 2012; JM. O' Donnell et al., 2007; JM O' Donnell & Goode, 2008) O'Donnell et al leveraged HTA to demonstrate parallel improvements in adherence to the process steps of an ergonomically sound patient transfer protocol in the simulation and patient care environments. (JM. O' Donnell et al., 2007) What became apparent from this work was the power of HTA to structure and guide multiple phases of the simulation education continuum. The process maps derived from an HTA analysis can be used to help design educational materials and the structure of the simulation intervention itself. (J. S. Goode Jr & O'Donnell, 2016; JM O' Donnell & Goode, 2008) In this current study, development of the OTS-SIGA impacted multiple aspects of the delivery of educational materials regarding the induction process. A detailed mapping of the important steps of the process allowed us to focus on these elements in training. SRNAs were given an in-depth introduction to the details of each step of the OTS-SIGA in the classroom. Partial task simulations (e.g., basic airway management workshops) allowed for focused practice of individual steps. Practice sessions allowed for exploration of and familiarization with the OTS-SIGA in an individualized manner. Every aspect of the Mock Induction was guided by the process steps, including post-simulation debriefings. One other advantage is that leveraging this tool can provide for consistency among faculty instructors and guide both horizontal and vertical integration of simulation components into the larger curriculum.

The rigor of the iterative HTA process resulted in a high degree of certainty that classroom content and the simulation training mapped well onto the actual real-world process of a standard induction of general anesthesia, though, as discussed earlier there were some process steps that

demonstrated variability in when they occurred and had no strong support in the literature for a definitive placement in the HTA process map. However, from a research perspective, this meant that we could be relatively confident that we would be comparing ‘apples-to-apples’ when comparing SRNA performance in the simulation lab versus the clinical setting. Additionally, using the OTS-SIGA to inform the development of our web based applications for data collection meant that we would also be using the same evaluation tools across observational settings. The work of other researchers has also reinforced our belief that we could be confident in our description of the induction process. Phipps et al performed a similar task analysis for anesthesia processes. (Phipps, Meakin, Beatty, Nsoedo, & Parker, 2008) Their work encompassed the whole of the anesthetic experience from pre-operative assessment to transfer to the recovery room. However, the steps described with regard to induction include many of our own. The level of descriptive detail of our OTS-SIGA is significantly less than that of Phipps. There are two main reasons for this. First, our process map is intended for the training of novice SRNAs. Mock Induction is just one part of the previously described intensive two months of preparation that include sixty hours of classroom content and forty hours of simulation. The intent in this circumstance is to provide a ‘boot-camp’ to achieve a minimum competency that will allow for the safe entry of SRNAs to clinical practice. Phipps et al intended to achieve granular detail with their HTA, creating “a framework for promoting good practice and highlight areas of concern.” Additionally, they used the Systemic Human Error Reduction and Prediction Approach (SHERPA) to examine portions of the process at risk for human error genesis. Other differences with our HTA product may be related to differences in clinical practice between the United Kingdom and the United States. For example, Phipps describes a process where induction takes place away from the operating room; for the most part inductions in the U.S. take place in the operating room.

One other notable aspect of this study was the ability to examine the combined role of simulation and increasing clinical experience on knowledge retention. We could do so because we repeatedly administered an assessment tool across the span of the SRNAs first clinical rotation and interjected a booster simulation for one cohort. Educators often speak of the difference between ‘knowing and doing’. Our study design and our adaptable data collection tools provided the opportunity to begin to examine changes in knowledge over time mapped over changes in clinical performance. In this case, we surmised that anesthesia preceptors in the clinical setting provided information that conflicted with the HTA process map for some sub-goal process steps. This impacted the results obtained from our knowledge assessment tool.

Similar opportunities were presented in regard to self-efficacy. An accepted assumption has been that as clinical experience increases, so does self-efficacy. It was unclear however if repeated exposure to simulation exercises would augment this and our approach in concurrently assessing SRNA self-perception of performance and preceptor evaluations of performance provided some initial answers. There was no significant impact on self-confidence in performing the OTS-SIGA process steps with regard to the Refresher Mock Induction. What was clear was that for the study group as a whole, preceptor evaluation of performance correlated very highly with SRNA self-confidence ratings. Preceptors indicated with their scoring that our training process prior to the start of the SRNAs clinical rotation prepared them to function at very high level from the beginning of their clinical experiences. As opportunities to perform clinical tasks increased, success in performing those tasks increased, resulting in a concurrent increase in self-confidence.

5.2 Summary

The purpose of this pilot study was to assess the potential impact of a simulation intervention on the subsequent clinical performance of first term Student Registered Nurse Anesthetists during their first clinical rotation. In addition to their standard training, including a Mock Induction simulation exercise, one cohort of this study received a refresher simulation midway through the first clinical rotation; a second cohort served as the control group. Data was collected on a variety of outcome measures. While performance across all measures improved over the course of the 17-week rotation, no consistent statistically significant differences were detected between the study groups. Possible explanations for these results are: 1) the intervention was not deployed at an optimal time; and 2) the magnitude of the effect of the clinical experience on learning was so large that the impact of the intervention, if any, was not detectable.

Despite this, important lessons were learned that could guide future research. In the course of this protocol, a robust method for training raters to score a highly complex psychomotor task in real time was developed. Integral to this training and to the scoring in both the simulation and clinical settings was the development of a series of data collection tools that were accessible across multiple device platforms. These tools were based on the HTA-derived Optimum Task Set for Standard Induction of General Anesthesia (OTS-SIGA) and operationalized for use on a web-based platform by WISER Information Technology personnel and proved to be reliable and easy to use. Finally, while no statistically significant differences between the two study cohorts were detected with regard to the refresher simulation intervention, outcomes data provided insight into how learning in the clinical setting occurs over time for the highly complex task set of induction of general anesthesia.

Appendix A Hierarchical Task Analysis Derived Optimal Task Set for Standard Induction of General Anesthesia (OTS-SIGA) © 2006 Goode, Jr. JS, O'Donnell JM; revised January

2016

Preop: *Discuss allergies, current medications, any concerns regarding the patient's presenting condition and current health status with your assigned CRNA.*

1. **Equipment verification-** Verbalize or demonstrate
 - 1.1. System pressure check
 - 1.2. Prepare machine (flows, open APL valve)
 - 1.3. Place suction, oral airway, tongue blade, laryngoscope, ETT, ETT tree at bedside
 - 1.4. Manual Resuscitator, manual cuff, independent light source, alarm system

2. **Interview and Airway Assessment-** Ask pertinent questions (keep short-get facts)- Demonstrate
 - 2.1. Patient ID
 - 2.2. Oral Opening
 - 2.3. Mallampati Classification
 - 2.4. Thyromental distance
 - 2.5. Cervical range of motion
 - 2.6. Mandibular length
 - 2.7. Mandibular mobility
 - 2.8. Identification of cricothyroid membrane
 - 2.9. Other exam and questions appropriate to case

3. **Apply Monitors-** Demonstrate
 - 3.1. Check that safety strap is in place on patient
 - 3.2. Pulse oximetry
 - 3.3. B/P cuff
 - 3.4. EKG
 - 3.5. Precordial stethoscope
 - 3.6. Other monitors as appropriate

4. **Pre-induction Preparation-** Demonstrate
 - 4.1. **Time out/ procedural pause**
 - 4.1.1 Includes discussion of antibiotic administration (drug, dose, timing)
 - 4.2. Adjust height of bed
 - 4.3. Axes alignment maneuvers (sniffing position etc.)
 - 4.4. Call MD
 - 4.5. Sedation medications- (midazolam, fentanyl or other appropriate medication)
 - 4.6. Other Medications
 - 4.6.1 defasciculating or pretreatment dose of muscle relaxants or other appropriate medications
 - 4.7. Pre-oxygenation (may administer at an earlier step): **must follow 4.6**
 - 4.8. Hold mask at all times or place mask straps

5. **Induction Steps-** Demonstrate
 - 5.1. State medications, dosage, and sequence
 - 5.2. Assess patient and patient monitor while induction drugs are given
 - 5.3. Observe/assess for apnea, tell patient to open eyes, check lash, adjust APL (1/4 turns)
 - 5.4. Tape eyes

6. **Mask ventilate one minute -** Demonstrate
 - 6.1. C and E positions- ('E' fingers on bony structures)
 - 6.2. Nose, chin, pull-in- clear lips and tongue
 - 6.3. Use one-provider technique for BVMV
 - 6.4. Avoid PAP > 20 cm H₂O
 - 6.5. Ineffective- Reposition head, tongue blade, oral airway

