2019

Adaptation and Validation of the Situation Awareness Global Assessment Technique for Student Registered Nurse Anesthetists

Deniz Dishman
Virginia Commonwealth University

Adaptation and Validation of the Situation Awareness Global Assessment Technique

for Student Registered Nurse Anesthetists

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Virginia Commonwealth University.

by Deniz Dishman

DNAP Virginia Commonwealth University, August 2014
MSN Columbia University, August 2000
BSN Molloy College, August 1994
BA Hofstra University, May 1987

Director: Michael D. Fallacaro, DNS, CRNA, FAAN
Professor
Director of Special Projects
Department of Nurse Anesthesia

Virginia Commonwealth University Richmond, Virginia
January 2019

Acknowledgement

The author would like to thank several people without whom this dissertation could not have been completed. First and foremost, I would like to extend heartfelt thankfulness to my husband, Richard M. Dishman III, who always helped me make time for this work, always encouraged (and continues to encourage) this research journey without wavering in his support, and for the countless trips to Richmond and days (and nights) in hotels without complaint. I would like to thank my children, Danyal Brink and Richard M. Dishman IV, for all of their support and faith in my abilities. I would like to offer my profound gratitude to Dr. Michael Fallacaro for being someone I have and always will look up to, and for all his priceless insight and inspiration. Additionally, I would like to thank Dr. Suzanne Wright, Dr. Diane Dodd-McCue, and Dr. Nicole Damico for all of their teaching, guidance and encouragement. All of you at VCU Nurse Anesthesia are responsible for my drive for patient safety and research, and hold a high bar of professionalism and knowledge to which I continue to strive – you are truly a special bunch! Thank you to everyone at the VCU PhD in Health Related Sciences department for giving me the knowledge and tools to complete this research study. I must give special acknowledgement and sincere gratitude to Dr. Melanie Wright who graciously answered my email without knowing me in the slightest and generously allowed me to work on this research with a great deal of her support and guidance. I would be remiss if I failed to thank my colleagues at the UTH Houston Health Science Center, Cizik School of Nursing – Myron Arnaud, Matt Lewis, Peter Slivinski, and Amy Blok - for putting up with me (and all this work) for the past two and a half years. It has been so rewarding to work with and get to know you all. Thank you to the AANA Foundation for their generosity in providing funding for this study.

TABLE OF CONTENTS

List of Tables

List of Figures

Abstract

ADAPTATION AND VALIDATION OF THE SITUATION AWARENESS GLOBAL ASSESSMENT TECHNIQUE FOR STUDENT REGISTERED NURSE ANESTHETISTS

By Deniz Dishman, PhD, DNAP, MSN

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Virginia Commonwealth University

Virginia Commonwealth University, 2019

Major Director: Michael D. Fallacaro, DNS, FAAN
Professor, Director of Special Projects
Department of Nurse Anesthesia

Anesthesia is a health care specialty fraught with high workload demands, including stressful work environments, increased production pressure, work areas with many distractions, an increasing use of advanced technology, and the constant need to prioritize work actions. Proper patient management requires skillful clinical judgment particularly in this dynamic environment during anesthetized conditions. Effective clinical judgment includes not only appropriate interventions but also recognition that condition changes are occurring. Additionally, proficient clinical judgment must incorporate the ability to project what may occur secondary to actual or potential condition changes. These key elements operationalize situation awareness (SA).

Successful and safe anesthetic patient management requires high level SA to meet these workload needs. High level SA in student registered nurse anesthetists (SRNAs) is an important characteristic in the development of future, effective anesthesia providers. With Endsley's "Theory of Situation Awareness" as the foundation, the goal of this study was to adapt and validate the "Situation Awareness Global Assessment Technique" (SAGAT), according to her

recommendations, to quantify the SA of the nurse anesthesia graduate student (SRNA) subset of anesthesia trainees during the simulation of the induction of general anesthesia with the associated placement of an oral endotracheal tube.

After attaining IRB exempt review status approval, this study used purposeful sampling to identify a sample of CRNA, nurse educator subjects. An exploratory sequential mixed methods design was utilized. Delphi methods during qualitative data collection and validation used a seven-member sample. Following content analysis of these results, items for the adapted SAGAT were created. Quantitative methods applied to these items utilized data collected from a 40-member sample to determine item content validity and scale content validity indices (S-CVI/Ave. 0.92). Additionally, exploratory factor analysis was performed on these findings, providing further reliability with a Cronbach's alpha of 0.937.

This study's findings revealed that a SAGAT specific to the anesthesia domain and the nurse anesthesia graduate student subgroup was amenable to adaptation and validation. The resultant adapted and validated items from this study are appropriate and applicable for use with SRNAs during specific simulated exercises. These results have positive implications in SRNA education and training, and can also be extended to other anesthesia trainees and practicing providers. Additionally, this study provides support of the further adaptation, validation, and use of this instrument in other anesthetic content areas, as well as other health care domains.

Chapter One: Introduction

In 1999, the Institute of Medicine (IOM) delivered an agenda to rid the nation of an alarming crisis: preventable medical errors (Makary & Daniel, 2016). The IOM implicated all aspects of health care delivery in the United States as culpable in the alarming rates of death caused by medical error (Makary & Daniel, 2016). During that time, death rates related to medical error were estimated by the IOM at approximately 94,000 per year (Makary & Daniel, 2016). More recent estimates adjusted to capture as much available data as possible, place current cases closer to 440,000 deaths per year, positioning medical errors as the third leading cause of death in the nation (Makary & Daniel, 2016).

There is little evidence these shocking figures are discussed in mainstream groups or media outlets. Though annually more die by preventable medical errors than the equivalent of ten crashing jumbo jets full of passengers, these deaths occur one at a time in locations spread across the nation, making this an insidious epidemic (Reason, 1995). The release of the IOM's report marked the beginning of a national awareness of this alarming patient-safety problem, and many improvement measures have since been put into place throughout the health care system (van den Bos et al., 2011). There has been increased attention to infection prevention, the implementation of electronic records, more robust system designs, improved interdisciplinary team dynamics, and enhanced communication, for example. Reimbursement to health care providers is beginning to be tied to meeting quality improvement benchmarks. In the past two plus decades since the release of the IOM report, perhaps no other medical specialty has embraced the challenges of improving patient safety as greatly as the field of anesthesia (Chang et al., 2017).

Background of the Problem

Anesthesia is the medical specialty leading the way in the research in and the implementation of patient safety measures. Much of this work is directly linked to the safety work of high reliability organizations (HROs) in industries outside of health care (Gaba, 2000a). Pioneers in anesthesia patient safety appropriately link the complex system of anesthesia to that of other high risk, highly reliable, complex systems, and use the safety practices of these HROs in developing standards for anesthesia training and practice (Gaba, 2000a).

Improvements and advancements in medications and technology over the past 50 years have made anesthesia safe (Matveevskii & Gravenstein, 2008). In the wake of new anesthetic drugs, anesthesia techniques, advancing technology, and improved training, anesthesia related mortality risk has continually improved from approximately one death in 1000 anesthetics in the 1940s to an estimated one death in 100,000 procedures in the 1990 through early 2000s (Lampotang, 2008). More recently, the estimated anesthetic-related mortality rate has been reported at one death in 200,000 to 300,000 anesthetics (Chang et al., 2017; Morgan, Kurrek, Bertram, LeBlanc, & Przybyszewski, 2011; Morgan et al., 2015). Despite being increasingly reliable, anesthesia still carries the potential for serious injury and death. Anesthesia-related morbidity estimates indicate that more than one out of ten patients have an intraoperative medical error-related incident, and one out of 1000 experience an injury (Haller, Laroche, & Clergue, 2011).

Much of anesthesia safety efforts are focused at measures mirroring those of other high-risk industries, such as aviation and the military (Gaba, 2000b). Following organizational safety theories and human factors' specialists, anesthesia patient safety efforts are concentrated in areas such as technology, systems' design, simulation, organizational systems, and human factors

(Gaba, 2000b).

James Reason, a leading organizational safety researcher and theorist, demonstrates the commonalities of anesthesia and intensive care medicine with other high-risk industries, the latter of which were long the study of human factors specialists (Reason, 2005). He states that the shared attributes between these groups lie at the "sharp end", or the "human-system" and "doctor-patient interface", secondary to the complex, dynamic environments within which the respective operators function; and at the larger, organizational end in which these interactions at the "sharp end" occur (Reason, 2005). The interactions of anesthesia and intensive care medicine providers all occur within the "… complex, tightly coupled institutional setting" (Reason, 1995, p. 80).

Relating his "Swiss Cheese Model" of error to medical error, and more specifically anesthesia mistakes, Reason states that human error's contribution to accidents in these medical specialties is more attributable to opportunity lining up the weaknesses in established safety barriers than to provider carelessness or recklessness (Reason, 2005). He distinguishes human contribution to error with two terms: active and latent human failures. Active failures, according to Reason, are those acts causing immediate consequence committed at the "doctor-patient interface" and the "human- system" interaction (Reason, 1995). Latent failures begin at the organizational level and may not demonstrate adverse consequences for a long-time, until they by chance merge with other factors within the complex system (Reason, 2005).

By understanding these principles of human error, health care providers and organizations can take measures aimed at mitigating the arrangement of the "Swiss Cheese" holes that allow dissolution of defense mechanisms (Reason, 2005). According to James Reason, "… the goal of effective risk management is not to minimize particular errors and violations as to

enhance human performance at all levels of the system" (Reason, 1995, p. 85).

Since the time of Reason's publication, the anesthesia specialty has instituted mechanisms aimed at enhancing human performance at all levels, including: increased use of simulation in training; focus on improved systems' design; implementation of electronic records; communication and team dynamic training; and crisis resource management tools for high risk, low occurrence events. Other health care specialists such as surgeons, emergency medicine providers, emergency medical technicians, and nurses have followed suit and now embrace these practices (Schulz, Endsley, Kochs, Gelb, & Wagner, 2013; Schulz et al., 2015).

Human factors' engineering can be applied to patient safety efforts in areas related to organizational, systems, environmental, workload, and behavioral characteristics to improve patient safety and outcomes (Weinger & Gaba, 2014). Examples include: anesthesia machine system design with forcing functions; structured training using simulation to improve decision making in dynamic environments; incorporation of checklists to decrease reliance on memory; and training in communication and teamwork dynamics, particularly during handovers (Weinger & Gaba, 2014).

There is now an increased emphasis on these human-systems' interactions and human factors' engineering implications in error, causing the adoption of human factors' techniques to understand patient safety risk and associated system performance, including the use of simulation (Weinger & Slagle, 2002). In anesthesia and all other health care specialties, the focus of patient safety efforts should lie in "…understanding avoidable threats to patients due to individual and systems failures …" and the implementation and improvement of "… systems that will respond resiliently to non-routine operating conditions" (Weinger & Gaba, 2014, p. 801).

Situation Awareness

The relationship between human factors and error is drawn from organizational safety theory in HROs such as aviation and the military, and the importance of situation awareness (SA) is well established in these works. Situation awareness, or having a knowledge and understanding of what is happening in the environment, is a key construct in anesthesia safety (Gaba, Howard, & Small, 1995). James Reason's publication demonstrating the applicability of his organizational safety theory to medicine and the anesthesia specialty comes soon after that of David Gaba, an anesthesia safety pioneer. Gaba and his team were the first to highlight the importance of ensuring anesthesia providers attain high levels of SA (Gaba et al., 1995). Mica Endsley, a founding researcher on SA in aviation and the military, maintains that high levels of SA are a key determinant of decision-making, particularly in dynamic environments, such as that found in anesthesia (M. Wright, Taekman, & Endsley, 2004).

Attaining and maintaining high levels of SA is recognized as critical for anesthesia providers. Situation awareness is key to proper clinical management, maintenance of patient safety, and minimization of medical error (Shelton, Kinston, Molyneux, & Ambrose, 2013). This is particularly true for anesthesia providers who work in an environment of ever increasing information load, high task requirements, advanced technology with multiple levels of data, and a rapidly changing situation (Wright et al., 2004; Wright, 2015).

Situation awareness is defined as, "… the perception of elements of the environment within a volume of time and space, the comprehension of their meaning and the projection of their status in the near future." (Endsley, 1995a, p. 5) During World War I, the importance and role of SA was recognized in the course of military aviation missions (Manz, Hercinger, Todd, Hawkins, & Parsons, 2013). Military pilots' ability to maintain mindfulness of enemy aircraft in

the dynamic environment of air flight was recognized as a vital characteristic (Manz et al., 2013).

Situation awareness is described in three ordered levels (McKenna et al., 2014). Level 1, also known as perception, encompasses the ability to collect relevant information from the dynamic events occurring within one's environment. Comprehension is the state of SA at Level 2, incorporating the integration of the relevant information perceived in Level 1 with formulation of an understanding of its meaning and importance in the current situation. Level 3, projection, is the highest level of SA and reflects the ability, based on the perception and comprehension of the current state, to anticipate potential imminent events and the consequences that may develop. Level 3 SA is vital for appropriate decision-making, team coordination and communication, and work flow management. Table 1 depicts the hierarchical levels of SA.

In high-risk industries such as aviation, department of defense, transportation, health care, and nuclear power for example, SA is a well-described necessity: the personnel in these environments must be aware of conditions in their vicinity at any given moment (Manz et al., 2013). High level SA is particularly important in these areas, as there is great potential for injury and death. The increasing workload secondary to advanced technology, now a foundation in these industries, potentiates the need for high level SA (Tolley, Marks-Maran, & Burke, 2010). Possessing full SA requires the ability to project future events and react to them accordingly (Tolley et al., 2010).

Importance of situation awareness in anesthesia. An anesthetized patient is typically in an uneventful, steady state with his/her physiological condition changing very slowly over the course of the procedure (McKenna et al., 2014). During a rare critical event however, the patient's physiological state can change rapidly and unexpectedly. Unless recognized and managed correctly, these high-risk, low occurrence

Table 1

Levels of Situation Awareness (SA)

SA	Characteristics
Level 1 Perception	Collect relevant data in dynamic environment
Level 2 Comprehension	Formulate contextual meaning and importance of data
Level 3 Projection	Project potential imminent events and consequences

Note. Definitions of characteristics for levels of SA from McKenna et al. (McKenna et al., 2014).

events can lead to severe injury and death. High level SA must be ever present in an anesthesia provider to properly recognize, successfully manage, and project potential future events and consequences during changes in the usual static state of anesthesia (McKenna et al., 2014). High level, or Level 3 SA in anesthesia providers is paramount to patient safety and effective teamwork (McKenna et al., 2014).

Statement of the Problem

Situation awareness develops with experience and training, and this is particularly true of anesthesia providers (Schulz et al., 2016). Endsley underscores the relationship of the ability to rapidly attain and maintain high level SA with experience and training (Kaber & Endsley, 2004). Student nurse anesthetists (SRNAs), for example, have limited experience in anesthesia-related

patient care situations, but do have experience as critical care nurses from which they can draw upon in certain critical situations. It is from these experiences that they can potentiate Level 1 SA and conceivably further develop higher levels of SA with proper anesthesia education and training (Blandford & Wong, 2004).

Anesthesia is a health care specialty fraught with high workload demands, including stressful work environments, increased production pressure, work areas with many distractions, an increasing use of advanced technology, and the constant need to prioritize work actions (Wetmore et al., 2016). Proper patient management requires skillful clinical judgment, particularly in this dynamic environment during anesthetized conditions (Schulz et al., 2013).

Effective clinical judgment includes not only appropriate interventions but also recognition that condition changes are occurring (Schulz et al., 2013). Additionally, this clinical judgment must incorporate the ability to project what may occur because of actual or potential condition changes. These key elements of clinical judgment operationalize situation awareness (Wright et al., 2004). Successful and safe anesthetic patient management requires higher order SA to meet these workload needs (Chang et al., 2017).

Simulation in anesthesia training. Beginning in the 1960s with the advent of "Resusci-Anne" and "SimMan", simulation training has long been recognized as an educational tool for the improvement in practitioner performance, clinical management skills, team dynamics, and overall patient outcomes (Cooper & Taqueti, 2004). Simulation is increasingly being used in anesthesia education and training to provide practice in clinical management without actual patient harm. Emulating the practices of other complex system industries such as aviation, its use is focused in education and technical skills training, as well as minimizing the factors associated with medical error including: poor communication; work load and task prioritization; and systems' design, with an emphasis on technology's effect on performance (Chang et al., 2017).

Recent evidence demonstrates that utilizing simulation in health care provides important education regarding all aspects of patient management (Sollid et al., 2016). For more than a decade, simulation's use in anesthesia education and training is increasingly aimed at providing practice in patient management (Sollid et al., 2016). Beginning with the work of Gaba in 1995, simulation use in education and training continues, with its current focus on the education and training of human factors' techniques and their relation to medical error and patient safety (Chang et al., 2017).

Based on the growing evidence of simulation training's positive impact on patient safety and optimal clinical management, the Standards for Accreditation of Nurse Anesthesia Programs – Practice Doctorate (Council of Accreditation of Nurse Anesthesia Programs [COA]), Curriculum Standard 11, requires that "Simulated clinical experiences are incorporated into the curriculum" (Council on Accreditation of Nurse Anesthesia Educational Programs, 2018, p. 22). To support the incorporation of these changes in nurse anesthesia curricula, valid measures are needed to document the impact of simulation on training and performance (Wright et al., 2004).

Assessing situation awareness. Situation awareness assessment has been attempted by both direct and indirect approaches (Orique & Despins, 2018). Direct measurement refers to those methods that purport to quantitatively assess an individual's SA (Orique & Despins, 2018). These instruments assess the subject's recognition of developing events, and include both objective and subjective techniques (Orique & Despins, 2018). Indirect methods measure behavioral and performance outcomes, and thereby infer an individual's SA (Orique & Despins, 2018). As the importance of SA in health care providers grows and is now highlighted in current education and training, particularly for those in the anesthesia specialty, a means to measure this construct is of paramount importance.

To use simulation in the evaluation of the skills and training of anesthesia providers, an objective measurement tool of higher order SA, the foundation of sound clinical judgment and decision-making, is necessary (Wright et al., 2004). Finding instruments that quantify performance by computing level of SA, or direct measures, during simulation is important but difficult, with a limited array of existing tools (Wright et al., 2004). Currently available measurement tools are subjective, using expert rater observations of behavior as the basis for assessments or utilize retrospective self- report (Wright et al., 2004). These methods also bear high costs as they employ experts to observe and assess performance. At this time there are no direct, objective tools by which SA can be measured in anesthesia trainees.

Situation Awareness Global Assessment Technique (SAGAT). The Situation Awareness Global Assessment Technique (SAGAT), developed by Mica R. Endsley, is an SA measurement tool intended to directly and objectively measure SA (O'Brien & O'Hare, 2007; Orique & Despins, 2018). Most other SA measurement instruments are subjective, using rater observations of behavior or retrospective self-report (Lavoie, Cossette, & Pepin, 2016).

Previously validated and applied to many domains, this measurement tool utilizes queries related to unfolding events and is administered during a simulated patient care scenario (Orique & Despins, 2018).

This assessment tool involves questioning subjects at predetermined freeze times during a developing situation, hence relegating its application to simulation, which presents its main drawback (O'Brien & O'Hare, 2007). The subject is asked questions directly related to events as they unfold, and these questions are purported to measure perception of pertinent data as well as higher levels of SA. Developing goal specific questions that reflect appropriate levels of SA is another limitation of this instrument (Najjar, Docherty, & Miehl, 2016). To date, there is no SAGAT instrument developed to assess SA in anesthesia trainees (SRNAs or physician residents).

Purpose of the Study

The purpose of this study is to validate a modified Endsley's Situation Awareness Global Assessment Technique to quantify the SA of nurse anesthesia graduate students during the simulated induction of general anesthesia with the associated placement of an oral endotracheal tube. The primary goal of this study is to adapt and validate the SAGAT, a direct and objective measurement tool used in other high-risk industries, for use as a means to quantify nurse anesthesia graduate students' SA during a specific anesthesia simulation. This tool will give educators the ability to quantify for the first time an anesthesia trainee's level of SA. This instrument will highlight the decision- making processes of SRNAs during a specific period of patient management, to which training and education efforts can be focused.

Research Question

Can Endsley's SAGAT be adapted and validated to quantify the SA of nurse anesthesia

graduate students, or SRNAs, during the simulation of a specific anesthesia event?

Significance of the Study

Through expert consensus, the SAGAT is adapted and validated for use with SRNAs during the simulation of the induction of general anesthesia with associated insertion of an oral endotracheal tube. Once adapted and validated, application of this instrument as a formative assessment tool can benefit nurse anesthesia education programs. The SAGAT can also be used as a summative assessment tool to quantify learning at the culmination of training. Understanding the level of SA in SRNAs, and any impact simulation training and didactic education has on SA, can thus steer successful SRNA curricula. The SAGAT can be used to assess groups of SRNAs across different programs as well as individually.

Theoretical Framework

The most widely cited theory underpinning SA is by Mica R. Endsley and presented within her 1995 publication, *Toward a Theory of Situation Awareness in Dynamic Systems* (Endsley, 1995a; Orique & Despins, 2018). In this work, Endsley describes three levels of SA that impact appropriate decision-making and performance in dynamic environments; she further outlines the destructive effect of inaccurate SA on outcomes; and the impact it has on effective team cohesion and dynamics (Endsley, 1995a). Since its debut, the theory of SA has been heralded in human factors' research as the backbone of safety and effective situation management (Wickens, 2008).

Situation awareness is a component of a feedback system in which decisions are made based on observed data and predicted outcomes of events (Endsley, 1995a, Endsley, 2015). As decisions are made and executed, the outcomes further impact data and continued decision-making such that every taken action has an impact on the situation and observable data, and subsequently the next planned action (Schulz et al., 2013). Decisions are based on past experience and training, and the execution of these decisions are based on abilities, stress, and

workload (Schulz et al., 2013). The feedback system of situation awareness as described by Endsley's (1995a) "Theory of Situation Awareness" is depicted in Figure 1.

Applying Endsley's theory to the health care domain, SA becomes a critical determinant of clinical judgment (Lavoie et al., 2016). Clinical decision-making can be related to the levels of SA as Endsley describes them: perception, comprehension, and projection (Wright et al., 2004). As an extension of this process to the anesthesia domain, and specifically the SRNA subset, sound patient management decisions are found in higher order SA (Level 3). Higher order SA is crucial to effectively manage patients in a dynamic physiological state (such as when anesthetized) within a dynamic environment (i.e., the operating room). Table 2 correlates

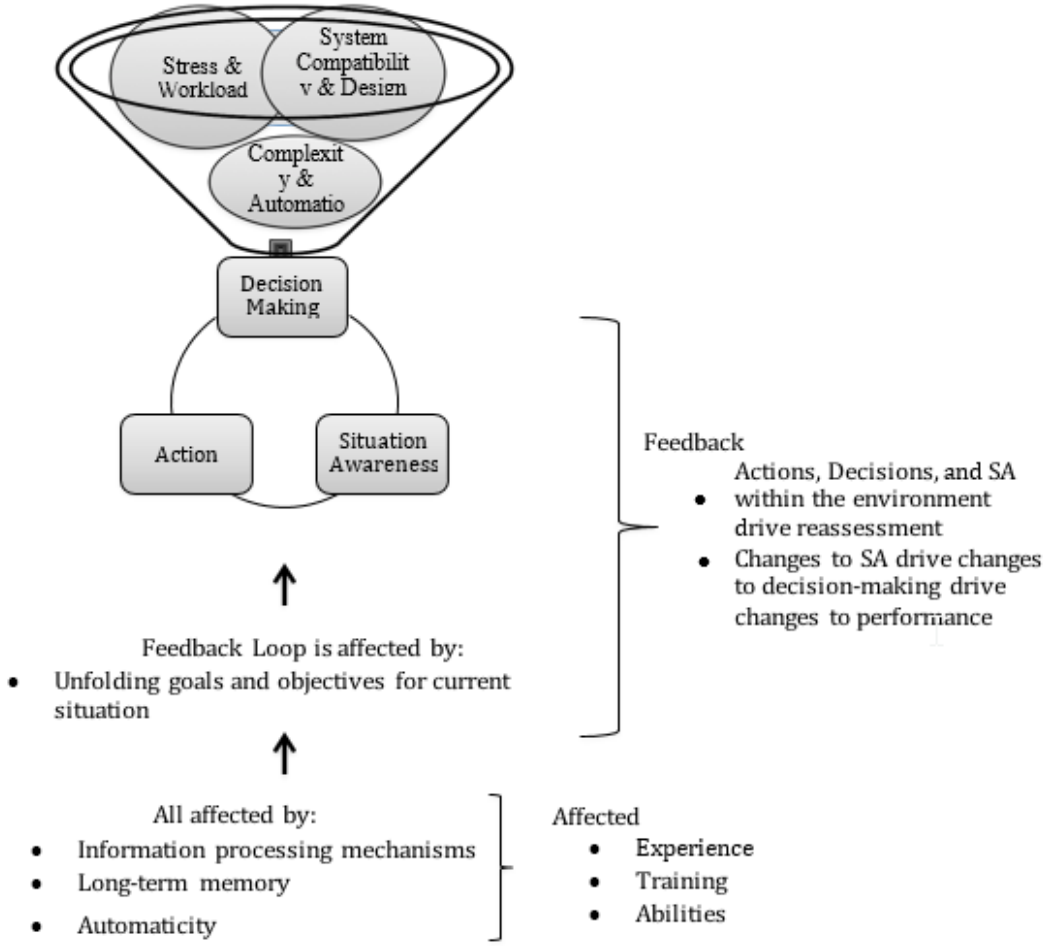


*Figure 1.*Depiction of feedback system described in Endsley's "Theory of Situation Awareness" (Endsley, 1995a). Data for the depiction from Endsley (1995a).

appropriate SRNA clinical judgment to levels of SA. Situation awareness determines the clinical judgment of SRNAs at each level in this population.

Table 2

Levels of SA in Relation to SRNA Level of Clinical Judgment

Level of SA	SRNA Clinical
Correlation Level 1: Perception	Recognition of abnormal cues:
i.e.	abnormal heart rate/rhythm, blood pressure, oxygen saturation, blood loss
Level 2: Comprehension	Drawing on past experience, i.e. didactic and simulation education, critical care nursing experience, understand the crisis or events abnormal cues represent
Level 3: Projection	Take measures to stabilize patient; Plan and execute proper clinical management of critical event, anticipate unfolding events and potential sequelae from actions taken.

Note: Clinical correlation to SA level adapted from Melanie Wright et al. (2004), *Objective measures of situation awareness in a simulated medical environment* (Wright, et al., 2004).

Definition of Terms

Ergonomics."… an applied science concerned with designing and arranging things people use so that the people and things interact most efficiently and safely – called also human engineering, human factors engineering." ("Ergonomics | Definition of Ergonomics by Merriam-Webster," n.d.).

Situation awareness. "The perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future." (Endsley, 1995a, p.5)

Simulation. "… a technique – not a technology- to replace or amplify real experiences with guided experiences that evoke or replicate substantial aspects of the real world in a fully interactive manner" (Gaba, 2007, p. 12).

Nurse anesthesia/nurse anesthetist. Nurse anesthesia is the practice of anesthesia by an advanced practice nurse with graduate level training. Nurse anesthetists have the ability to provide all forms of anesthesia in any setting. ("CRNA Fact Sheet," American Association of Nurse Anesthetists [AANA], 2018).

Student registered nurse anesthetist (SRNA). A registered nurse with a minimum of a bachelor's degree in nursing or a related field, one year of intensive care nursing experience and enrolled in a masters or clinical doctorate program of nurse anesthesia accredited by the Council on Accreditation of Nurse Anesthesia Programs (COA, 2018).