7. **Laryngoscopy and Intubation-** Demonstrate and verbalize
 - 7.1. Mouth open, clear lips, insert R, sweep and lift, joust
 - 7.2. Do not take eyes off of vocal cords once in view
 - 7.3. Insert with DV, stylet removal with approximation
 - 7.4. Insert to 22 cm at teeth (if present)
 - 7.5. Inflate cuff

8. **Verify correct ETT placement-** Demonstrate and verbalize
 - 8.1. ETCO₂ x 4 breaths
 - 8.2. Condensation and chest rise
 - 8.3. Listen to breath sounds x 5: (anterior X 2, lateral X 2, epigastric)

9. **Post-induction management-** Demonstrate
 - 9.1. Check proper settings, place on ventilator, adjust flows, turn on gas
 - 9.2. Appropriate starting volume for ventilation in ml/kg
 - 9.3. Secure ETT- appropriately
 - 9.4. OG/Esophageal/Airway
 - 9.5. Additional equipment (BIS, warmer, PNS)

10. **Manage common complications-** Verbalize and demonstrate

- 10.1. Hypotension with bradycardia
- 10.2. Hypotension with tachycardia
- 10.3. Hypertension with tachycardia
- 10.4. Hypertension with bradycardia
- 10.5. Desaturation
- 10.6. Bronchospasm
- 10.7. Arrhythmias – atrial
- 10.8. Arrhythmias- ventricular
- 10.9. Adverse drug reaction

Appendix B Example of the Standard Induction of General Anesthesia Knowledge

Assessment Instrument (SIGA-KAI) © 2001 Goode, Jr. JS, O'Donnell JM

1. According to the Induction 10-Point checklist, the equipment verification step would include:

1. Doing a system pressure check
2. Checking that the safety strap is on the patient
3. Verifying that there is active/working suction available
4. Confirming availability of a manual resuscitation bag and mask
5. Confirming availability of appropriate, functioning laryngoscope handle and blades
6. Confirming the presence of a succinylcholine for potential rapid sequence induction

- A. 1, 2, 4, 6
- B. 1, 3, 4, 5
- C. 2, 3, 4, 5
- D. 2, 3, 5, 6
- E. 3, 4, 5, 6

2. The correct statement regarding the complete anesthesia gas machine checkout is:

- A. The complete AGM checkout must be performed prior to every case
- B. Only the low pressure system check needs to be performed before each case
- C. A leak check of the breathing system is done with flows set to minimum and pressurization to at least 30 cm H₂O
- D. Between cases high flow of oxygen or medical air should be maintained to remove residual inhaled agent from the system

3. Which statement is a World Health Organization recommendation for proper patient identification?
- A. The OR nurse is the only provider required to verify the identity of patients to match them to the correct care
 - B. At least three identifiers should be used to verify each patient's identity
 - C. Each institution should establish their own patient identification processes within a healthcare system
 - D. Patients should be encouraged to participate in the identification process
4. There are _____ critical airway assessments that should be done prior to the delivery of any anesthetic. Of these _____ is the most commonly documented in the USA and the _____ is the most difficult parameter for most anesthesia providers to assess.
- A. seven, thyromental distance, thyrohyoid membrane
 - B. seven, Mallampati, cricothyroid membrane
 - C. six, mandibular length, Mallampati
 - D. six, Mallampati, mandibular displacement
5. Which monitor is not required for a standard induction of general anesthesia?
- A. BIS monitoring
 - B. Blood pressure cuff
 - C. EKG
 - D. pulse oximetry
6. Which monitor is typically the first to be placed by anesthesia providers?
- A. Blood pressure cuff
 - B. EKG
 - C. Precordial stethoscope
 - D. Pulse oximetry

7. Which statement is correct regarding the performance of the Time Out/Procedural Pause:
- A. All members of the operative team except the anesthesiologist should stop their activities and participate
 - B. Known allergies and the need for antibiotics should be confirmed
 - C. The time-out can be delayed until after initial incision when the attending surgeon is not present
 - D. Can be skipped if unilaterality of the surgical site is not an issue (e.g. cervical spine surgery)
8. Which pre-induction maneuvers should be utilized to optimize the success of laryngoscopy and intubation?
- A. Administering at least 0.05 mcg/kg of fentanyl intravenously
 - B. Adjusting the bed height to the level of the intubator's umbilicus
 - C. Aligning the oral, pharyngeal and laryngeal axes (sniffing position)
 - D. Aligning the nasopharyngeal, pharyngeal and laryngeal axes (sniffing position)
9. According to the Induction 10-Point checklist, which statement is not a step in the induction process?
- A. Stating and confirming the medications to be administered, their dosage and the sequence of administration
 - B. Observing for active respirations and directing the patient to take a breath
 - C. Assessing vital signs continuously via standard monitors during induction
 - D. Assessing lash reflex immediately after administration of muscle relaxation
10. Taping (or other protective measures) of the eyes _____.
- A. is recommended for the prevention of corneal abrasion.
 - B. occurs after sedation but before complete loss of consciousness.
 - C. should only occur **after** initiating mask ventilation.
 - D. is done to prevent visual stimuli from contributing to recall.

11. What is the maximum peak airway pressure (in cmH₂O) during mask ventilation?
- A. 15
 - B. 20
 - C. 30
 - D. 35
12. When using the 'C' and 'E' finger positioning technique during mask ventilation, the 'E' fingers are placed:
- A. on the soft tissue below the mandible to allow for tracheal stabilization.
 - B. between the bottom edge of the mask and the mandible.
 - C. on the mandible proper (bony structures).
 - D. with all three fingers on the bony structure behind the angle of the mandible.
13. When placing an endotracheal tube with the use of a stylette, when should the stylette be withdrawn?
- A. Just before the tip of the endotracheal tube is placed between the vocal cords.
 - B. When the endotracheal tube cuff begins to pass the vocal cords.
 - C. When the tip of the endotracheal tube is 5 to 6 cm below the vocal cords.
 - D. When the endotracheal tube reaches the final insertion depth (~22cm).
14. Which statement correctly reflects the process steps for insertion of a laryngoscope into the oral cavity?
- A. Open the mouth, clear the lips, insert from the right side, sweep the tongue to the left
 - B. Tilt the head toward the chest, clear the lips, insert the blade midline, joust
 - C. Move the head to neutral position, open the mouth, insert from the right side, sweep the tongue to the left
 - D. Open the mouth, clear the lips, insert from the left side, sweep the tongue to the right
15. Which of these assessments most reliably confirms endotracheal intubation?
- A. Auscultation bilaterally of the chest
 - B. Four or more consistent ETCO₂ waveforms
 - C. Moisture or condensation in the endotracheal tube
 - D. Visible chest rise

16. What is the optimal auscultation pattern to assist in confirming endotracheal tube placement?
- A. Listen to breath sounds X 4 (anterior X 2, lateral X 2)
 - B. Listen to breath sounds X 2 on the right side anteriorly and laterally
 - C. Listen to breath sounds x 5: (anterior X 2, lateral X 2, epigastric)
 - D. Listen for absence or presence of air movement over the abdomen
17. According to the Induction 10-Point checklist, which procedure would not be part of immediate post-induction management?
- A. Securing the endotracheal tube
 - B. Selecting an appropriate starting tidal volume and respiratory rate for the patient
 - C. Placement of additional equipment such as (e.g. BIS monitoring, peripheral nerve stimulator)
 - D. Taping the eyes
18. What is an appropriate starting tidal volume for mechanical ventilation?
- A. 5-9 ml/kg
 - B. 10-15 ml/kg
 - C. 16-20 ml/kg
 - D. 21-25 ml/kg
19. Post-induction your patient is hypotensive and tachycardic. Which drug would be the best initial choice to treat these symptoms?
- A. Ephedrine
 - B. Epinephrine
 - C. Nitroglycerine
 - D. Phenylephrine
20. Post-induction your patient is hypotensive and bradycardic. Which drug would be the best initial choice to treat these symptoms?
- A. Ephedrine
 - B. Epinephrine
 - C. Nitroglycerine
 - D. Phenylephrine

Appendix C Operational Definitions

Hierarchical Task Analysis Derived Optimal Task Set for Standard Induction of General Anesthesia (OTS-SIGA)

operational definitions for scoring (simulation and clinical)

Scoring options and their operational definitions are as follows:

Y = YES the step was completed

For a step to be scored as ‘Y’ the SRNA study participant must have done the step in its entirety, completely unprompted by preceptors or other persons.

N = No the step was not completed

For a step to be scored ‘N’ it must not have been done by the SRNA study participant. This should be a rare event because it indicates that no one (participant or preceptors) did the task step. More likely if a step was not done, it is because another provider has done the step, taking the opportunity away from the SRNA. This would then be scored ‘NA’ instead of ‘No’ (see below).