Formative assessment. A measurement tool in which learning measurement is used to formulate teaching plans and learning activities (Krage & Erwteman, 2015).

Summative assessment. A measurement tool that quantifies learning at the culmination of education/training; most often designated by a grade that affects the student/trainee's ability to progress within a program (Krage & Erwteman, 2015).

Expert. "… having, involving, or displaying special skill or knowledge derived from training or experience." (McPherson, Reese, & Wendler, 2018, p. 405)

Consensus. "… a general agreement; the judgment arrived at by most concerned" (McPherson et al., 2018, p. 405).

Induction of general anesthesia. The act of bringing about a "… drug-induced, reversible condition that includes specific behavioral and physiological traits – unconsciousness, amnesia, analgesia, and akinesia – with concomitant stability of the autonomic, cardiovascular, respiratory, and thermoregulatory systems" (Brown, Lydic, & Schiff, 2010, p. 2638).

Oral endotracheal tube. A medical device, most of which are made of polyvinylchloride (PVC), that can be placed through the mouth into the trachea during general anesthesia; provides the anesthesia provider the ability to support the patient's respirations during general anesthesia.

Assumptions

This study was conducted using the Delphi method to elicit expert consensus and collect expert opinion for the content reflected in the items generated for the adapted SAGAT. This expert consensus was also used for the assessment of relevancy of the created items with respect to an anesthesia provider's required SA for decision-making. Those surveyed in this study were nurse anesthesia program directors and any other nurse anesthesia program faculty, with a minimum title of assistant professor, whose contact information could be found publicly from university, web-based resources, or from the public listing found on the COA website. It was assumed that the sample of nurse anesthesia educators' opinions were reflective of those of all

anesthesia providers, particularly nurse anesthetists. Also, as this study was conducted by electronic means, specifically email, it was assumed that the intended recipients were indeed those who answered the surveys.

Conclusion

This chapter highlights the importance of patient safety in health care, with a particular emphasis on the anesthesia specialty. The link between HROs and organizational safety theory to anesthesia is presented. Discussion focuses on SA as an important construct and its role in anesthesia patient safety. The importance of simulation in patient safety endeavors is also presented. The foundation for the need of a direct, and objective SA assessment tool in anesthesia simulation is introduced.

The following chapter presents a review of the literature with a focus on these highlighted topics. The appropriateness of Endsley's, "Theory of Situation Awareness" (1995) underlying the construct of SA is established. The SAGAT instrument is explored in depth in relevant literature. Additionally, the increasing use of the SAGAT as a quantitative measure of SA in health care specialties is highlighted. Identification of the lack of a SAGAT instrument adapted and validated for the anesthesia domain is demonstrated.

Using this literature review as the foundational support for this study, the remaining chapters directly describe the research study design, methods, and results. Chapters Three and Four articulate the methods and results of this research study in accordance with the accepted standards for reporting qualitative research (Tong, Sainsbury, & Craig, 2007). These criteria are applicable to this study as the design, though mixed methods, is heavily weighted in qualitative methods. The latter of the aforementioned chapters will present the study outcomes according to accepted reporting guidelines for both qualitative and quantitative results, as additional

validation of the findings are provided by quantitative measures. The final chapter of this publication discusses the study's results in terms of feasibility, reproducibility, and generalizability to other anesthesia trainees and perhaps providers; identifies the limitations of this work, and gives suggestions for further research.

Chapter Two: Literature Review

Over a decade since the Institute of Medicine's (IOM) report on the state of patient safety in health care, human factors as a root cause of medical error is now well established. Training with simulation is currently part of the health care education curriculum across many specialties in an effort to minimize the effects these constructs have on patient safety. The increasing use of simulation in health care as a training tool grows from its positive impact on operations in other complex industries, such as aviation, nuclear power, and the military.

Simulation is used in health care training to improve outcomes in high-risk, low occurrence events, for provider and or trainee assessment, and to practice routine situations and system operations. More so than mechanical failures, human factors such as system design and an operator's cognitive and physical capabilities are implicated in preventable accidents in complex systems, including complex health care domains such as anesthesia. For greater than a decade, there is an increasing use of simulation for the assessment of and training in the human factors' components of performance.

This chapter provides the link between adept clinical decision-making and patient safety in anesthesia: human factors and safety in anesthesia are discussed, as is the role of simulation in health care and anesthesia education and training, with respect to human factors and error prevention. These discussions form the basis for the review of literature of SA in anesthesia, a crucial human factors' construct necessary for proficient clinical decision-making and positive patient outcomes. Reviewing the current literature leads to the examination of the importance of measuring SA, and a comparison of available SA measurement tools. The most widely accepted, applied, and cited theory of SA, Mica R. Endsley's "Theory of Situation Awareness", is presented. This chapter concludes with a thorough description of "The Situation Awareness

Global Assessment Technique", or SAGAT, one of the few available direct and objective SA measurement tools.

Complex Systems

Aristotle is credited with the basis for the adage "the whole is greater than the sum of its parts", with its origins attributed in his writing, "The totality is not, as it were, a mere heap, but the whole is something besides its parts", as translated in his work, *Metaphysics* (Metaphysics,VIII, 1045a8-10). This statement demonstrates the most simplistic definition of a complex system. The IOM defines complex systems as those that have many parts that interact with each other and are tightly coupled: these system fragments have interdependencies, interrelationships, and interactions among them and with the systems' environment (Donaldson, Corrigan, & Kohn, 2000). The IOM further defines complex systems as being "non-linear" as they have many feedback loops and are highly specialized (Donaldson, et al., 2000).

Health care as a complex system. Peter Petros (2003) describes all of medicine as made up of complex systems, and offers evidence from biological processes that interactions of parts, no matter the system, are not linear in fashion. Petros states that in these complex systems, one plus one does not equal two because of the prevalence of feedback mechanisms (Petros, 2003). The nature of these systems favors a high tendency towards potentially lethal or disabling activities and outcomes (Gaba, 2000a). The dynamic nature and complexity of these systems place a large weight on the cognitive abilities of operators (Gaba, 2000a). Because of the requisite increased performance capabilities within these systems, the important role of human factors in aptitude is underscored (Gaba, 2000a).

James Reason (1995) defines health care as a complex system as it shares many traits of other high-risk industries. He describes health care as being a complex, tightly coupled system

with multiple interactions between different groups (Reason, 1995). Health care in the United States is a high-risk system, on par with the aviation, military, and nuclear power industries. All of these systems share high-risk contexts with many points of highly active and changing parts: none are straightforward; all have a high degree of inter-relationships; and are each tightly coupled (Wright, 2015).

Anesthesia as a complex system. The specialty of anesthesia is identified as comparable to other industries that are considered complex systems, such as aviation, nuclear power control, and the military: it is a system with interdependencies, interrelationships, and interactions within its integral parts and within the environment. Anesthesia is a system of high information load, multiple task requirements, a potential for rapidly changing situations and conditions, and an inherent increased risk (Glavin & Flin, 2012). It is a health care specialty fraught with high workload demands, including stressful work environments, increased production pressure, work areas with many distractions, increasing use of advanced technology, and the constant need to prioritize work actions (Chin & Lagasse, 2017). It necessitates high cognitive ability and task demands, a slim margin of error with the potential for severe adverse outcomes, and crucial decisions made with sometimes inadequate or unavailable information, all within a rapidly changing environment (Glavin & Flin, 2012).

A patient is exposed to high risk with every moment he or she is subjected to the complex system of anesthesia (Wright, 2015; Gaba, 2000b). David Gaba (2000b) describes patients themselves as complex systems with every physiologic process potentially or inevitably influencing other processes. As Gaba illustrates in this review of patient safety in anesthesia, the necessity of interaction between these two complex systems increases the potential for mortality and morbidity, and further demonstrates the impact human factors have on outcomes (Gaba,

2000b).

Anesthesia likened to aviation. Holzman et al. (1995) assert that the anesthesia provider is working in a complex system: multi-factorial, "tightly coupled", and high task environment. In this publication, they describe their "Anesthesia Crisis Resource Management" ("ACRM") course for anesthesia providers: educational offerings aimed as an aid to learn and practice skills necessary in successfully managing high risk, low occurrence events (Holzman et al., 1995). "ACRM" is the author's adaptation of crisis resource management teachings utilized in the aviation industry. The authors contend that anesthesia is unlike any other health care specialty, with its successful practice based on "event-driven decision making" (Holzman et al., 1995). They give support to the similarities between the anesthesia specialty and aviation, with the need for specific cognitive and physical abilities, recognition and adaptation to an ever-changing environment, and the potential for catastrophe should an error occur (Holzman et al., 1995).

Zausig et al (2009) present the implications of non-technical skills training for anesthesiologists and its effect on patient outcomes in a randomized controlled trial. The authors attempt to link, through the use of simulation, the implications of non-technical skills training for anesthesiologists and patient outcomes (Zausig et al., 2009). Within this work, they contend that anesthesia is a "high-risk, complex work system" comparative to aviation (Zausig et al., 2009).

Human error in complex systems. The IOM report highlights that complex systems in industries with inherent risk are prone to accidents and errors mostly because of human error (Donaldson, Corrigan, & Kohn, 2000). Even the slightest failure in any part of the system, because of its tight linkages, can cause catastrophic effects in another part of the system. Additionally, Endsley and Kiris (1995) point out that as technological advances increase in these systems, automaticity increases the risk of a disastrous event. The further removed the provider, or operator, is from the system, the authors contend, the more likely he or she will be unable to detect a problem (Endsley & Kiris, 1995).

Preventable Medical Errors

Makary and Daniel (2016), provide a study extrapolating established data regarding preventable medical error in the United States (Makary & Daniel, 2016). The authors state that medical errors occur at an alarming rate, drawing evidence that places medical error as the third leading cause of death in this country (Makary & Daniel, 2016). Presenting the 1999 IOM report *To Err is Human,* the authors implicate preventable medical errors as contributing to the deaths of over 90,000 patients per year (Makary & Daniel, 2016). It also includes more recent studies, such as the 2004 report by the Agency for Healthcare Quality and Research (AHRQ) Patient Safety Indicators in Medicare patients, that place this number closer to 194,000 annually (Makary & Daniel, 2016). Recently, Claussen et al. reported a 1.13% rate of death caused by medical error (Makary & Daniel, 2016). Applying this rate to the 2013 number of hospital admissions in the United States, Makary and Daniel (2016) show medical errors as the root of preventable death at a rate of over 400,000 deaths per year. The authors also provide evidence that the medical errors associated with these preventable deaths find their root cause in human factors, not progression of illness or technical failures (Makary & Daniel, 2016).

Jones, Corbett, Morton, Lister, and Mercer (2018) report statistics from the United Kingdom demonstrating that human factors account for one error in every 133 anesthetics. The authors cite two prospective studies conducted by the National Audit Project in the United Kingdom, examining major airway events and occasions associated with awareness under anesthesia (Jones et al., 2018). Human factors are implicated at the root of these cases, specifically lack of communication and poor team dynamics, as well as poor or absent situation awareness (Jones et al., 2018).

Preventable anesthesia-related error. In his 2000 paper regarding patient safety in anesthesia, David Gaba explains the difficulty in estimating the number of preventable deaths related to medical errors in the delivery of anesthesia (Gaba, 2000b). This is, as Gaba offers, in part due to the difficulty in estimating the total number of anesthetics given in the United States annually, and the use of different methodological constructs and design techniques in previously reported studies (Gaba, 2000a). However, by studying trends, an estimate of anesthesia-related medical errors can be extrapolated (Gaba, 2000a). Estimates from an epidemiological study of data from 1999 to 2005 implicate anesthesia as the direct cause of 34 deaths in the United States per year, and as a contributing factor in 281 additional deaths per year (Li, Warner, Lang, Huang, & Sun, 2009).

Fortunately, improvements and advancements in medications and technology over the past 50 years have made anesthesia very safe (Matveevskii & Gravenstein, 2008). In the wake of new anesthetic drugs, anesthesia techniques, advancing technology, and improved training, anesthesia related mortality risk has continually improved from approximately one death in 1000 anesthetics in the 1940s to an estimated one death in 100,000 procedures in the 1990s to early 2000s (Lampotang, 2008). Despite an exemplary safety statistic, anesthesia related morbidity

remains an issue in health care and patient safety (Chin & Lagasse, 2017). The increasing complexity of the anesthesia environment secondary to technological advances and automation add to the increased risks to patient safety and the commission of preventable medical errors (Holzman et al., 1995).

Matveevskii and Gravenstein, (2008) present anesthetic related morbidity as divided into three levels of injury. Minor injury refers to those that do not excessively prolong hospital stay or cause permanent complications, for example post-operative nausea and vomiting. Intermediate injury applies to those events that increase hospital stay without causing permanent complications such as dental injury. Serious injury includes those events that increase hospital length of stay and cause permanent injury, i.e. nerve injury resulting in loss of function (Matveevskii & Gravenstein, 2008).

Data from 2011 yields estimates of minor injuries occurring in 18 to 22% of anesthetics, and serious injury in 0.45 to 1.4% of anesthetics (Chin & Lagasse, 2017). Severe, negative outcomes with permanent injury occur in one of every 170 to 500 cases (Staender & Mahajan, 2011). A study by the AANA Foundation (2015) of malpractice closed claims between 2003 and 2012 (n = 245), found that the most common adverse events leading to liability claims are: death (35.1%), respiratory causes (31.8%), and central nervous injuries (41.6%) (Jordan & Quraishi, 2015). The authors of the study estimate that 45.5% of the adverse events cited were preventable (Jordan & Quraishi, 2015). The most common injuries underpinning malpractice claims as evidenced by the American Society of Anesthesiologists (ASA) Closed Claim Project (2009, n=9,214) are: death 29% nerve injury 21%, permanent brain damage 9%, airway injury 6%, mental duress 6%, and eye injury 4% (Schulz et al., 2014). Notably, dental injuries, an intermediate level of injury, are excluded from this data.