P = Prompted the SRNA study participant was *prompted* or *assisted* by someone to complete the step

For a step to be scored ‘P’ the SRNA study participant must be either ***verbally or non-verbally reminded or directed to do the step***. This would include being assisted to do a step (e.g. the clinical preceptor puts on some of the EKG leads or directs where they should be placed) or being corrected while doing a step (the clinical preceptor verbally and/or physically adjusts hand position on the mask during mask ventilation). If the SRNA study participant does the task independently after prompting, it is preferable to score such a step as ‘P’ and then indicate ‘Y’ when the task is completed.

NA = Not Available/Not Applicable

For a step to be scored as ‘NA’ the task must have been done by another person (usually the CRNA or Anesthesiologist preceptor) before the SRNA study participant has had a chance to do it (e.g. someone puts on the blood pressure cuff while the SRNA is doing another task like putting on the pulse oximeter). ‘NA’ can also mean “not applicable” (e.g. a precordial stethoscope is usually not used in the adult setting).

There are naturally some small differences scoring certain steps between simulation events and scoring live in the clinical setting. Below in the step-by-step scoring the operational definitions for both simulation and clinical are given. These definitions are for what is required for a step to be scored as ‘Yes’. The major differences between simulation and clinical scoring are highlighted for each step.

1. **Equipment verification-** Verbalize or demonstrate
 - 1.1. System pressure check
 - 1.2. Prepare machine (flows, open APL valve)
 - 1.3. Place suction, oral airway, tongue blade, laryngoscope, ETT, ETT tree at bedside
 - 1.4. Manual Resuscitator, manual cuff, independent light source, alarm system

Simulation Scoring:

All steps (1.1 through 1.4) must be physically done or verbally acknowledged by the SRNA to be scored as ‘Yes’ (e.g. verbally stating presence of equipment/location of code alarm button in step 1.4).

Clinical Scoring:

Ideally the SRNA should be observed doing all the steps (1.1 through 1.4), however, in the clinical setting this might not be possible (especially for first cases of the day). It is then acceptable to either: 1) ask the CRNA preceptor if the equipment verification steps were done appropriately; and/or 2) query the SRNA about these steps (“Tell me everything that you did to verify that all of your equipment was ready and available”) A comment should be made in the notes section (e.g. “Asked CRNA preceptor about equipment verification”).

2. **Interview and Airway Assessment-** Ask pertinent questions (keep short-get facts)-
Demonstrate
 - 2.1. Patient ID
 - 2.2. Oral Opening
 - 2.3. Mallampati Classification
 - 2.4. Thyromental distance
 - 2.5. Cervical range of motion
 - 2.6. Mandibular length
 - 2.7. Mandibular mobility
 - 2.8. Identification of cricothyroid membrane
 - 2.9. Other exam and questions appropriate to case

Simulation Scoring:

2.1 Patient identification should be done by confirming the information on the patient armband on the mannequin.

2.2– 2.8 All steps must be physically done or verbally stated by the SRNA. The exception is the cricothyroid membrane, which the SRNA must physically palpate on the mannequin.

2.9 At a minimum, *other exam and questions* should include the following to be scored as ‘Yes’:

- Allergies
- Time of last meal or liquid intake
- Confirmation of the procedure to be performed

Other relevant questions (cardiac history, pulmonary history, family history of anesthetic complications) may have been discussed pre-simulation with the person playing the role of the CRNA and thus might not be on the video recording.

Clinical Scoring:

2.1 Patient identification should be done using a minimum of two forms of patient ID.

2.2– 2.8 All airway assessments must be physically done or verbally stated, including palpation of the cricothyroid membrane.

2.9 At a minimum, the other exam questions should include:

- Allergies
- Time of last meal or liquid intake
- Confirmation of the procedure to be performed
- Assessment of cardiac status/history
- Assessment of pulmonary status/history
- Family history of anesthetic complications

3. **Apply Monitors-** Demonstrate

3.1. Check that safety strap is in place on patient

3.2. Pulse oximetry

3.3. B/P cuff

3.4. EKG

3.5. Precordial stethoscope

3.6. Other monitors as appropriate

Simulation Scoring:

3.1 The SRNA must verbally acknowledge the presence of the safety strap or give physical indications that they have checked that it is in place (e.g. touching or clearly being seen to look at the safety strap).

3.2-3.5 The SRNA must independently place each of these monitoring devices. We emphasized placement of the precordial stethoscope for the simulation exercise.

3.6 Other monitors would include anything else that the anesthesia team determines is needed to be in place prior to induction (e.g. pre-induction placement of an arterial line).

Clinical Scoring:

3.1 The SRNA must verbally acknowledge the presence of the safety strap or give physical indications that they have checked that it is in place (e.g. touching or clearly being seen to look at the safety strap).

3.2-3.5 The SRNA must independently place each of these monitoring devices. As precordial stethoscopes are not used very often in the adult setting, step 3.5 will generally be scored as “NA”.

3.6 Other monitors would include anything else that the anesthesia team determines is needed to be in place prior to induction (e.g. pre-induction placement of an arterial line).

4. **Pre-induction Preparation-** Demonstrate

4.1. **Time out/ procedural pause**

4.1.1. Discussion of antibiotic administration (drug, dose, timing)

4.2. Adjust height of bed

4.3. Axes alignment maneuvers (sniffing position etc.)

4.4. Call MD

4.5. Sedation medications- (midazolam, fentanyl or other appropriate medication)

4.6. Other Medications

4.6.1 defasciculating or pretreatment dose of muscle relaxants or other appropriate medications

4.7. Pre-oxygenation (may administer at an earlier step): must follow 4.6

4.8. Hold mask at all times or place mask straps

Simulation Scoring:

4.1 The SRNA should stop doing all other tasks and focus on the Time-Out process. The SRNA should ask about or appropriately discuss the plan for surgical antibiotic prophylaxis.

4.2 The SRNA should adjust the bed to the appropriate height, approximately the level of the xiphoid process. Ideally this will occur during the pre-induction preparation but could occur at a later point. *It must occur before Step 6 (Mask Ventilation).*

4.3 The head should be raised approximately 10cm (4 inches) by placing a foam pillow or another object under the occiput. This brings the laryngeal and pharyngeal axes into alignment. The head should then be extended on the atlanto-occipital joint, bringing the oral axis closer into alignment with the laryngeal and pharyngeal axes. Note that this may not be done physically in simulation because the mannequin's head is usually already on the foam pillow and SRNAs are aware from practice sessions that the mannequin head frequently does not stay in the adjusted position. At a minimum, verbal explanation of axes alignment should be given.

4.4 The SRNA should call or direct someone else to call the attending anesthesiologist to inform them that the patient is ready for induction.

4.5 If appropriate, the SRNA should administer sedative/anxiolytic medications and/or

analgesics such as fentanyl. This may occur at an earlier or later step. Note that it is appropriate for the SRNA to ask the preceptor before doing so and this would not be considered as being prompted. Additionally, the SRNA should cleanse the injection port with alcohol ('scrub the hub') or place injection ports on the stop cock and then 'scrub the hub'. It is acceptable to use another port on the IV tubing, but again, this port should be appropriately cleansed before administering medications.

4.6 If appropriate, the SRNA should administer or ask about administering other medications (e.g. defasciculating or pre-treatment doses of muscle relaxant). This may occur at an earlier or later step. Note that it is appropriate for the SRNA to ask the preceptor before doing so and this would not be considered as being prompted. Injection ports should be cleansed as described in step 4.6 above.

4.7 The SRNA should begin pre-oxygenation via face mask. This may occur at an earlier time but must occur after either step 4.6 or step 4.7.

4.8 The SRNA should hold the face mask on or near the face at all times. Alternatively, mask straps may be used or the SRNA may ask another provider to do so if another task needs to be done.

Clinical Scoring:

4.1 The SRNA should stop doing all other tasks and focus on the Time-Out process. The SRNA should ask about or appropriately discuss the plan for surgical antibiotic prophylaxis.

4.2 The SRNA should adjust the bed to the appropriate height, approximately the level of the xiphoid process. Ideally this will occur during the pre-induction preparation but could occur at a later point. It must occur before Step 6 (Mask Ventilation).

4.3 The head should be raised approximately 10cm (4 inches) by placing a foam pillow or another object under the occiput. This brings the laryngeal and pharyngeal axes into alignment. The head should then be extended on the atlanto-occipital joint, bringing the oral axis closer into alignment with the laryngeal and pharyngeal axes. Note that final head positioning may not occur until Step 5: Induction.