Human Factors in Medical Error

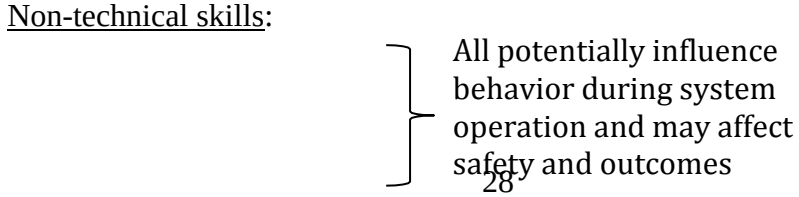
To Err is Human: Building a Safer Health System, a report by the IOM and its follow-up report, *Crossing the Quality Chasm,* call attention to the alarming rate of medical error in the United States, patient safety concerns, and highlight the human component of medical errors (Corrigan, 2005; Donaldson et al., 2000). Considerable amounts of organizational safety research are based on the IOM's report, making attempts to model safety efforts after those seen in HROs. James Reason presents a "systems approach" to patient safety, citing the example of HROs, such as nuclear aircraft controllers, air traffic control systems, and nuclear power plants (Reason, 2000). He offers that these systems are reliable as "…safety is preserved by timely human adjustments…" (James Reason, 2000, p. 395).

James Reason offers guiding work in studying organizational safety in his hallmark publication *Human error: Models and Management* (Reason, 2000). The "Swiss Cheese Model" he presents demonstrates how active and latent errors can find weaknesses in defense barriers within organizational systems, including health care (Reason, 2005). Reason draws the comparison of human factors in complex industries to those in the complex system of health care (Reason, 2005). A synopsis of different organizational safety theories is found in work by David Gaba (Gaba, 2000b), who highlights James Reasons' "Swiss Cheese Model" of error, in which active and latent failures in complex systems line up in such a way as to present "holes" in defense barriers that are in place; and Charles Perrow's "Normal Accident Theory", relating the complexity of a system and its "tight coupling" of interactions and sub-systems to the occurrence of error or accidents as the end result of a run-off of these interactions. Failures or lapses of error are identified as the human factors implicated in preventable medical errors and poor or unintended patient outcomes (Gaba, 2000b; Jones et al., 2018).

Non-technical skills (NTS). Human factors refer to the "… psychological, social, and physical nature of human beings and the system(s) in which they function." (Wright, 2015 p.4). Figure 2 depicts the elements commonly associated with human factors. These elements include task and time pressures, stress level and tiredness of the operator, and an awareness of all the elements within the environment in which the system(s) are operating – termed situation awareness (Wright, 2015). Human factors are constructs, not skills or tasks that can be learned by performance. As constructs, these notions consist of many elements, some without tangible means of measurement or expression.

Figure 2. Components of human factors data for depiction from Wright (2015).

The human factors of medical error have come to be known in the medical field as "non-technical skills", or NTS, a misnomer imparting the fallacy that these are skills or tasks that can be mastered and checked off from a list, akin to placing an intravenous line or an oral endotracheal tube. Flin, Patey, Glavin, and Maran (2010) distinguished NTS from "technical errors", as described in other complex system industries. Non-technical skills in health care are historically accepted as including: "situation awareness, decision- making, team-work, communication, and the management of stress and fatigue" (Flin et al., 2010, p.39). Figure 3 represents those elements recognized as components of non- technical skills.



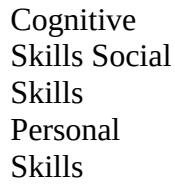


Figure 3. Components of non-technical skills. Data for depiction from Flin, Patey, Glavin, and Maran (2010).

Behavioral performance. David Gaba et al. (1998) assert that successful, safe anesthesia requires technical performance as well as behavioral performance. These authors define technical performance as a demonstration of competence from a medical and technical perspective. Competence in technical performance can be assessed by traditional written examination and observation of clinical management (Gaba et al., 1998). Behavioral performance relates to the decision-making and quality of interactions with other team members during patient care management (Gaba et al., 1998). This behavioral component of performance is critical for competent clinical decision-making and adaptation to changing conditions (Endsley, 1995b; Gaba et al., 1998; Schulz et al., 2013).

The 1978 publication of Cooper et al. implicates "human factors" in anesthesia error at 70-80% (Cooper, Newbower, Long, & McPeek, 1978). Since the time of this publication, the study and acceptance of human factors in anesthesia related error is well established (Schulz et al., 2013; Zausig et al., 2009). The behavioral performance indicators put forth by David Gaba et al. (1998) are developed and formally defined in Flin et al.'s 2010 work, "Anesthetist's Non-Technical Skills Handbook" (Flin et al., 2010; Gaba et al., 1998). Four categories of NTS are established in this publication: task management, team working, situation awareness, and decision-making (Flin et al., 2010).

As in aviation, NTS are implicated in patient outcomes, whether positive or negative

(Morgan et al., 2011). Preventable errors in anesthesia are intertwined to human factors (Gaba, 2000b). Human error and team dynamic breakdown are often cited as the cause of poor outcomes during the perioperative period (Matveevskii & Gravenstein, 2008; Gaba, 2000b). Since David Gaba introduced the concept in the 1989 publication, "Human Error in Anesthetic Mishaps," NTS grew to be the focus of attaining positive patient outcomes and the subject of patient safety research for greater than 20 years (Gaba, 1989; Wright, 2015). With increasing automation and advances in technology, human factors research is timelier now than ever before.

Over the past decade, a number of research studies have highlighted the impact of the teaching of NTS on clinical outcomes. Studies such as the randomized control trial conducted by Yule et al. (2015) demonstrated that briefing anesthesia trainees (physician residents) regarding NTS before the management of a simulated event improved outcomes over the control group that did not receive the NTS coaching. More recently, Hagemann et al. (2017) demonstrate similar results in a randomized, double blind trial with a pretest and posttest design. The authors show a positive impact on outcomes in a simulation event with a training session on NTS as the intervention (Hagemann et al., 2017).

Simulation

From these decades' long interest and focus in human factors and patient safety, simulation in health care emerges as a large component of training and education in many medical specialties. Medical training using simulation is borrowed from aviation, as well as other complex HROs, emulating their safety and training practices. Simulation is a well-established training tool in aviation, nuclear power control, and military operations. Looking back at the history of simulation in health care, its evolvement from the 1960s into a training tool is linked to the imitation of safety programs in other complex, high- risk systems (Cooper & Taqueti, 2004).

From the fleeting use of "Sim One" by Denson and Abrahamson in the 1960s comes the integration of computer adaptive technology with life-like mannequins in the two ground breaking anesthesia high fidelity simulators: one created by David Gaba in 1987 and the other by Michael Good, developed almost simultaneously (Cooper & Taqueti, 2004).

Simulation in anesthesia. Since these early anesthesia simulators, the use of simulation training to improve behavioral performance in clinical management is now well established. In 1991, David Gaba provided an editorial in the ASA's journal, *Anesthesiology*, stating, "… it is unreasonable to expect anesthesia trainees to learn sensible responses to critical events purely by 'osmosis' from reading, conferences, rare personal experiences…" (Gaba, 1992, p. 493). In 2004, David Gaba published his view on the usefulness of simulation in teaching, training, and assessing behavioral performance of clinicians and teams. He purports that simulation is ideally suited to improve patient safety in anesthesia, as evidenced by its use in other HROs (Gaba, 2004). According to Gaba, the traditional focus of anesthesia education must shift from an individual's knowledge to that of performance assessment, particularly clinical team performance (Gaba, 2004).

Green, Tariq, and Green (2012) present a literature review that provides a summary of the history of simulation in anesthesia. This work highlights recent studies demonstrating the benefits of simulation use in anesthesia training. Advantages of simulation in this review focus on the realization of technical skills and retention of these abilities, as evidenced by recent studies: anesthesia skills for high risk, low occurring events, such as difficult airway management, obstetric emergencies, and cardiothoracic surgery; and for learning new uses of technology, for example, ultrasound for regional anesthetic techniques. The authors mention that outcomes of recent studies also attest to the benefits of simulation in the transfer of NTS to the

learner, specifically communication skills in interprofessional crisis situations. (Green et al., 2016).

Krage and Erwteman present support for the use of simulation in anesthesia training, particularly in the teaching of NTS (Krage & Erwteman, 2015). This publication provides strong evidence for the use of simulation in anesthesia training to positively affect patient outcomes (Krage & Erwteman, 2015). The authors highlight the "crucial role" human factors' play in patient safety and positive outcomes. They call for more simulation-based human factors related education and training in the "so-called" NTS required of competent anesthesia providers, particularly as these professionals depend on effective communication and clinical management: they often work in interdisciplinary teams during crises (Krage & Ereteman, 2015).

Sollid et al. (2016) developed an expert consensus statement on simulation's most important contributions to health care content areas that improve patient safety. Using a Delphi technique, the authors established an expert consensus identifying the top five areas most improved by simulation. Interestingly, this list distinguished subject areas including technical skills, system probing, assessment, and effectiveness – each one intertwined with a common subject, NTS (Sollid et al., 2016).

Assessment in simulation. Using objective structured clinical examinations (OSCEs), Sidi, Baslanti, Gravenstein, and Lampotang (2014), demonstrate that behavioral assessment requires the use of simulation coupled with an objective assessment tool. The authors use an evaluation model with two tiers, integrating lower level knowledge and skill evaluation with cognitive and behavioral assessment using simulation (Sidi et al., 2014). They place the qualities of assessment, evaluation, projection, and decision-making at the top of those skills necessary to reach clinical performance competence (Sidi et al., 2014).

Zausig et al. (2009) completed a randomized controlled study of anesthesiologists in simulation using medical management with and without NTS briefings. The authors' results did not demonstrate an improvement in scores between groups, one with NTS pre-training and the other without. Both groups did receive training in medical management. Expert consensus was used to identify markers of performance for NTS and medical management. A previously validated NTS rating scale, the Anesthesia Non-Technical Skills Scale or ANTS, was used to measure NTS, which is divided into four categories, including "resource management", "planning", "leadership", and "communication" (Zausig et al., 2009). These four elements are factors that can certainly impact SA, but are not the only contributors, according to Endsley's theory (Endsley, 1995b, 2015).

Simulation in training and education. Simulation training affords the benefit of performance assessment and gives the opportunity to teach health care trainees during dynamic clinical situations without actual patient harm (Sollid et al., 2016). Simulated clinical events can improve patient safety and outcomes by systematically exposing trainees, such as SRNAs, to established and effective patient management protocols in a controlled environment (Sollid et al., 2016). Simulated events can be stopped at key intervals to foster discussion and reflection, thereby enhancing learning (Sollid et al., 2016). Recent evidence demonstrates that utilizing simulation in health care improves practitioner performance, patient safety, and quality of care (Sollid et al., 2016; Zausig et al., 2009).

It is very difficult to demonstrate clear evidence that clinical decision-making is improved in real practice. This difficulty lies mostly in the essence of high-risk, low occurrence events: they occur infrequently, are not reported in a standardized manner, and have multi-factorial causation making the identification of any action as root cause near impossible. The research concept of clearly providing a link to improved patient outcomes is signified as "T4", with reference to the National Institute of Health (NIH) definition that includes the translation of research outcomes to the much broader measure of impacting population health (Vukotich, 2016). Excluding a few recent research studies, proving simulations' effect on patient outcomes is elusive. Of late, three studies were published showing such a link to patient outcomes: Wayne et al. (2008), Draycott et al. (2006), and Andreatta et al. (2011). All showed improvement in cardiac arrest survival rates in patients after providers received cardiac arrest specific simulation training (Andreatta, Saxton, Thompson, & Annich, 2011; Draycott et al., 2006; Wayne et al., 2008).

Simulation in education and training is well established in aviation, military, and even

judicial education (Riley, 2015). Since its inception almost 30 years ago as a teaching approach, simulation is currently incorporated in many health care education programs, including graduate medical education (Riley, 2015) and nurse anesthesia graduate programs (Council on Accreditation of Nurse Anesthesia Educational Programs, 2018). The Standards for Accreditation of Nurse Anesthesia Programs – Practice Doctorate (Council of Accreditation of Nurse Anesthesia Programs or COA) require that "simulated clinical experiences are incorporated into the curriculum" (Council on Accreditation of Nurse Anesthesia Educational Programs, 2018).

Situation Awareness

Stanton, Chambers, and Piggott (2001) offer a historical perspective of SA, describing it as a construct originating from World War II military operations, though the authors' contend it is a concept traceable to World War I military strategy. The first publications exploring SA emerged in the 1980s and describe the meaning of SA, factors that contribute to SA, the importance of maintaining SA, and the detrimental effects of losing SA (Stanton et al., 2001).

The 1990s brought a surge of attention and research into the construct of SA, attributable to the changes at that time in airplane system design: as automation of airplane control systems increased, the pilot was further removed from being in touch with what was occurring in the system – a loss of SA (Stanton et al., 2001). Other complex system industries, such as nuclear power control, navigation, and health care recently honed focus on the importance of SA for safety (Stanton et al., 2001). This leap from military and aviation underpinnings to other industries, particularly health care, Stanton et al. (2001) contend, traces back to the commonalities among these complex systems: the operator has multiple, simultaneous goals; the operator's attention is required for multiple tasks, all with some relevance to the goals; and the

operator is performing under time constraints and increased stress (Stanton et al., 2001). The authors also provide in this publication, evidence of the relevance of SA to dynamic, tightly coupled systems (Stanton et al., 2001).

Stanton, Salmon, Walker, Salas, and Hancock (2017) assert that the recent growth in research, publications, and application of SA in many industries creates the need for a "state-of the-science" review of different models of SA. Stanton et al. (2017) illustrate the contentious debate among different scholars and industries in defining SA, stating that this review is borne from necessity as the concept of SA has recently become the focus of cognitive psychology and health care. According to the authors, the definition of SA has evolved as the application of SA has grown over the past few decades: its meaning progressed from an individual focus to also include teams and systems (Stanton et al., 2017).

The authors demonstrate there is room for all SA models by linking the applicability of each to different problematic themes evident in recent literature (Stanton et al., 2017). This work provides an important comparison of the different models of SA, demonstrating their similarities and differences. The author's state that Endsley's model of SA is centered on the individual operating within a man-made system, originates from the domain of aviation, utilizes human information processing as the foundational theory, and is defined as the, "…perception of elements, comprehension of meaning and projection of future status" (Stanton et al., 2017, p. 454). They compare this SA model to others: one by Salas et al. (1995) that is centered on teams, meaning more than one operator; and the other by Stanton et al. (2006) that is focused on systems, defined by the authors as "… human and non-human agents" (Stanton et al., 2017).

Mica R. Endsley provides the most widely accepted definition of situation awareness (Endsley, 2015). In its simplest terms, SA is defined as knowing what is occurring around you.

Endsley's definition adds the elements of space and time, stating that it is an individual's, "…perception of the elements in the environment within a volume of space and time, the comprehension of their meaning, and the projection of their status in the near future." (Endsley, 1995b, p.36).

Due to the increasing exploration and competing models of SA, there is some controversy surrounding each of Endsley's assertions regarding SA. Endsley offers rebuttal to each, as she refers to them, "misconception and misunderstanding" of her interpretation of SA by competing authors (Endsley, 2015). Table 3 presents each purported "miscommunication and misunderstanding", with the attributed author(s) and work(s), and Endsley's rebuttal. In this paper, the importance of operator goals in a situation is underscored (Endsley, 2015). Situation awareness, she states, is a by-product of perceived operator goals, in that he or she directs his or her attention to that data/information believed to be relevant to the goal (Endsley, 2015). Therefore, SA is not linear in nature, necessitating attainment of a prior level to reach the next; rather it is more of a feedback type mechanism (Endsley, 2015).

Table 3

Controversies and Rebuttals of Elements of Endsley's Theory of SA

"Miscommunication and Misunderstanding"	Attributed Author(s) and Cited Works	Endsley's Rebuttal
3 Levels of SA are linear	Sorenson, Stanton, and Banks (2010); Salmon, Stanton, and Young (2012); Dekker and Lutzhoft (2004)	3 SA Levels are ascending: complex, dynamic systems utilize data driven and goal driven processing: data drives projection but projection and goals also drive data-seeking in an iterative manner; mental models and schema link SA Levels.

Data-driven Information-Processing Model	Salmon et al. (2012); Chiappe et al. (2011); Klein, Phillips, Rall, and Peluso (2007)	Model integrates some cognitive processes as traditional information- models but emphasizes goal-driven processing – linking goals to activation of mental models guiding interpretation and projection; environmental changes can change goals – a data-driven processing; mental models direct attention to get needed information; people play active role in getting and fostering own SA.
SA as Product versus SA as a Process	Salmon et al. (2008); Klein et al. (2007); Chiappe et al. (2011)	SA is process intertwined with resultant state of knowledge; processes involved in achieving and maintaining SA also affect product of SA which then affects processes.
Model is Not Cyclical or Dynamic	Salmon et al. (2008)	SA is dynamic feedback loop of information seeking and acting on environment; SA is constantly updated as environment changes; model stresses importance of time – SA based on events of the moment based on events of past and future.
Fails to Account for Meaning	Dekker and Lutzhoft (2004)	Expert analyses determines subject matter specific SA Levels and determines what is meaningful for successful operator; all levels meaningful; concept of meaning is SA's foundation.
All of SA Found in Working Memory	Chiappe, Rorie, Moran, and Vu (2012); Chiappe et al. (2011); Chiappe, Vu, and Strybeel (2012)	Working memory causes barrier for novices, as does any exposure to novel situation; long term memory – schemata and mental models – frees up perceptive abilities and increases SA attainment; experts rely more on long- term memory, working memory constrained in novices, Level 1 SA

Situation awareness in anesthesia. With advances in and the increased use of simulation, emerging research points to SA as a critical cognitive construct for patient safety and positive outcomes (O'Brien & O'Hare, 2007). Situation awareness' importance as a construct for the anesthesia provider originates with David Gaba and his 1995 publication with Howard and Small, "Situation Awareness in Anesthesia" (Gaba, et al., 1995). In this pivotal work on SA, Gaba et al. (1995) conclude that in dynamic situations, such as those occurring in anesthesia, the ability to perceive and read clues in an ever-changing environment is crucial to decision-making. The authors contend that this need for good SA, the key determinant of decision-making, provides the link between anesthesia and aviation, thus supporting the anesthesia specialty's need to imitate aviation safety practices, including simulation and crisis management (Gaba et al., 1995). The hallmark of high level SA, according to Endsley (2015), is the ability to constantly interpret key information and make accurate projections from a continuously changing situation.

Jones at al. (2018) present a systematic review of the human factors' implications of complications in anesthesia. The authors contend anesthetic related lapses in all three levels of SA are demonstrated in an "error taxonomy", ranging from unavailable data, failure to observe data, and misperception of data, to poor mental models, memory failures, inability to maintain multiple goals, and "habitual schema" (Jones et al., 2018). The authors address key clinical practice areas and circumstances in which human factors are integral to patient safety (Jones et al., 2018).

Measures of situation awareness. As important as SA is as a human factors' construct, so is the ability to assess the SA of operators who function in complex systems. Since the pivotal work of Endsley and others in the late 1980s through 1995, the importance of assessing SA has become the focus of researchers, leading to the development of SA measurement tools (Endsley, 1995a). Endsley describes various available SA measurement techniques along with their advantages and disadvantages (Endsley, 1995a). The author uses this review to advocate for the "Situation Awareness Global Assessment Technique" (SAGAT), which she developed as an instrument to objectively and directly measure SA (Endsley, 1995a). Endsley describes the bias inherent in subjective tools such as self-report and those that utilize observer rating (Endsley, 1995a). In this work, she presents two studies of the SAGAT, testing and demonstrating its validity and usefulness as a SA measure (Endsley, 1995a).

The importance of measuring SA finds its root in systems design, particularly in fighter pilot control display (Salmon, Stanton, Walker, & Green, 2006). Salmon et al. (2006) review SA measurement techniques in relation to SA assessment in command, control, communication, computers, and intelligence, also known as "C4i", environments. The authors describe various assessment techniques, including direct and indirect measures. These assessment techniques are presented in Tables 4 and 5, divided into direct and indirect measures, respectively.

Table 4

Direct Situation Awareness Measurement Tools

Measure	Instrument	How Performed	Advantages	Disadvantages
Subjective Measures	Self-rating Questionnaires: SARS, SART, posttest questionnaires	Post-scenario questions assessing scenario specific knowledge	Post-scenario, not obtrusive; easy to use, low cost	Subjects over-estimate SA in hindsight

Objective Measures	Freeze-Probe: SAGAT Real-Time Probe Verbal Protocols	Queries posed during freezes of simulated event Queries during unfolding events Subject "thinks" out loud as performing	Scant evidence performance impacted performance; anyone can administer (non-expert) No need to freeze task: useful for real-time (clinical) performance Performed real-time (clinical)	Pausing scenario may impact performance; needs creating context specific queries; needs simulation May clue subject to events; creating context specific queries; needs expert "actor" Dubious content validity; unable to measure SA by observation alone; subject may modify behavior as observed; needs expert observer

Note. Data for table compiled from Bolstad, Cuevas, and Cuevas (2010), McKenna et al. (2014) Endsley (1995b), Orique and Despins (2018), and Salmon et al.(2006).

Direct measures, as listed in Table 4, include: those that employ "freeze probe" methods used in simulated environments, whereby the situation is "frozen" or paused and the operator asked questions directly related to the unfolding scenario; and real time probe methods that are similar to the freeze probe but questions are asked in real-time without pausing a scenario, making this measure useful during a real event (Salmon et al., 2006).

Table 5

Indirect Situation Awareness Measurement Tools

	Instrument: Example	How Performed	Advantages	Disadvantages
Subjective measures	Behavioral marker systems: ANTS, NTS,	Expert observer rates performance using established markers	Can be utilized for actual events (clinical)	No relationship between behavioral markers

	NTS for surgeons, teams			and SA; time consuming; needs expert rater
Objective Measures	Performance outcome measures: Wombat-CS Process Indices: Eye- tracking, physiologic techniques (i.e. EEG)	SA inferred based on predefined outcome/ standard Eye tracking device worn during task performance measuring area of focus, time spent looking at area	Can be utilized for actual events (clinical) Can be combined with other physiologic data, i.e. heart rate, blood pressure, EEG to infer stress and cognitive perception	No relationship between performance outcomes and SA Cannot be used during actual clinical performance; lengthy data extraction Questionable link between eye fixation and data actually perceived; costly

Note. Data for table compiled from Bolstad and Cuevas (2010), McKenna et al. (2014) Endsley (1995a), Orique and Despins (2017), and Salmon et al. (2006).

Indirect instruments, as presented in Table 5, include: self-rating tools, which are self-reports of the operator's perception of his or her SA and performance; observer-ratings that employ expert observers to rate the operator's performance either in real time or from a recorded event; performance measures, which only measure the efficiency of performance; and process indices, such as eye tracking technology, that can track operator eye focus and use it to imply perception (Salmon et al., 2006). These instruments, indirect and direct, are those most often cited in health care literature (Bolstad & Cuevas, 2010; Cooper, Porter, & Peach, 2014; Endsley, 1995; Orique & Despins, 2017; P. Salmon, Stanton, Walker, & Green, 2006).

Situation Awareness Global Assessment Technique (SAGAT). Endsley's SAGAT is the only SA measurement tool that directly and objectively quantifies SA during simulated events (Bolstad & Cuevas, 2010). Other available SA measurement instruments are subjective, using self or expert rater observations of behavior (Lavoie et al., 2016). This tool requires no training or expertise to administer, further adding to its value and usability. Direct measure refers to an instrument that quantifies SA by evaluating information collected against what is actually occurring, not inferring it from behavior or performance (Orique & Despins, 2018). Indirect measures are those that draw an inference of SA based on certain behavioral markers or performance outcomes (Cooper et al., 2013).

Salmon et al. (2009) describe the reliability of an instrument as related to results being repeated either by a different test giver or at another time under different conditions. They define validity as the precision of an instrument at measuring what it sets out to measure (Salmon et al., 2009). This publication reports evidence presented by Jones and Kaber (2004) who offer numerous studies suggesting the validity and reliability of the SAGAT as a metric for SA (Salmon et al., 2009). Hogan, Pace, Hapgood, and Boone (2006) studied SA in subjects using the SAGAT, comparing participants' performance to a checklist of standards and management steps for similar scenarios. These authors reported a Cronbach's alpha of 0.77 and a Pearson's correlation of 0.81 with a sample of 16 subjects (Hogan et al., 2006).

Previously validated and applied to many domains, the SAGAT is known to be reliable, with test-retest reproducibility of results. Endsley gives examples demonstrating the sensitivity of this instrument to the differences in subject expertise, operational concepts, and system and automation operations across a wide range of industries (Endsley, 2000). More recently, testing of SA in the medical domain using the SAGAT shows its reliability and validity across many

specialties, including nursing, trauma, obstetrics, emergency physicians, and with medical students. LaVoie, Cossette, and Pepin (2016) use the SAGAT to evaluate nursing students' SA during a patient deterioration simulation. Their findings demonstrate that most of the queries developed (*n* = 21, 65.6%) were above threshold for proper discrimination using Ebel's criteria for D (discrimination) across three cohorts (n=109, 77, and 48) of testing (Lavoie et al., 2016).

Gardner, Kosemund, and Martinez (2017) demonstrated the validity and reliability of the SAGAT with testing of medical students during simulation. The authors employ the SAGAT with two trauma simulation scenarios in ten team-training sessions. Results demonstrate composite SAGAT scores predicted team performance in both scenarios, reported as R^2 of 0.30 for the first scenario and R^2=0.38 for the second (Gardner et al., 2017). Salmon et al. compare the results of the SAGAT to the Situation Awareness Rating Technique (SART), a SA self-rating tool. The SAGAT proved to be most accurate at assessing SA in this evaluation of the two tests (Salmon et al., 2009).

Endsley cautions that instrument sensitivity is dependent on using a broad range of queries during testing (Endsley, 2000). According to Endsley, to increase sensitivity, the SAGAT should contain a randomly selected assortment of queries that measure all Endsley's described SA Levels 1, 2, and 3 (Endsley, 2000). The SAGAT is a global measure of SA, Endsley asserts, and thus sensitivity of the test is enhanced with a comprehensive range of questions representing the three levels of SA (Endsley, 2000). Endsley suggests this is best accomplished using context specific, goal-directed task analysis, a form of cognitive task analysis.

Goal Directed Task Analysis (GDTA). Endsley suggests that procedures for goal directed task analysis (GDTA), in which context specific goals, sub-goals, key decisions, and SA requirements are used for item development, confer both face validity and content validity to the

instrument. Polit and Beck (2012) echo this assertion and emphasize that expert analysis provides face and content validity for questionnaires or "inventories" (Polit & Beck, 2012).

Polit, Beck, and Owen (2007) describe indices of item content validity (I-CVI) and item clarity (I-CI). The authors suggest using a panel of experts (*n* = 8-12) and a Delphi method for each item to achieve an optimal computed value of greater than 0.78. The entire scale content validity index can then be computed (S-CVI/Ave) with the optimal score being at least 0.90 with this large sample of experts (n = or > 8) (Polit et al., 2007).

Endsley's work maintains that using established objectives and guidelines for identification of key context specific goals strengthens face, content, criterion, and construct validity (Endsley, 2000). Crozier and co-authors used Miller's *Anesthesia*, a well-recognized textbook of anesthesia, for the objectives and goals of airway management from which they developed SAGAT queries (Crozier et al., 2015). Following Endsley's recommendations for SAGAT query development, the authors describe using these objectives and goals for GDTA (Crozier et al., 2015).