4.4 The SRNA should call or direct someone else to call the attending anesthesiologist to inform them that the patient is ready for induction.

4.5 If appropriate, the SRNA should administer sedative/anxiolytic medications and/or analgesics such as fentanyl. This may occur earlier (e.g. in the pre-operative holding area) or later. Note that it is appropriate for the SRNA to ask the preceptor before doing so and this would not be considered as being prompted. Additionally, the SRNA should cleanse the injection port with alcohol ('scrub the hub') or place injection ports on the stop cock and then 'scrub the hub'. It is acceptable to use another port on the IV tubing, but again, this port should be appropriately cleansed before administering medications.

4.6 If appropriate, the SRNA should administer or ask about administering other medications (e.g. defasciculating or pre-treatment doses of muscle relaxant). This may occur at an earlier or later

step (but not prior to placement of all monitors). Note that it is appropriate for the SRNA to ask the preceptor before doing so and this would not be considered as being prompted. Additionally, the SRNA should cleanse the injection port with alcohol ('scrub the hub') or place injection ports on the stop cock and then 'scrub the hub'. It is acceptable to use another port on the IV tubing, but again, this port should be appropriately cleansed before administering medications.

4.7 The SRNA should begin pre-oxygenation via face mask. This may occur at an earlier time but must occur after either step 4.6 or step 4.7.

4.8 The SRNA should hold the face mask on or near the face at all times. Alternatively, mask straps may be used or the SRNA may ask another provider to do so if another task needs to be done.

5. **Induction Steps-** Demonstrate

5.1. State medications, dosage, and sequence

5.2. Assess patient and patient monitor while induction drugs are given

5.3. Observe/assess for apnea, tell patient to open eyes, check lash, adjust APL (1/4 turns)

5.4. Tape eyes

Simulation Scoring:

5.1 The SRNA should either state or be able to correctly answer questions (if asked) about induction medication dosages and sequence of administration.

5.2 The SRNA should verbally state or display indications (e.g. looking at the monitor) that they are assessing the patient and vital signs during the induction.

5.3 The SRNA should verbally state or display indications that they are assessing for the onset of apnea (e.g. looking at the patient for chest excursion, looking at the monitor for capnography tracings). They should be directive in assessing for unconsciousness (e.g. stating "take a breath, open your eyes") and should correctly assess the lash reflex.

5.4 The SRNA should correctly protect the patient's eyes. This can be done with tape or with other protective devices that may be the standard at different clinical sites (specific eye coverings or goggles, etc.)

Clinical Scoring:

5.1 The SRNA should either state or be able to correctly answer questions (if asked) about induction medication dosages and sequence of administration.

5.2 The SRNA should verbally state or display indications (e.g. looking at the monitor) that they are assessing the patient and vital signs during the induction.

5.3 The SRNA should verbally state or display indications that they are assessing for the onset of apnea (e.g. looking at the patient for chest excursion, looking at the monitor for capnography tracings). They should be directive in assessing for unconsciousness (e.g. stating

“take a breath, open your eyes’) and should correctly assess the lash reflex.

5.4 The SRNA should correctly protect the patient’s eyes. This can be done with tape or with other protective devices that may be the standard at different clinical sites (specific eye coverings or goggles, etc.). *Note that this may occur at a later time depending on the clinical site and/or clinical preceptor preference.*

- 6. Mask ventilate one minute - Demonstrate**
- 6.1. C and E positions- ('E' fingers on bony structures)
 - 6.2. Nose, chin, pull-in- clear lips and tongue
 - 6.3. Use one-provider technique for BVMV
 - 6.4. Avoid PAP > 20 cm H₂O
 - 6.5. Ineffective- Reposition head, tongue blade, oral airway

Simulation Scoring:

6.1 The SRNA should have the fingers of the left hand in the 'C' (thumb and index fingers on the cone of the mask) and 'E' (remaining fingers on the mandible) positions. The 'E' fingers should only be on the mandible proper and not on the soft tissue of the neck.

6.2 The mask should be placed and seated from the nose to the chin. The 'E' position fingers should primarily be used to pull the mandible up and into the mask.

6.3 the SRNA should be able to provide adequate ventilation using the one-provider, one-handed mask seal technique without assistance.

6.4 The SRNA should adjust the anesthesia gas machine APL valve to provide adequate ventilation while avoiding airway pressures in excess of 20 cm H₂O.

6.5 If there is difficulty with delivering breaths of adequate tidal volume, the SRNA should recognize this, reposition the mask and/or the head or alternatively place an appropriately sized oral airway using a tongue blade. ***The inverted insertion technique for the oral airway is not considered appropriate.***

Clinical Scoring:

6.1 The SRNA should have the fingers of the left hand in the 'C' (thumb and index fingers on the cone of the mask) and 'E' (remaining fingers on the mandible) positions. The 'E' fingers should only be on the mandible proper and not on the soft tissue of the neck.

6.2 The mask should be placed and seated from the nose to the chin. The 'E' position fingers should primarily be used to pull the mandible up and into the mask.

6.3 The SRNA should be able to provide adequate ventilation using the one-provider, one-handed mask seal technique without assistance.

6.4 The SRNA should adjust the anesthesia gas machine APL valve to provide adequate ventilation while avoiding airway pressures in excess of 20 cm H₂O.

6.5 If there is difficulty with delivering breaths of adequate tidal volume, the SRNA should recognize this and reposition the mask and/or the head. Alternatively, an appropriately sized oral airway should be placed using a tongue blade. ***The inverted insertion technique for the oral airway may only be considered appropriate if the clinical preceptors direct the SRNA to do so.***

7. Laryngoscopy and Intubation- Demonstrate and verbalize

- 7.1. Mouth open, clear lips, insert R, sweep and lift, joust
- 7.2. Do not take eyes off of vocal cords once in view
- 7.3. Insert with DV, stylet removal with approximation
- 7.4. Insert to 22 cm at teeth (if present)
- 7.5. Inflate Cuff

Simulation Scoring:

7.1 The SRNA should open the mouth using either the cross-finger technique or by pressing on the mentum. The lips should be cleared of compression by the laryngoscope blade as the blade is inserted from the right side of the mouth. The tongue should be swept from right to left. After the blade is in its final position, the SRNA should use a jousting-like lift technique (no ‘breaking’ of the wrist which moves the blade toward the mandibular teeth). Ideally, these steps should be verbalized.

7.2 Once the glottis is visualized, the SRNA should not avert their eyes from the view of the glottis.

7.3 Once the cuff of the endotracheal tube is approximated to the vocal cords, the SRNA should ask that the stylette be removed from the endotracheal tube. The SRNA must maintain a secure grip on the ETT.

7.4 While continuing to maintain a view of the glottis, the SRNA should advance the endotracheal tube to an appropriate depth (approximately 22 cm at the teeth).

7.5 While maintaining the position of the endotracheal tube, the SRNA should inflate the endotracheal tube cuff or direct someone else to do so.

Clinical Scoring:

7.1 The SRNA should open the mouth using either the cross-finger technique or by pressing on the mentum. The lips should be cleared of compression by the laryngoscope blade as the blade is inserted from the right side of the mouth. The tongue should be swept from right to left. After the blade is in its final position, the SRNA should use a jousting-like lift technique (no ‘breaking’ of the wrist which moves the blade toward the teeth). Ideally, these steps should be verbalized.

7.2 Once the glottis is visualized, the SRNA should not avert their eyes from the view of the glottis.

7.3 Once the cuff of the endotracheal tube is approximated to the vocal cords, the SRNA should ask that the stylette be removed from the endotracheal tube. The SRNA must maintain a secure grip on the ETT.

7.4 While continuing to maintain a view of the glottis, the SRNA should advance the endotracheal tube to an appropriate depth (approximately 22 cm at the teeth).

7.5 While maintaining the position of the endotracheal tube, the SRNA should inflate the endotracheal tube cuff or direct someone else to do so.

8. **Verify correct ETT placement-** Demonstrate and verbalize
 - 8.1. ETCO₂ x 4 breaths
 - 8.2. Condensation and chest rise
 - 8.3. Listen to breath sounds x 5: (anterior X 2, lateral X 2, epigastric)

Simulation Scoring:

8.1 The SRNA should observe for and verbally confirm that a minimum of 4 consecutive and consistent End Tidal CO₂ waveforms are seen on the monitor. Ideally the SRNA should provide the actual ventilations via manual resuscitation bag or the anesthesia circuit, but it is acceptable if the SRNA directs someone else to perform the ventilations.

8.2 The SRNA should verbally confirm the presence of chest rise. While the mannequin does not provide condensation in the endotracheal tube, the SRNA should verbally state that this should be seen with correct endotracheal tube placement.