Endsley recommends that the output of GDTA be used as the foundation for item, or query, development. She does not provide limits as to how many queries need be developed, but stipulates that an exhaustive analysis of all that an operator would ideally "need to know" to successfully attain a goal be included (Endsley, 2000). The queries selected for the instrument need only be inclusive of the content related to the situation: no extraneous items should be included. Endlsey (2000) also does not recommend a total number of queries to include in a given assessment: however, the assortment of queries must represent all levels of SA to increase the sensitivity of the instrument (Endsley, 2000). During test administration, context specific items particular to the chosen scenario should be randomized, according to Endsley's

recommendations (Endsley, 2000).

Lavoie et al. (2016) developed a SAGAT instrument with a total of 31 items. Crozier et al. (2015) list a sampling of 8 scenario specific items. Salmon et al. (2009) developed a SAGAT with 24 items. Wright, Taekman, and Endsley (2004) identify a study by Zhang et al (2002) in which only four Level 1 queries, two Level 2 queries, and two Level 3 queries were used during SAGAT testing. The authors speculate that during this study with a small set of items, the subjects may have been able to predict the questions during the simulation freezes (Wright et al., 2004). Table 6 identifies common terms in instrument reliability and validity testing, providing key references that utilize the SAGAT to demonstrate these attributes.

Table 6

Reliability and Validity Terms, Definitions Correlated to SAGAT and Relevant Citations

Term	Definition	Demonstrated in SAGAT	Evidence (Citation) Recent Healthcare Literature
Face Validity	Instrument appears to measure what it is supposed to	Items derived by GDTA	Endsley, 2000; Gardner et al., 2017; Lavoie et al., 2016; Morgan et al., 2015; Salmon et al., 2009
Content Validity	Items that make up theinstrument adequately represent the variable being measured.	Items derived by GDTA	Endsley, 2000; Gardner et. al, 2017; Lavoie et al., 2016; Morgan et al., 2015; Salmon et al., 2009
Criterion Validity	Degree of correlation between scores on an instrument and an established standard <u>Predictive validity</u> Tool predicts observed measure in the future	GDTA based on established standards as agreed upon by content experts.	Crozier et al., 2014; Endsley, 2000; Gardner et al., 2017; Hogan et al., 2016; Lavoie et al., 2016

	Concurrent validity: correlation between scores on instrument with external criterion measured at same time		
Construct Validity	Instrument measures what it purports to measure	Established by the content experts	Crozier et al., 2014; Endsley, 2000; Gardner et al., 2017; Hogan et al., 2016
Sensitivity & Specificity	Sensitivity = correctly finds true positive cases Specificity= correctly identifies untrue cases		Endsley, 2000; Hogan et al., 2006; Lavoie et al., 2016
Test-Retest Reliability	Results repeated under separate occasions, and/or different test giver, subjects, scenario		Endsley, 2000; Gardner et al., 2017

Note: Definitions of terms from Polit and Beck (2012)

SAGAT use in anesthesia. The SAGAT utilizes queries related to unfolding events and is administered during a simulated operator situation (Gardner et al., 2017). The assessment tool involves questioning subjects at predetermined freeze times during a developing event, hence relegating its application to simulation, the main drawback of this instrument (Gardner et al., 2017). The subject is asked questions directly related to developing events, which are purported to measure perception of pertinent data as well as higher levels of SA (Lavoie et al., 2016). The requirement of developing goal specific questions that reflect appropriate levels of SA in a specific context represents another limitation of this instrument (Lavoie et al., 2016).

Wright et al. (2004) attest to the applicability and appropriateness of the SAGAT as a SA measurement tool for anesthesia providers (Wright et al., 2004). The authors present the

anesthesia specialty as event-driven and dynamic, making the SAGAT well suited for this domain: SA is the footing of cognitive and behavioral abilities in this provider population (Wright et al., 2004). The authors describe using GDTA to create this anesthesia specific SAGAT (Wright et al., 2004). An exhaustive review of the literature reveals that, to date, there is no completed SAGAT instrument developed to assess SA in anesthesia trainees (SRNAs or physician residents) or anesthesia providers.

Theoretical Framework

World War I military theory demonstrates a recognition of the importance of high level situation awareness related to aircraft and engaging the enemy, identifying a division between an operator's understanding of a system's status and the actual status of that system (Manz et al., 2013). References to SA are evident in military literature as early as the 1900s, yet in the technical and academic realms the focus does not escalate until the late 1980s (Manz et al., 2013). In 1995, SA is heralded in human factors' research as the backbone of safety and effective situation management (Stanton et al., 2001).

The recent increasing interest in SA in the many domains outside of the military and aviation evolves from the rapidly expanding use of advanced technology in these complex, dynamic systems, including nuclear power, automotive, and most recently health care (Stanton et al., 2017). Technological advances may increase an operator's SA, making more timely information available, or detract from SA because of poor design or furthering the individual from understanding the operations of the system (Endsley & Kiris, 1995).

Mica Endlsey presents the theory of situation awareness in her 1995 work, "Toward a Theory of Situation Awareness in Dynamic Systems" (Endsley, 1995b). Endsley developed this "Theory of Situation Awareness" from years of military and aviation research, including

research of SA and the effects of automation on SA in the aviation industry (Endsley & Garland, 2000). Since this publication, Endsley's work is the focus of research and utilized in other complex system industries, such as automobile, nuclear power, and medicine.

Situation awareness and decision-making. Situation awareness, Endsley's theory posits, is the critical precursor of decision-making, which in turn is vital for action performance (Endsley, 2000). Situation awareness, decision-making, and performance are a feedback loop, each being re-evaluated and adjusted as the loop cycles (Endsley, 2000). This all occurs within a dynamic environment, meaning within rapidly changing circumstances and within a given time frame, precipitating reassessment of the elements of the feedback loop as shown in Figure 4 (Endsley, 2000).

Levels of situation awareness. Endsley (2000) describes the construct of SA as having three hierarchical levels that impact appropriate decision-making and performance in dynamic environments. In this publication, the author refers to Level 1 SA as the intake of information within the environment, or as Endlsey defines it: perception of elements in the current state (Endsley, 1995b). As SA

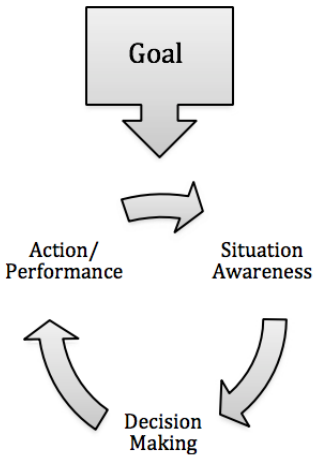


Figure 4. Endsley's "Theory of Situation Awareness Goal-Dependent Feedback Mechanism" as

adapted from Endsley (1995b).

progresses, it reaches Level 2, defined as the understanding of what this information means within a given context (Endsley, 1995b). The author places the highest level of SA, or Level 3, at the ability to forecast future events based on the operator's comprehension of the information at hand and the evolving circumstances (Endsley, 1995b). As Endsley reports, the levels of SA are non-linear and "ascending", as this all occurs within a complex system that is in a state of flux (Endsley, 2015).

Complex, dynamic systems necessitate goal-driven processing for proficient operation (Endsley, 2015). This SA model is an iterative process whereby operatives mostly utilize goal-driven processing, but also use a linear data-driven input of understanding to further their comprehension of the unfolding events (Endsley, 2015). Purely data-driven linear processing is termed bottom-up processing, and conversely, the goal-driven processing is labeled top-down processing (Stanton et al., 2017). Endsley's model of SA utilizes both bottom-up and top-down processing, with the higher levels of SA using top-down, goal driven processing. Endsley's feedback loop demonstrates that goal-driven processing initiates the push to find new data, and the new data fuels adjustments to comprehension and goals (Endsley, 2015).

This dynamic environment of SA, decision-making, and action performance is influenced by external or "task/system factors" and internal or "individual factors" (Endsley, 2000). These factors influence the way in which operators choose, understand, and translate information (Endsley, 2000). The external factors include: "system capability", "interface design", "stress and workload", "complexity", and "automation" (Endsley, 1995b). Each of these external factors has the potential to impact SA, decision-making, and performance (Endlsey, 2000). Goals and presumptions are part of the internal factors and can influence SA and decision-making (Endsley,

1995b). How an individual processes information, his or her long term memory, and automaticity are internal/individual factors that can not only influence goals and presumptions, but also bear weight on SA, decision-making, and performance (Endsley, 2000). Skill level, training, and experience will all have influence on information processing, long-term memory, and automaticity (Endsley, 1995b).

As this model demonstrates, SA is not interchangeable with decision-making. Situation awareness, according to Endsley's theory, is instead the vital precursor to decision-making (Endsley, 2000). Endsley states that having good, or high-order SA does not necessarily equate to good decision-making. As the operator must make a conscious choice in formulating a decision, he or she may have excellent SA yet still make an incorrect decision (Endsley, 2000). This incorrect decision-making despite high level SA may be due to inadequate resources, organizational constraints, or personality factors such as impulsiveness, indecisiveness, or risk-taking behavior (Endsley, 2000). As the theory depicts in its iterative, feedback loop process, SA drives decision- making and decision-making impacts SA (Endsley, 2000). There is a strong relationship, nonetheless, between sound SA and good decision-making, ultimately leading to appropriate actions (Endsley & Bolstad, 1994). This is most often seen in operators with good training and experience (Endsley & Bolstsad, 1994).

According to Endsley (1995b), SA is derived from the input of all an operator's senses. Operators, she stipulates, are active participants in information acquisition (Endsley & Bolstad, 1994). The operator, based on his or her goals, chooses the information on which to focus, thus controlling SA Level 1: data acquisition (Endsley & Bolstad, 1994). The operator may also control, depending on the system design, what information is available (adjusting screen layouts, etc.), also bearing impact on his or her SA (Endsley & Bolstad, 1994). This may negatively or

positively impact SA depending on whether or not the operator's preconceived goals are accurate or if the information feeding the goal is selected appropriately (Endsley, 2000). With proper training and experience, these factors should lead to appropriate goal selection, dependent on the scenario (Endsley, 2000). Endsley points out that SA can be enhanced and improved upon with proper training and experience (Endsley & Bolstad, 1994; Endsley, 2000).

Conclusion

This chapter, drawing on the one preceding, links medical error and anesthesia error to SA and presents the relationship between the integration of simulation in health care education, with an emphasis on SA, to improved patient safety and outcomes. The importance and need for a valid, reliable, direct, and objective SA measurement tool is highlighted. The SAGAT is introduced, along with evidence of its reliability, sensitivity, and validity. The increasing use of this instrument in medical simulation is also presented. Endsley's "Theory of Situation Awareness" is thoroughly discussed at the conclusion of this chapter.

Moving forward, the next chapter presents the methodology of adaption and validation of the SAGAT for nurse anesthesia graduate student assessment, the goal of this research study. An in-depth description of study design, sampling, data collection and analysis, limitations and assumptions are discussed. Methodology is presented according to the accepted criteria of qualitative research reporting (Kisely & Kendall, 2011; Tong et al., 2007). From this presentation, chapters follow presenting this study's results in explicit detail, an ensuing discussion of its relevance and implications, and as a foundation for future research.

Chapter Three: Methods

The primary goal of nurse anesthesia education and training is to develop anesthesia providers (certified registered nurse anesthetists, or CRNAs) that can effectively and proficiently recognize, understand, and respond to the dynamic physiologic and environmental conditions of the anesthetized patient and the environment (Wright & Fallacaro, 2011). Patient safety in terms of the anesthesia specialty requires a provider awareness of the entire milieu within which he or she is operating, keeping a patient's well-being the focus: this is the essence of high level SA (Schulz et al., 2013; Wright et al., 2004; Wright & Fallacaro, 2011). To minimize anesthetic morbidity and mortality, high level SA is required in anesthesia providers to optimize critical thinking, and is a key characteristic of a proficient, safe anesthesia provider (Schulz et al., 2013; Wright et al., 2004).

Endsley developed the SAGAT, and as an exhaustive literature search affirms, it is the only tool validated to directly and objectively measure SA (Endsley, 1995b; Schulz et al., 2013; Wright et al., 2004). As good SA is built on training and experience, the importance of a quantitative, objective SA measurement tool is clear. An in-depth review of the literature, as described in the previous chapter, demonstrates no SAGAT, or any other direct and objective SA measurement tool adapted and/or validated to assess SA in anesthesia trainees (SRNAs or physician residents), or any anesthesia provider.

The primary goal of this study is to adapt and validate the SAGAT as a means to quantify the level of SA in the SRNA subset of anesthesia trainees, during the simulation of the induction of general anesthesia with the associated placement of an oral endotracheal tube. The final product emerging from this study will consist of a list of context specific items related to a particular experimental simulation setting and relevant to those SA requirements necessary to

successfully attain the established context specific goals and sub-goals. After the conclusion of this study, this final list of validated items becomes the pool of queries from which questions for SAGAT testing are drawn.

In the previous chapter, a review of the literature demonstrates a gap in direct and objective SA measurement tools for the anesthesia domain, particularly for nurse anesthesia graduate students. The necessity of measuring SA in this population is supported by the thorough review of the literature provided in the previous chapter. The review gives evidence of anesthesia as a complex system, along with the inherent risk of preventable errors in this system. This is directly linked to the recent emergence of simulation as a key component within health care education, particularly for anesthesia trainees. Growth in research related to the effect of human factors, specifically related to SA and its implications in appropriate decision-making, is offered as the foundation for the need of its assessment. The comparison of SA theoretical models provided demonstrates Endsley's "Theory of Situation Awareness (1995) as most appropriate to the anesthesia domain. Additionally, the review of currently available SA measurement tools affirms Endsley's SAGAT as the only direct and objective means by which SA can be quantified.