8.3 The SRNA should listen for breath sounds on the anterior chest wall (both left and right), on the lateral chest wall (both left and right) and over the epigastrium. Some providers prefer a sequential order (left anterior, right anterior, right lateral, left lateral, epigastrium) but this exact sequence is not required for scoring this step as 'Yes'. Ventilations can be provided by the SRNA while they are listening or can be done by another provider. It would also be acceptable for the SRNA to specifically direct another provider to auscultate.

Clinical Scoring:

8.1 The SRNA should observe for and verbally confirm that a minimum of 4 consecutive and consistent End Tidal CO₂ waveforms are seen on the monitor. Ideally the SRNA should provide the actual ventilations via manual resuscitation bag or the anesthesia circuit, but it is acceptable if the SRNA directs someone else to perform the ventilations.

8.2 The SRNA should verbally confirm the presence of chest rise. The SRNA should verbally state whether condensation in the endotracheal tube is seen.

8.3 The SRNA should listen for breath sounds on the anterior chest wall (both left and right), on the lateral chest wall (both left and right) and over the epigastrium. Some providers prefer a sequential order (left anterior, right anterior, right lateral, left lateral, epigastrium) but this exact sequence is not required for scoring this step as 'Yes'. Ventilations can be provided by the SRNA while they are listening or can be done by another provider. It would also be acceptable for the SRNA to specifically direct another provider to auscultate.

9. **Post-induction management-** Demonstrate
 - 9.1. Check proper settings, place on ventilator, adjust flows, turn on gas
 - 9.2. Appropriate starting volume for ventilation in ml/kg
 - 9.3. Secure ETT- appropriately
 - 9.4. OG/Esophageal/Airway
 - 9.5. Additional equipment (BIS, warmer, PNS)

Simulation Scoring:

9.1 The SRNA should confirm the settings on the Anesthesia Gas Machine ventilator and turn on the ventilator. Fresh gas flows should be adjusted, and an inhalational agent initiated at an appropriate concentration.

9.2 Confirmation of settings should include the proper weight-based tidal volume and an appropriate respiratory rate. Ideally this should be verbally stated. *Note that these settings may have been adjusted in Step 1.2 above.*

9.3 The SRNA should secure the endotracheal tube. There are a variety of appropriate taping styles (and sometimes ETT oral fixation devices), but at a minimum the taping should not allow for significant movement of the ETT and the tape should not be adhered to the surface of the lip proper (tape should be above the vermilion of the upper lip).

9.4 If appropriate, the SRNA should place an orogastric or nasogastric tube. Additional equipment to be placed orally could include an esophageal temperature probe and an appropriate bite block.

9.5 The SRNA should place any additional monitoring or patient safety equipment at this time. This might include, but is not limited to, a temperature probe (if an esophageal probe was not placed), BIS (awareness) monitoring, a peripheral nerve stimulator and a warming blanket.

Clinical Scoring:

9.1 The SRNA should confirm the settings on the Anesthesia Gas Machine ventilator and turn on the ventilator. Fresh gas flows should be adjusted, and an inhalational agent initiated at an appropriate concentration.

9.2 Confirmation of settings should include the proper weight-based tidal volume and an appropriate respiratory rate. Ideally this should be verbally stated. *Note that these settings may have been adjusted in Step 1.2 above.*

9.3 The SRNA should secure the endotracheal tube. There are a variety of appropriate taping styles (and sometimes ETT oral fixation devices), but at a minimum the taping should not allow for significant movement of the ETT and the tape should not be adhered to the surface of the lip proper (tape should be above the vermilion of the upper lip).

9.4 If appropriate, the SRNA should place an orogastric or nasogastric tube. Additional equipment to be placed orally could include an esophageal temperature probe and an appropriate bite block.

9.5 The SRNA should place any additional monitoring or patient safety equipment at this time. This might include, but is not limited to, a temperature probe (if an esophageal probe was not placed), BIS (awareness) monitoring, a peripheral nerve stimulator and a warming blanket.

10. Manage common complications- Verbalize and demonstrate

- 10.1. Hypotension with bradycardia
- 10.2. Hypotension with tachycardia
- 10.3. Hypertension with tachycardia
- 10.4. Hypertension with bradycardia
- 10.5. Desaturation
- 10.6. Bronchospasm
- 10.7. Arrhythmias – atrial
- 10.8. Arrhythmias- ventricular
- 10.9. Adverse drug reaction

Simulation Scoring:

10.1 – 10.9 All Mock Induction participants will experience one of the common post-induction complications listed here. These are automatically triggered in the Laerdal SimMan software when Step 9 (Post-Induction Management) is marked as completed. The primary objective is for the SRNA to recognize the physiologic derangement. To be scored as ‘Yes’ the SRNA should be able to identify and describe any physiologic changes that have occurred. The Mock Induction is structured so that once the complication is identified the SRNA will be walked through answering questions about the etiology and an appropriate intervention by the in-room preceptor. ‘ P ’ would be used if the SRNA had to be prompted to identify the physiologic derangement (e.g. “what are you seeing on the monitor”). Etiologies and suggested interventions for these common adverse events are discussed in the Mock Induction Faculty Manual.

Clinical Scoring:

10.1 – 10.9 It is possible for a given induction of general anesthesia in the clinical setting that the SRNA *will not* experience one of the common immediate post-induction complications listed here. If this is the case, the step should either not be scored or alternately ‘NA’ can be scored for each option. If a complication occurs, to be scored as ‘Yes’ the SRNA should be able to identify and describe any physiologic changes that have occurred. Once the complication is identified the SRNA should be able to answer questions about the etiology and an appropriate intervention. ‘ P ’ would be used if the SRNA had to be prompted to identify the physiologic derangement (e.g. “what are you seeing on the monitor”) or needed significant prompting for the etiology and/or treatment.

Bibliography

- Alinier, G. (2011). Developing high-fidelity health care simulation scenarios: A guide for educators and professionals. *Simulation & Gaming, 42*(1), 9-26.
- Annett, J., Cunningham, D., Mathias-Jones, P., Annett, J., Cunningham, D., & Mathias-Jones, P. (2000). A method for measuring team skills. *Ergonomics, 43*(8), 1076-1094.
- Bailer-Jones, D. (2002). Models, Metaphors and Analogies. In P. K. Machamer & M. Silberstein (Eds.), *The Blackwell guide to the philosophy of science* (pp. 108-127). Oxford: Blackwell.
- Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *Journal of Personality & Social Psychology, 51*(6), 1173-1182.
- Barsuk, J. H., Ahya, S. N., Cohen, E. R., McGaghie, W. C., & Wayne, D. B. (2009). Mastery learning of temporary hemodialysis catheter insertion by nephrology fellows using simulation technology and deliberate practice. *Am J Kidney Dis, 54*(1), 70-76. doi:10.1053/j.ajkd.2008.12.041
- Barsuk, J. H., Cohen, E. R., McGaghie, W. C., & Wayne, D. B. (2010). Long-term retention of central venous catheter insertion skills after simulation-based mastery learning. *Academic Medicine, 85*(10 Suppl), S9-12. doi:10.1097/ACM.0b013e3181ed436c/00001888-201010001-00006 [pii]
- Barsuk, J. H., Cohen, E. R., Mikolajczak, A., Seburn, S., Slade, M., & Wayne, D. B. (2015). Simulation-Based Mastery Learning Improves Central Line Maintenance Skills of ICU Nurses. *J Nurs Adm, 45*(10), 511-517. doi:10.1097/nna.0000000000000243
- Barsuk, J. H., McGaghie, W. C., Cohen, E. R., Balachandran, J. S., & Wayne, D. B. (2009). Use of simulation-based mastery learning to improve the quality of central venous catheter placement in a medical intensive care unit. *J Hosp Med, 4*(7), 397-403. doi:10.1002/jhm.468
- Barsuk, J. H., McGaghie, W. C., Cohen, E. R., O'Leary, K. J., & Wayne, D. B. (2009). Simulation-based mastery learning reduces complications during central venous catheter insertion in a medical intensive care unit. *Crit Care Med, 37*(10), 2697-2701.
- Batterman, R. W. (2002). *The devil in the details : asymptotic reasoning in explanation, reduction, and emergence*. Oxford ; New York: Oxford University Press.
- Batterman, R. W., & Rice, C. C. (2014). Minimal Model Explanations. *Philosophy of Science, 81*(3), 349-376. doi:10.1086/676677
- Benumof, J. L. (1996). The American Society of Anesthesiologists' Management of the Difficult Airway Algorithm and Explanation-Analysis of the Algorithm. In J. L. Benumof (Ed.), *Airway Management: Principles and Practice* (pp. 143-156): Mosby.
- Beyea, S. C., von Reyn, L. K., & Slattery, M. J. (2007). A nurse residency program for competency development using human patient simulation. *J Nurses Staff Dev, 23*(2), 77-82.
- Britt, R. C., Reed, S. F., & Britt, L. D. (2007). Central line simulation: a new training algorithm. *Am Surg, 73*(7), 680-682; discussion 682-683.
- Burke, H., & Mancuso, L. (2012). Social cognitive theory, metacognition, and simulation learning in nursing education. *J Nurs Educ, 51*(10), 543-548. doi:10.3928/01484834-20120820-02