Rationale and Assumptions for Qualitative Design

With the application of Endsley's "Theory of Situation Awareness" (1995) as the foundational support for this study's main goal as highlighted in the identified research question, it is appropriate that a qualitative approach was employed to develop query items for the SRNA population and the specific event: induction of general anesthesia with associated oral endotracheal tube placement. Specifically the research question stated, Can Endsley's SAGAT be adapted and validated to quantify the SA of nurse anesthesia graduate students, or SRNAs,

during the simulation of a specific anesthesia event?

According to Endsley's recommendations for creation of SAGAT query items, GDTA was performed. Goal directed task analysis is a form of cognitive task analysis: because SA is based on goals which are the foundation of decision-making, the key decisions necessary to attain context specific goals are analyzed, not the requisite tasks. Endsley states that once each goal and sub-goal is identified, the SA required to make the decisions must be delineated (Endsley & Garland, 2000). Determining these key decisions requires the identification of the information an operator needs to successfully manage these goals, as well as how this information is integrated by the subject in approaching each decision (Endsley & Garland, 2000). Endsley recommends accomplishing GDTA by expert interview, observation, and review of the literature (Endsley & Garland, 2000). This methodology is most appropriately achieved by qualitative research approaches.

Qualitative research methods systematically collect evidence, just as is accomplished by quantitative methods. Qualitative research diverges from quantitative designs in that it examines a research problem from the perspective of the population involved (Kisely & Kendall, 2011). Qualitative methods induce an in-depth understanding of specific phenomena most often by employing analysis of observations, interviews, and written text (Kisely & Kendall, 2011). Eliciting the information accurate GDTA requires is best accomplished, therefore, by qualitative research techniques.

The resultant list of SAGAT queries from this study emerged from the key decisions and necessary SA requirements identified by GDTA. As Endsley outlines for ideal item development, the queries were, "…phrased as similar as possible to how the person thinks…" to aid in the ease of administration of and response to items (Endsley & Garland, 2000, p.1). The

basis of developing and writing these items came from content analysis methodology, a qualitative research paradigm. Validation of this final set of items, according to Delphi methods, stems from the iterative process of expert consensus on each item's relevance to specified subject matter and correlated SA.

This study provided further validation of the generated item list by qualitative and quantitative methods. A larger, geographically diverse sample of experts assessed the items relevancy to the specific scenario. Then, quantitative techniques applied to these results further validated the list of items. Content validity indices (I-CVI) for each item as well as the entire list of items (S-CVI/Ave) bore further proof of validity, as did an exploratory factor analysis (EFA) performed on this same inventory. For this study, the qualitative methods predominated and were used in the majority of study design. Qualitative methods bore all data that generated the final list of items. The quantitative methods provided an additional layer of validation for the items and came at the end of the design.

Type of design. This is a mixed methods instrument adaptation study utilizing a sequential exploratory mixed methods design: a majority of qualitative methods are used until the final aspects of the study, in which quantitative methods are employed for further validation of results (Fetters, Curry, & Creswell, 2013). Figure 5 depicts this study's design. The predominant qualitative methods are the basis for data collection, content analysis, and item development. The resultant items generated from this study are further validated by quantitative techniques, with the goal of their eventual use to quantify SA in SRNA's during the simulation of the standard induction of general anesthesia with associated placement of an oral endotracheal tube. Adaptation and validation of this instrument occurred over three distinct phases and in accordance with Endsley's recommendations: the first phase used qualitative approaches; the second phase used both qualitative and quantitative methods; and the third phase quantitative techniques.

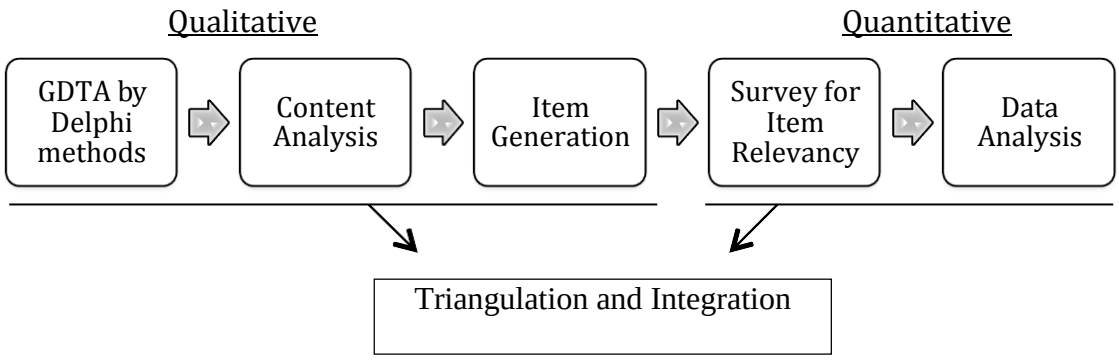


Figure 5. Depiction of sequential exploratory methods design used in this study. Depiction adapted from Creswell et al. (2003), Advanced Research Mixed Methods Designs (Creswell, Clark, Gutmann, & Hanson, 2003).

Content analysis was the foundational framework used for the qualitative portions of this study design. All resultant data from the first two distinct phases of the design originated from

the systematic organization of participant replies into a methodical format (Tong et al., 2007). This theoretical framework supported the exploration of the research question and goals of this study (Tong et al., 2007). Goal directed task analysis, Endsley's recommended means by which SAGAT items should be developed, is best completed using a combination of, "…expert elicitation, observation of operator performance of tasks, verbal protocols, analysis of written materials and documentation, and formal questionnaires…" (Endsley & Garland, 2000, p.2). This is the essence of content analysis, a systematic framework for evaluating text and communications (Kisely & Kendall, 2011).

Researcher's Role and Reflexivity

As this study bases its results on the expert opinions of nurse anesthesia educators, they themselves being nurse anesthetists (CRNAs) with both clinical and education/academic experience, the researcher must divulge that she is also a CRNA and educator in a nurse anesthesia program. Other members of this study team, including the dissertation committee chair and two other committee members, are also CRNAs and nurse anesthesia educators. Though it is through these characteristics of the researcher that led to the choice of research study and area of focus, the data collection performed by the research assistant helped minimize any bias that may unintentionally be placed on results. Providing the participants in this study serial iterations of the same survey, including their verbatim answers from previous rounds, also minimized any bias the researcher may impart in data content analysis. Utilization of a research assistant to manage all participant electronic communication and data collection, who is neither a CRNA nor a nurse anesthesia educator, ensured anonymity of all participants and facilitated the distancing of the researchers from the participants.

The smaller sample utilized in the first phase of this study design, the GDTA, was an

expert panel of CRNA, nurse anesthesia educators closely associated with the researcher. Though electronic survey utilization increased anonymity, the researcher had frequent interactions with this sample group, unrelated to this specific study. The second phase of this study design also employed a CRNA, nurse anesthesia educator sample, however this sample was a larger, geographically diverse pool of experts. The researcher did not engage in direct communication with this sample, and anonymity was maintained by electronic survey and data collection techniques. Information regarding the researcher's credentials and interest in this research subject were disclosed to all participants before their consent to participate was formally attained.

Participant Selection

Utilization of purposive sampling for the qualitative components of this study garnered study participants able to provide expert opinion regarding the research questions' topical focus. A purposive sample is a nonprobability sample: subjects are selected based on their expertise or knowledge (expert sample) (Kisely & Kendall, 2011). Purposive sampling is typically employed in qualitative research to identify experts in a domain of interest who are available, will freely participate, and are well versed and reflective of the content area and their opinions (Palinkas et al., 2015). Also, as this study employed Delphi methods to elicit expert response, purposive sampling is preferred. The Delphi technique does not employ random sampling to represent a targeted population; rather it uses a panel of experts with each participant having key knowledge in the area of research interest (Keeney, Hasson, & McKenna, 2001).

Participants in all phases of this study design included only CRNAs who are currently nurse anesthesia educators. CRNAs without a minimum rank of assistant professor in an accredited nurse anesthesia program on university web-listings were excluded from participating

in this study. This benchmark of academic rank lends the participant a level of experience within academia. Though this may actually not be in fact a tenure track position, this study design assumed that the assistant professor title represents a greater academic experience level than those with lower rank academic titles. A participation incentive was offered to the second phase recruitment sample: those in the first phase sample did not receive this offer as their university employment policy prohibits faculty from accepting any incentives.

As the GDTA is based on a customary and well-established anesthesia practice of induction of general anesthesia with associated oral endotracheal tube placement, nurse anesthesia educators are an ideal sample for recruitment: they represent anesthesia professionals' charged with teaching these accepted, standard practices to SRNAs. "Studies employing the Delphi make use of individuals with knowledge of the topic being investigated" (Hasson, Keeney, & McKenna, 2000). Identification of the goals, sub-goals, key decisions, and SA requisite during the induction of general anesthesia also requires no experience in simulation. For these reasons, and in recognition of the criteria for trustworthiness in qualitative research, using nurse anesthesia educators affords credibility to the items generated in the study. The criteria of credibility in qualitative research highlights that content validity of instrument items can be assured by use of those with item content specific knowledge in their creation (Shenton, 2004). Table 9 delineates the inclusion and exclusion criteria of all those asked to participate in this study.

Phase one participants who completed the GDTA originated from a purposive sample of nurse anesthesia educators at a large, urban university. Participants included were CRNAs currently teaching in this nurse anesthesia program, and contact for recruitment was made by email directly from the research assistant. Those recruited have a minimum title of assistant

professor listed in university resources. All seven solicited agreed and consented to participate in this phase of the study, and completed the Delphi process of GDTA.

The second phase of this study design again employed purposive sampling of CRNAs who are nurse anesthesia educators. The Council on Accreditation of Nurse Anesthesia Educational Programs (COA) provides, at no cost, a public list of directors for all 121 accredited nurse anesthesia programs across the United States. As these programs are geographically diverse, spanning all regions of the United States, purposive sampling in this phase of the design matched that of a "maximum variation strategy" (Palinkas et al., 2015). Purposeful sampling with a "maximum variation strategy" attempts to elicit important homogenous patterns in the sample despite the participants themselves being heterogenetic based on regional influence implications: this method imparts significance to the subjects' shared opinions (Palinkas et al., 2015).

Some of the COA listed program director's email addresses provided a generic, university-based email address. To assure delivery of the solicitation to the intended recipient, a web-based search was performed for specific, individual email addresses utilizing public, university-based web-resources. This method revealed other potential participants meeting inclusion criteria: individual's holding CRNA certification and having a listed minimum designation of assistant professor at an accredited nurse anesthesia program. These additional individuals were included in the recruitment sample. No individual email for some individuals listed in the COA directory could be found using public web-based resources, and thus those individuals were not included in the sample. A final recruitment sample of 165 potential subjects was compiled.

As in the previous phase, a research assistant contacted all subjects electronically (email)

for recruitment. After three attempts at recruitment, 49 individuals consented to participate in this phase of the study design. Those who did not participate did not provide rationale: no response was received to any of the three attempted recruitment solicitations. Of these 49 consented subjects, nine completed only the demographic survey. This resulted in 40 subjects completing both the demographic and item relevancy surveys. Figure 6 demonstrates the sample recruitment process in this phase of the study.

Setting. Both phases of this design, the GDTA by those at the large, urban, university nurse anesthesia program and the subsequent survey of items for relevancy rankings to the larger and

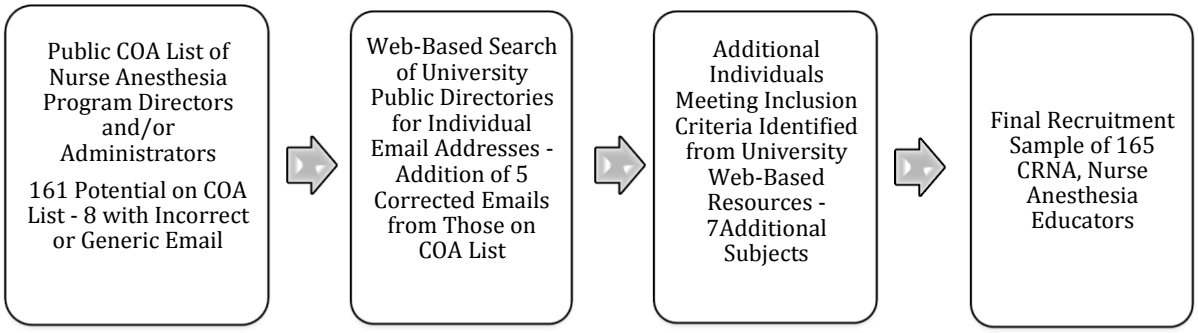


Figure 6. Depiction of methods utilized for phase three recruitment sample identification.

geographically diverse sample, took place entirely by electronic means. Three iterations of surveys for the GDTA were distributed to all phase one participants via email utilizing REDCap electronic data capture tools. Consent and data collection for all three iterations of this GDTA instrument also were collected electronically via REDCap. The study design's third phase, the survey of subjects for relevancy of created items from phase one, also was disseminated electronically utilizing REDCap electronic data capturing tools.

Study data were collected and managed using REDCap electronic data capture tools hosted at Cizik School of Nursing at the University of Texas Health Science Center at Houston, Houston, Texas. REDCap (Research Electronic Data Capture) is a secure, web-based application designed to support data capture for research studies, providing: an intuitive interface for validated data entry; audit trails for tracking data manipulation and export procedures; automated export procedures for seamless data downloads to common statistical packages; and procedures for importing data from external sources (REDCap.(n.d.); Harris et al., 2009).

This study design approach included classical Delphi methods. Classical Delphi techniques utilize a first round open-ended survey questionnaire distributed via email or postal services to identified and recruited topical experts. Content analysis performed on the first round is synthesized into another round of questions to this same panel, soliciting their confirmation or rebuttal of concepts. In this specific study, a third round was employed to ensure the incorporation of all respondent's feedback.