- Chamberlain, B. (2008). Remembering Mrs. Chase... "Mrs Chase" (in the February/March 2008 edition of Imprint). *Imprint (00193062)*, 55(3), 8-8.
- Chan, P. S., Jain, R., Nallmothu, B. K., Berg, R. A., & Sasson, C. (2010). Rapid Response Teams: A Systematic Review and Meta-analysis. *Arch Intern Med*, 170(1), 18-26. doi:10.1001/archinternmed.2009.424
- Chan, P. S., Khalid, A., Longmore, L. S., Berg, R. A., Kosiborod, M., & Spertus, J. A. (2008). Hospital-wide code rates and mortality before and after implementation of a rapid response team. *JAMA*, 300(21), 2506-2513. doi:10.1001/jama.2008.715
- Chen, J. C., Amar, A. P., Levy, M. L., & Apuzzo, M. L. (1999). The development of anatomic art and sciences: the ceroplastica anatomic models of La Specola. *Neurosurgery*, 45(4), 883-891; discussion 891-882.
- Cheng, A., Kessler, D., Mackinnon, R., Chang, T. P., Nadkarni, V. M., Hunt, E. A., . . . Education Reporting Guidelines, I. (2016). Reporting Guidelines for Health Care Simulation Research: Extensions to the CONSORT and STROBE Statements. *Simul Healthc*, 11(4), 238-248. doi:10.1097/SIH.0000000000000150
- Chi, M. T. H., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science*, 13(2), 145-182. doi:https://doi.org/10.1016/0364-0213(89)90002-5
- Clark, C. (1976). Simulation gaming: a new teaching strategy in nursing education. *Nurse Educator*, 1(4), 4-9.
- Clark, C. (1977). Learning outcomes in a simulation game for associate degree nursing students. *Health Education Monographs*, 5 suppl 1, 18-27.
- Cohen, Feinglass, J., Barsuk, J. H., Barnard, C., O'Donnell, A., McGaghie, W. C., & Wayne, D. B. (2010). Cost savings from reduced catheter-related bloodstream infection after simulation-based education for residents in a medical intensive care unit. *Simul Healthc*, 5(2), 98-102. doi:10.1097/SIH.0b013e3181bc8304
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, N.J.: L. Erlbaum Associates.
- Cohen, J. (1992). A power primer. *Psychol Bull*, 112(1), 155-159.
- Coleman, J., Livingston, S., Fennessey, G., Edwards, K., & Kidder, S. (1973). The Hopkins Games Program: Conclusions From Seven Years of Research. *Educational Researcher*, 2(8), 3-7.
- Colyvan, M., Regan, H. M., & Ferson, S. (2001). Is it a Crime to Belong to a Reference Class. *Journal of Political Philosophy*, 9(2), 168-181. doi:10.1111/1467-9760.00123
- Cook, D. A., Brydges, R., Hamstra, S. J., Zendejas, B., Szostek, J. H., Wang, A. T., . . . Hatala, R. (2012). Comparative effectiveness of technology-enhanced simulation versus other instructional methods: a systematic review and meta-analysis. *Simul Healthc*, 7(5), 308-320. doi:10.1097/SIH.0b013e3182614f95
- Cook, D. A., Brydges, R., Zendejas, B., Hamstra, S. J., & Hatala, R. (2013a). Mastery Learning for Health Professionals Using Technology-Enhanced Simulation: A Systematic Review and Meta-Analysis. *Acad Med*, 88(8), 1178-1186. doi:10.1097/ACM.0b013e31829a365d
- Cook, D. A., Brydges, R., Zendejas, B., Hamstra, S. J., & Hatala, R. (2013b). Technology-enhanced simulation to assess health professionals: a systematic review of validity evidence, research methods, and reporting quality. *Acad Med*, 88(6), 872-883. doi:10.1097/ACM.0b013e31828ffdcf

- Cook, D. A., Erwin, P. J., & Triola, M. M. (2010). Computerized virtual patients in health professions education: a systematic review and meta-analysis. *Academic Medicine*, 85(10), 1589-1602. doi:10.1097/ACM.0b013e3181edfe13
- Cook, D. A., Hamstra, S. J., Brydges, R., Zendejas, B., Szostek, J. H., Wang, A. T., . . . Hatala, R. (2013). Comparative effectiveness of instructional design features in simulation-based education: systematic review and meta-analysis. *Med Teach*, 35(1), e867-898. doi:10.3109/0142159X.2012.714886
- Cook, D. A., Hatala, R., Brydges, R., Zendejas, B., Szostek, J. H., Wang, A. T., . . . Hamstra, S. J. (2011). Technology-enhanced simulation for health professions education: a systematic review and meta-analysis. *JAMA*, 306(9), 978-988. doi:10.1001/jama.2011.1234
- Cook, D. A., Levinson, A. J., Garside, S., Dupras, D. M., Erwin, P. J., & Montori, V. M. (2010). Instructional design variations in internet-based learning for health professions education: a systematic review and meta-analysis. *Academic Medicine*, 85(5), 909-922. doi:10.1097/ACM.0b013e3181d6c319
- Cooper, J. B., & Taqueti, V. R. (2004a). A brief history of the development of mannequin simulators for clinical education and training. *Qual Saf Health Care*, 13 Suppl 1, i11-18.
- Cooper, J. B., & Taqueti, V. R. (2004b). A brief history of the development of mannequin simulators for clinical education and training. *Quality & Safety in Health Care*, 13 Suppl 1, i11-18.
- Cooper, J. B., & Taqueti, V. R. (2008). A brief history of the development of mannequin simulators for clinical education and training. *Postgraduate Medical Journal*, 84(997), 563-570.
- Crofts, J. F., Ellis, D., Draycott, T. J., Winter, C., Hunt, L. P., & Akande, V. A. (2007). Change in knowledge of midwives and obstetricians following obstetric emergency training: a randomised controlled trial of local hospital, simulation centre and teamwork training. *BJOG: An International Journal of Obstetrics & Gynaecology*, 114(12), 1534-1541.
- Crofts, J. F., Fox, R., Ellis, D., Winter, C., Hinshaw, K., & Draycott, T. J. (2008). Observations from 450 shoulder dystocia simulations: lessons for skills training.[see comment]. *Obstetrics & Gynecology*, 112(4), 906-912.
- Dacome, L. (2007). Women, wax and anatomy in the 'century of things'. *Renaissance Studies*, 21(4), 522-550. doi:10.1111/j.1477-4658.2007.00461.x
- DeVita, M. A., & Minnini, N. A. (2004). Improving medical crisis team performance. *Critical Care Medicine*, 32(2s), s61-65.
- Dewey, J. (1910). *How we think*. Boston,: D.C. Heath & Co.
- Draycott, T. J., Crofts, J. F., Ash, J. P., Wilson, L. V., Yard, E., Sibanda, T., & Whitelaw, A. (2008). Improving neonatal outcome through practical shoulder dystocia training. *Obstetrics & Gynecology*, 112(1), 14-20.
- Durham, C. F., Cato, M. L., & Lasater, K. (2014). NLN/Jeffries Simulation Framework State of the Science Project: Participant Construct. *Clinical Simulation In Nursing*, 10(7), 363-372. doi:10.1016/j.ecns.2014.04.002
- Gaba, D. M., & DeAnda, A. (1988). A comprehensive anesthesia simulation environment: recreating the operating room for research and training. *Anesthesiology*, 69(3), 387-394.
- Goldenfeld, N. (1992). *Lectures on phase transitions and the renormalization group*. Reading, Mass.: Addison-Wesley, Advanced Book Program.
- Goode, J., Schumann, T., Klain, M., & O'Donnell, J. M. (2007). Evaluation of a Novel System for Quantitative Assessment of Bag Valve Mask Ventilation (BVMV). *Simulation in Healthcare: The Journal of the Society for Simulation in Healthcare*, 1(1), 96.

- Goode Jr, J. (2016). Action, Practice, and Reflection: Dewey's Pragmatist Philosophy and the Current Simulation Healthcare Movement. In R. Bluhm (Ed.), *Knowing and acting in medicine* (pp. 239-257). Lanham: Rowman & Littlefield International.
- Goode Jr, J. S., & O'Donnell, J. M. (2016). Simulation in Nursing Education and Practice. In R. Riley (Ed.), *Manual of simulation in healthcare* (Second edition. ed., pp. 115-127). Oxford: Oxford University Press.
- Greenblat, C. S. (1971). Simulation Games and the Sociologist. *American Sociologist*, 6(2), 161-164.
- Greenblat, C. S. (1977). Gaming-simulation and health education an overview. *Health Education Monographs*, 5 suppl 1, 5-10.
- Groom, J. A., Henderson, D., & Sittner, B. J. (2014). NLN/Jeffries Simulation Framework State of the Science Project: Simulation Design Characteristics. *Clinical Simulation In Nursing*, 10(7), 337-344. doi:10.1016/j.ecns.2013.02.004
- Group., T. C. W. (2016). *Healthcare Simulation Dictionary*. Paper presented at the International Meeting for Simulation in Healthcare, Orlando, FL. <http://www.ssih.org/dictionary>.
- Hallmark, B. F., Thomas, C. M., & Gantt, L. (2014). The Educational Practices Construct of the NLN/Jeffries Simulation Framework: State of the Science. *Clinical Simulation In Nursing*, 10(7), 345-352. doi:10.1016/j.ecns.2013.04.006
- Haque, S., & Srinivasan, S. (2006). A meta-analysis of the training effectiveness of virtual reality surgical simulators. *IEEE Trans Inf Technol Biomed*, 10(1), 51-58.
- Harder, B. N. (2010). Use of simulation in teaching and learning in health sciences: a systematic review. *J Nurs Educ*, 49(1), 23-28. doi:10.3928/01484834-20090828-08
- Harris, K., Eccles, D. W., Ward, P., & Whyte, J. (2013). Theoretical framework for stimulation in nursing: answering Schiavenato's call. *J Nurs Educ*, 52(1), 6-16.
- Herrmann, E. K. (1981). Mrs. Chase: a noble and enduring figure. *American Journal of Nursing*, 81(10), 1836.
- Herrmann, E. K. (2008). Remembering Mrs. Chase: before there were SMART HOSPITALS and SIM-MEN, there was "Mrs. Chase". *Imprint*, 55(2), 52-55.
- Hoffmann, R., O'Donnell, J. M., & Kim, Y. (2007). The Effects of Human Patient Simulators on Basic Knowledge in Critical Care Nursing with Undergraduate Senior Baccalaureate Nursing Students. *Simulation in Healthcare: The Journal of the Society for Simulation in Healthcare*, 2(2), 110-114.
- Hunt, E. A., Shilkofski, N. A., Stavroudis, T. A., & Nelson, K. L. (2007). Simulation: translation to improved team performance. *Anesthesiol Clin*, 25(2), 301-319.
- INACSL, S. C. (2016). INACSL Standards of Best Practice: SimulationSM Simulation Design. In *Clinical Simulation In Nursing* (Vol. 12, pp. S5-S12): Elsevier.
- Issenberg, S. B., McGaghie, W. C., Petrusa, E. R., Lee Gordon, D., & Scalese, R. J. (2005). Features and uses of high-fidelity medical simulations that lead to effective learning: a BEME systematic review. *Medical Teacher*, 27(1), 10-28.
- Jaccard, J., & Jacoby, J. (2010). *Theory construction and model-building skills : a practical guide for social scientists*. New York: Guilford Press.
- Jeffries. (2005). A framework for designing, implementing, and evaluating simulations used as teaching strategies in nursing. *Nursing Education Perspectives*, 26(2), 96-103.
- Jeffries. (2007). *Simulation in Nursing Education* (1st ed.): National League for Nursing.

- Jones, A. L., Reese, C. E., & Shelton, D. P. (2014). NLN/Jeffries Simulation Framework State of the Science Project: The Teacher Construct. *Clinical Simulation In Nursing*, 10(7), 353-362. doi:10.1016/j.ecns.2013.10.008
- Kneebone, R. (2003). Simulation in surgical training : educational issues and practical implications. *Medical Education*, 37, 267-277.
- Kneebone, R. (2005). Evaluating clinical simulations for learning procedural skills: a theory-based approach. *Academic Medicine*, 80(6), 549-553.
- Kneebone, R. L. (2016). Simulation reframed. *Adv Simul (Lond)*, 1, 27. doi:10.1186/s41077-016-0028-8
- Knight, C. M., Moule, P., & Desbottes, Z. (2000). The grid that bridges the gap. [(30 ref)]. *Nurse Education Today*, 20(2), 116-122.
- Lawson, D. (1990). Anesthesia Simulator Curriculum Conference September 22-24, 1989 Rockville, Maryland. *Anesthesiology*, 72(3), 587.
- . Littell's Living Age. (1852, April 10, 1852), 33, 88-89.
- Mackinnon, D. P. (2011). Integrating Mediators and Moderators in Research Design. *Research on Social Work Practice*, 21(6), 675-681. doi:10.1177/1049731511414148
- Mc Gaghie, W., Siddall, V., Mazmanian, P., Myers, J., American College of Chest Physicians, H., & Science Policy, C. (2009). Lessons for continuing medical education from simulation research in undergraduate and graduate medical education: effectiveness of continuing medical education: American College of Chest Physicians Evidence-Based Educational Guidelines. *Chest*, 135(3 Suppl), 62S-68S. doi:http://dx.doi.org/10.1378/chest.08-2521
- McGaghie, W. C., & Harris, I. B. (2018). Learning Theory Foundations of Simulation-Based Mastery Learning. *Simul Healthc*, 13(3S Suppl 1), S15-S20. doi:10.1097/SIH.0000000000000279
- McGaghie, W. C., Issenberg, S. B., Petrusa, E. R., & Scalese, R. J. (2010). A critical review of simulation-based medical education research: 2003-2009. *Med Educ*, 44(1), 50-63. doi:10.1111/j.1365-2923.2009.03547.x
- McGovern, B., Lapum, J., Clune, L., & Martin, L. S. (2013). Theoretical framing of high-fidelity simulation with Carper's fundamental patterns of knowing in nursing. *J Nurs Educ*, 52(1), 46-49. doi:10.3928/01484834-20121217-02
- Messbarger, R. (2001). Waxing Poetic: Anna Morandi Manzolini's Anatomical Sculptures. *Configurations*, 9(1), 65-97.
- Motola, I., Devine, L. A., Chung, H. S., Sullivan, J. E., & Issenberg, S. B. (2013). Simulation in healthcare education: a best evidence practical guide. AMEE Guide No. 82. *Medical Teacher*, 35(10), e1511-1530.
- Nickerson, M., Morrison, B., & Pollard, M. (2011). Simulation in nursing staff development: a concept analysis. *J Nurses Staff Dev*, 27(2), 81-89. doi:10.1097/NND.0b013e3181a68abd
- Nicol, M., & Freeth, D. (1998). Assessment of clinical skills: a new approach to an old problem. *Nurse Education Today*, 18(8), 601-609.
- Niles, D., Sutton, R. M., Donoghue, A., Kalsi, M. S., Roberts, K., Boyle, L., . . . Nadkarni, V. (2009). "Rolling Refreshers": a novel approach to maintain CPR psychomotor skill competence. *Resuscitation*, 80(8), 909-912. doi:http://dx.doi.org/10.1016/j.resuscitation.2009.04.021
- O' Donnell J, M., Goode Jr, J. S., Odonohoe, O., & Choe, M. (2007). *The Impact of Intravenous Catheter Insertion Training Modalities on Clinical Intravenous Catheter Insertion*

- Performance in Graduate Nursing Students*. Paper presented at the 7th Annual International Meeting for Simulation in Healthcare, ; Lake Buena Vista, FL.
- O' Donnell, J. M., Goode, J. S., Henker, R. A., Kelsey, S., Bircher, N., Peele, P., . . . Sutton-Tyrrell, K. (2012). An Ergonomic Protocol for Patient Transfer That Can Be Successfully Taught Using Simulation Methods. *Clinical Simulation In Nursing*, 8(1), e3-e14.
- O' Donnell, J. M., Goode, J. S., Jr., Henker, R., Kelsey, S., Bircher, N. G., Peele, P., . . . Sutton-Tyrrell, K. (2011). Effect of a simulation educational intervention on knowledge, attitude, and patient transfer skills: from the simulation laboratory to the clinical setting. *Simul Healthc*, 6(2), 84-93. doi:10.1097/SIH.0b013e318212f1ef
- O' Donnell, J. M., Levett-Jones, T., Decker, S., & Howard, V. M. (2012). NLN-Jeffries Simulation Framework Project- Outcomes of Simulation Education. Presentation Abstracts from 2012 INACSL Conference. *Clinical Simulation In Nursing*, 8(8), e410.
- O' Donnell, J., Beach, M., & Hoffman, L. (2005). Human Simulation Training in the ICU: Applicability, Value and Disadvantages. *Critical Care Alert*, 13(6), 45-48.
- O' Donnell, J., Fletcher, J., Dixon, B., & Palmer, L. (1998). Planning and implementing an anesthesia crisis resource management course for student nurse anesthetists. *CRNA*, 9(2), 50-58.
- O' Donnell, J., Goode, J., Bradle, J., Henker, R., Kelsey, S., Peele, P., . . . Sutton-Tyrell, K. (2007). Use of Hierarchical Task Analysis Methodology to Develop a Simulation Educational Program and Unique Measurement Tools. *Simulation in Healthcare: The Journal of the Society for Simulation in Healthcare*, In press.
- O' Donnell, J., & Goode, J. J. (2008). Simulation in Nursing Education and Practice. In R. H. Riley (Ed.), *Manual of Simulation in Healthcare* (first ed., pp. 241-276). Oxford New York: Oxford University Press.
- O'Donnell, J., Rodgers, D., Lee, W., Gordon, D., Edelson, D., Haag, J., . . . Meeks, R. (2009). Structured and Supported Debriefing. Adapted from: Structured and Supported Debriefing. Dallas, TX: American Heart Association.
- Ogan, K., Jacomides, L., Shulman, M. J., Roehrborn, C. G., Cadeddu, J. A., & Pearle, M. S. (2004). Virtual ureteroscopy predicts ureteroscopic proficiency of medical students on a cadaver. *J Urol*, 172(2), 667-671.
- Parker, B. C., & Myrick, F. (2012). The pedagogical ebb and flow of human patient simulation: empowering through a process of fading support. *J Nurs Educ*, 51(7), 365-372. doi:10.3928/01484834-20120509-01
- Phipps, D., Meakin, G. H., Beatty, P. C. W., Nsoedo, C., & Parker, D. (2008). Human factors in anaesthetic practice: insights from a task analysis. *British Journal of Anaesthesia*, 100(3), 333-343.
- Popper, K. R. (1989). *Conjectures and refutations : the growth of scientific knowledge* (5th ed.). London ; New York: Routledge.
- Pusic, M. V., Boutis, K., & McGaghie, W. C. (2018). Role of Scientific Theory in Simulation Education Research. *Simul Healthc*, 13(3S Suppl 1), S7-S14. doi:10.1097/SIH.0000000000000282
- Ravert, P., & McAfoos, J. (2014). NLN/Jeffries Simulation Framework: State of the Science Summary. *Clinical Simulation In Nursing*, 10(7), 335-336. doi:10.1016/j.ecns.2013.06.002
- Rice, C. (2015). Moving Beyond Causes: Optimality Models and Scientific Explanation. *Noûs*, 49(3), 589-615. doi:10.1111/nous.12042

- Riskin, J. (2007). *Genesis redux : essays in the history and philosophy of artificial life*. Chicago: University of Chicago Press.
- Robertson, B. (2006). An obstetric simulation experience in an undergraduate nursing curriculum. *Nurse Educ*, 31(2), 74-78.
- Rosenberg, A. (2005). *Philosophy of science : a contemporary introduction* (Second ed.). New York: Routledge, Inc.
- Sawyer, T., Eppich, W., Brett-Fleegler, M., Grant, V., & Cheng, A. (2016). More Than One Way to Debrief: A Critical Review of Healthcare Simulation Debriefing Methods. *Simul Healthc*, 11(3), 209-217. doi:10.1097/SIH.0000000000000148
- Schiavenato, M. (2009). Reevaluating simulation in nursing education: beyond the human patient simulator. *J Nurs Educ*, 48(7), 388-394.
- Schoenherr, J. R., & Hamstra, S. J. (2017). Beyond Fidelity: Deconstructing the Seductive Simplicity of Fidelity in Simulator-Based Education in the Health Care Professions. *Simul Healthc*, 12(2), 117-123. doi:10.1097/SIH.0000000000000226
- Schwid, H. A., Rooke, G. A., Michalowski, P., & Ross, B. K. (2001). Screen-based anesthesia simulation with debriefing improves performance in a mannequin-based anesthesia simulator. *Teach Learn Med*, 13(2), 92-96.
- Seibert, D. C., Guthrie, J. T., & Adamo, G. (2004). Improving learning outcomes: integration of standardized patients & telemedicine technology. *Nursing Education Perspectives*, 25(5), 232-237.
- Shepherd, A. (1998). HTA as a framework for task analysis. *Ergonomics*, 41(11), 1537-1552.
- Siassakos, D., Hasafa, Z., Sibanda, T., Fox, R., Donald, F., Winter, C., & Draycott, T. (2009). Retrospective cohort study of diagnosis-delivery interval with umbilical cord prolapse: the effect of team training. *BJOG: An International Journal of Obstetrics & Gynaecology*, 116(8), 1089-1096.
- Stanton, N. A. (2006). Hierarchical task analysis: developments, applications, and extensions. *Applied Ergonomics*, 37(1), 55-79.
- Stayt, L. C. (2012). Clinical simulation: a sine qua non of nurse education or a white elephant? *Nurse Educ Today*, 32(5), e23-27. doi:10.1016/j.nedt.2011.06.003
- Steel, D. (2007). *Across the Boundaries: Extrapolation in Biology and Social Science*. New York: Oxford University Press.
- Stokes-Parish, J. B., Duvivier, R., & Jolly, B. (2017). Does Appearance Matter? Current Issues and Formulation of a Research Agenda for Moulage in Simulation. *Simul Healthc*, 12(1), 47-50. doi:10.1097/SIH.0000000000000211
- Sutton, R. M., Niles, D., Meaney, P. A., Aplenc, R., French, B., Abella, B. S., . . . Nadkarni, V. (2011a). "Booster" training: evaluation of instructor-led bedside cardiopulmonary resuscitation skill training and automated corrective feedback to improve cardiopulmonary resuscitation compliance of Pediatric Basic Life Support providers during simulated cardiac arrest. *Pediatric Critical Care Medicine*, 12(3), e116-121. doi:http://dx.doi.org/10.1097/PCC.0b013e3181e91271
- Sutton, R. M., Niles, D., Meaney, P. A., Aplenc, R., French, B., Abella, B. S., . . . Nadkarni, V. (2011b). Low-dose, high-frequency CPR training improves skill retention of in-hospital pediatric providers. *Pediatrics*, 128(1), e145-151. doi:http://dx.doi.org/10.1542/peds.2010-2105
- Tjiam, I. M., Schout, B. M., Hendrikx, A. J., Scherpbier, A. J., Witjes, J. A., & van Merriënboer, J. J. (2012). Designing simulator-based training: an approach integrating cognitive task

- analysis and four-component instructional design. *Med Teach*, 34(10), e698-707. doi:10.3109/0142159x.2012.687480
- Vaucanson, J. d. (1738). *Le Mécanisme Du Fluteur Automate: Présenté a Messieurs De L'académie Royale Des Sciences In. Paris: Jacques Guerin.*
- Walker, L. O., & Avant, K. C. (2005). *Strategies for theory construction in nursing* (4th ed.). Upper Saddle River, N.J.: Pearson/Prentice Hall.
- Weirich, P. (1986). Expected Utility and Risk. *The British Journal for the Philosophy of Science*, 37(4), 419-442.
- Wilson, R. D., & Hagler, D. (2012). Through the lens of instructional design: appraisal of the Jeffries/National League for Nursing Simulation Framework for use in acute care. *J Contin Educ Nurs*, 43(9), 428-432. doi:10.3928/00220124-20120615-27
- Wong, F. K. Y., Cheung, S., Chung, L., Chan, K., Chan, A., To, T., & Wong, M. (2008). Framework for adopting a problem-based learning approach in a simulated clinical setting. *Journal of Nursing Education*, 47(11), 508-514.
- Wood, G. (2002). *Edison's Eve : a magical history of the quest for mechanical life* (1st American ed.). New York: A.A. Knopf.
- Young, P. K., & Shellenbarger, T. (2012). Interpreting the NLN Jeffries Framework in the context of Nurse Educator preparation. *J Nurs Educ*, 51(8), 422-428. doi:10.3928/01484834-20120523-02