As this was performed using REDCap electronic data collection services, it is considered "e-Delphi" research (Toronto, 2017). "e-Delphi research", by using electronic data collection tools, provides very limited contact with the researcher: in this study, there was no direct contact between any participants and the researcher (Toronto, 2017). This also afforded the participants the ability to respond freely, without concern of other's opinions or loss of anonymity within the group, and removed any influence other respondents may have on a participant's response (Toronto, 2017). No other person involved with the research study was present when the research participants responded to surveys, another benefit of electronic data collection tools. Surveys can be answered in any location with internet access (Toronto, 2017).

Description of sample. Participating subjects in the first phase of the study design included seven CRNAs who are faculty of the nurse anesthesia program at a large, urban university in the southwestern region of the United States. According to Polit and Beck (2012), a sample of seven to 10 experts provides excellent content validity (Polit & Beck, 2012). All seven subjects met inclusion criteria, including that of having CRNA certification and being a nurse anesthesia educator. Table 7 lists the demographic information of these seven, first phase study participants.

Participating subjects in the study's second phase include 49 CRNA, nurse anesthesia educators from across the United States. All subjects hold faculty positions at a COA accredited nurse anesthesia program. Table 8 provides a summary of the demographic information of the

Table 7

Characteristics of Expert Panel Phase One N = 7

Characteristic		Response
Gender		
	Male	6
	Female	1
Age		
	Mean	38.8
	Range	33 to 43
Years as CRNA		
	5 to 10	4
	11 to 15	2
	16 to 20	1
Years as NA Educator		
	1 to 4	4
	5 to 10	1
	11 to 15	1
	16 to 20	1
Employment status		
	Full-time NA faculty	4
	Part-time NA faculty	1
	Adjunct/contract faculty	1
Job title		
	NA Program Director	1

	NA Program Assistant Director	5
	NA Educator/Faculty	5
Highest Academic Degree Completed		
	DNP/DNAP	7
Currently Practicing Clinically		
	Yes	7
HPS Incorporated into NA Program		
	Yes	7
Type of Sim Used in NA Program		
	TT	4
	SP	3
	HP mannequin	6
Participate in Sim Component of NA Program		
	Yes	5
	No	1
	No response	1
Training in Sim Education		
	Yes	2
	No	5

Note. CRNA = Certified Registered Nurse Anesthetist; NA = nurse anesthesia; TT = task trainer; SP = standardized patient; HPS = human patient simulation; Sim = simulation; FT = full-time; PT = part-time; phase two participants.

Table 8

Characteristics of Expert Panel Phase Two N = 49

Characteristic		Response
Gender		
	Male	16
	Female	32
	No response	1
Age		
	Mean	51.43
	Range	35 to 70
Years as CRNA		
	1 to 4	1

	5 to 10	4
	11 to 15	13
	16 to 20	9
	21 to 25	6
	26 to 30	4
	31+	2
Years as NA educator		
	1 to 4	8
	5 to 10	14
	11 to 15	10
	16 to 20	5
	21 to 25	6
	26 to 30	4
	31+	2
Employment status		
	Full-time NA faculty	43
	Part-time NA faculty	4
	Adjunct/contract NA faculty	1
	No response	1
Job title		
	NA Program Director	27
	NA Program Assistant Director	14
	NA Educator/Faculty	8
Highest Academic Degree Completed		
	DNP/DNAP	31
	PhD	17
	EdD	1
Currently Practicing Clinically		
	Yes	46
	No	3
If no, years since last practiced anesthesia		
	1 to 3	2
	4 to 6	1
Sim incorporated into NA program		
	Yes	44
	No	5
Type of sim used in NA program		
	TT	42
	SP	30
	HPS mannequin	42
	VR	8

Participate in sim component of NA		
	Yes	42
	No	5
	No response	5
Training in sim education		
	Yes	33
	No	16

Note. CRNA = Certified Registered Nurse Anesthetist; NA = nurse anesthesia; TT = task trainer; SP = standardized patient; HPS = human patient simulation; VR = virtual reality; Sim = simulation; FT = full-time; PT = part-time;

Data collection. The study received Investigational Review Board (IRB) approval with "exempt from review" status from both the University of Texas Health Science Center at Houston in Houston, Texas, and Virginia Commonwealth University in Richmond, Virginia. Research studies qualify for exempt status if they are deemed to pose minimal or no risk to subjects.

After receipt of IRB approval, the first phase of the study design rolled out to the seven identified subjects. At the completion of the phase one Delphi third round, generation of SAGAT items ensued by content analysis. This content analysis and creation of the list of items represents phase two of this study design. The created items were then sent as a survey for relevancy to the phase two participants. A detailed simulation scenario accompanied this survey to provide context for item relevancy ratings.

***Data sources: Collection and analysis*.** At the commencement of phase one, the seven identified subjects received email solicitation describing the study, the goal of this research, and the estimated time required for participation. A brief explanation of SA and GDTA was provided. The subjects also received a brief explanation of the Delphi technique, instructions regarding responses, and importance of participation in the three rounds. Consent to participate was implied upon entering the survey via the provided web link, as detailed in the consent form.

This is standard practice in low-risk surveys (Wiener, Chacko, Brown, Cron, & Cohen, 2009). Surveys created in REDCap, an electronic data collection tool, are the data source for the entirety of this project.

Phase one of this study consisted of three rounds of surveys. The goal of this entire phase one was development of the GDTA for the induction of general anesthesia with associated placement of an oral endotracheal tube. The induction of general anesthesia represents a critical time period of flux for a patient and anesthesia provider. During this event, medications are administered to induce a state of unconsciousness. It is a time fraught with changes in a patient's physiologic state coupled with high workload and task management requirements: the anesthesia provider must constantly prioritize actions. The induction of anesthesia, therefore, is a key event for which SA is critical. It also incorporates much of the key didactic content of nurse anesthesia programs: anatomy, physiology, pathophysiology, pharmacology, anesthesia equipment, problem solving, and communication. This event, therefore, is ideal to measure SA in the SRNA population. Also, this scenario provides easily identifiable and important goals according to established objectives: an event conducive to the delineation of key actions, decisions, and SA requirements.

To meet Endsley's recommendations for GDTA as well as to fulfill the four criteria of truthfulness of qualitative research design, this first phase of the study employed Delphi methods. The criteria of credibility in qualitative research highlight that content validity of instrument items can be assured by using those with item content specific knowledge in their creation (Shenton, 2004). Morgan et al. (2015) successfully utilized a Delphi technique in query development for the adaptation of the SAGAT in medical education (Morgan et al., 2015).

The Delphi technique is a widely used and accepted method for achieving consensus

among experts in health and social sciences (Hasson et al., 2000). This method uses serial iterations of questionnaires and evidence indicates two or more rounds are necessary to achieve consensus, with three rounds being ideal (Keeney, Hasson, & McKenna, 2001). For phase one of this study, the experts received three rounds of surveys, as the literature indicates the largest adjustments to items using Delphi processes occur in the first two rounds of surveys (de Villiers, de Villiers, & Kent, 2005).

The participants received email instructions including Endsley's definition of SA and her characterization of the three levels of SA. Instructions as to the determination of sub-goals, key decisions, and respective necessary SA requirements were also provided. The first round of surveys included open-ended questions regarding identification of the sub-goal, key decisions to reach this sub-goal, and the SA Level 1, 2, and 3 these decisions require. The questionnaire allowed subjects to enter as many sub-goals and corresponding decisions and SA requirements they deemed necessary. Data collection employed REDCap electronic data collection tools, the output of which was directly derived from subjects' exact input. Other studies adapting the SAGAT performed GDTA using established and published guidelines and protocols. This study enabled experts to deduce all GDTA elements from their own expertise.

The next round of the questionnaire was returned to the expert panel and included all subjects' anonymous responses verbatim. This round gave the group the opportunity to see everyone's responses, agree or disagree with a response, and provide additional responses if they wished. Data again was collected with REDCap technology, and the subsequent content analysis performed used manual extraction methods. The researcher hand tabulated text responses. The subjects were given the number of total experts in agreement or disagreement with each response. Two subjects did not participate in this round but were not eliminated from the next

round. Literature is lacking regarding the definition of consensus, with 51% presented as an acceptable benchmark. As these are readily identifiable sub-goals, decisions, and SA requirements based on a routine scenario, 70% consensus provided a stringent cutoff for agreement in this phase of study design.

The third and final round gave the expert panel all responses that met a minimum 70% consensus. They were also afforded the opportunity to once again agree or disagree with the groups' responses. Free text capability allowed respondents to add responses if they determined them requisite. Content analysis performed on the results of this round revealed the content for the key queries to include for a simulation scenario of the induction of general anesthesia with the associated placement of an oral endotracheal tube. The derivation of these items represents the first part of phase two of this study design.

Item development utilized content analysis methodologies. Manual extraction of the expert data collected elicited a list of 71 questions related to the three levels of SA. Once compiled, a review determined a redundancy of themes in some items and these were consolidated. A final list of 43 items was generated. A simulation scenario very specifically detailed to the unfolding events of a routine induction of general anesthesia with associated oral endotracheal intubation was created, specific to a healthy patient with no comorbidities and a clear-cut simulation event.

The second half of phase two of this study drew upon the expert opinion of the larger, geographically diverse sample of CRNA, nurse anesthesia educators. The expert panel for this part derived from those 165 CRNA nurse anesthesia educators solicited by recruitment email. The recruitment email identified the study, it's funding provided by the AANA foundation, the importance of SA and its measurement in the SRNA population, and a brief description of their

role as expert, including the requirements for participation. Estimated time to complete the survey was provided, as well as a demographic survey identical to that given to the phase one sample group. Unlike the first group, these potential subjects were offered a participation incentive of an undisclosed amount, according to IRB protocol.

The items sent to this sample were in no obvious order, and came with explicit instructions to rank them according to their relevancy to the provided scenario. Each item was listed with a four point Likert scale (0 to 3) from not relevant to somewhat relevant, relevant, and then highly relevant respectively. Of the original 49 consented subjects for this phase of the study, only 40 participated in the survey to rank items, and of those 40 only 34 completed the survey in its entirety (no missing responses).

Data analysis of these responses constitutes the third phase of this study.According to Polit, Beck, and Tatano (2012), for items to be considered excellent in terms of content validity, a consensus among 7 to 10 experts is required. As 40 subjects participated in the ranking survey, content validity is assured. After validity indices were computed, an EFA was performed as additional validation of results.

Though utilizing experts for item content imputes validity, further validation of items was provided by quantitative techniques. Each query was tested using content validity indices (I-CVI) and the entire scale (or group of items) was tested by scale content validity measures, including the scale content validity index average (S- CVI/Ave) (Polit & Beck, 2006). Item content validity scoring relates to the proportion of experts who deem an item relevant or highly relevant. The entire set of queries needs testing across all items for scale content validity (S-CVI) (Polit & Beck, 2006). It is recommended that for a scale to be considered as excellent in regards to content validity, it requires items with I-CVIs of 1.00 when using 3 to 5 experts, and I-CVI of

0.78 for 6 to 10 experts (Polit & Beck, 2006). Excellent content validity also necessitates the scale content validity (S-CVI) in relation to the average of the I-CVIs for all items on the scale (S-CVI/Ave) to be 0.90 or higher (Polit & Beck, 2006).

Additionally, to add rigor to the validity, EFA was performed to further establish reliability of results. Specifically, an EFA with principal axis factoring used as the extraction method determined the structure of the questionnaire (DeVellis, 2016). Factor analysis seeks to determine if any latent variables, or factors, can be identified from the items on an instrument by examining how specific items tend to correlate with each other. The content of the items that form the factors can be used to define the factors, or in this study, the level of situation awareness.

Methods for Verification of Trustworthiness

To provide robust verification of the trustworthiness of this study, the researchmethods described in this chapter are reported according to established criteria for reporting qualitative research (Kisely & Kendall, 2011; Tong et al., 2007). The quantitative methods performed to further validate results are provided in add.ition The "Consolidated Criteria for Reporting Qualitative Research", or COREQ, is a systematic checklist reporting framework utilized to assess and report qualitative studies. Developed and validated by Tong, Sansbury, and Craig (2007), the authors aim to improve the trustworthiness, reproducibility, reporting, and scientific acknowledgement of qualitative research (Kisely & Kendall, 2011). Table 9 presents a sample of the 32 items included in this checklist-reporting framework. For the quantitative methods utilized in this mixed methods design, provision of explicit detail regarding calculations of values included in this finalized document support the scale's content validity and quality, and support the four criteria of trustworthiness in qualitative research (Shenton, 2004).

Table 9

Sample Items from COREQ Checklist

No. Item	Guide
questions/Description Domain 2: study design	
Participant Selection	
10. Sampling	How were participants selected?
11. Method of Approach	How were participants approached?
12. Sample Size	How many participants were in the st
13. Non-participation	How many people refused to participate or dropped out? Reasons?

Note: Data for table taken from (Tong et al., 2007).

Four criteria of trustworthiness in qualitative research. As this study employs qualitative methods to adapt and validate a quantitative instrument, the quality of this research design must be judged by qualitative methods, namely trustworthiness: "Can the findings of this research be trusted?" (Korstjens & Moser, 2018). The terms internal validity, reliability, generalizability, and objectivity used in quantitative research to describe trustworthiness of results are analogous to the terms credibility, confirmability, transferability, and dependability in qualitative research. These terms are the four criteria of trustworthiness in qualitative research (Korstjens & Moser, 2018). Table 10 lists the four criteria of trustworthiness, the definition of each term, the analogous term in quantitative designs, and the element of this study that meets each criterion (Korstjens & Moser, 2018; Shenton, 2004).

Table 10

Four Criteria of Trustworthiness in Qualitative Research

Term	Definition	Anagalous Term	Element in Study Design
Credibility	Establishes plausibility that findings represent content drawn from participants'	Internal validity	·ocess of hnique; Triangulation using different methods of three

	original data and correctly portrays their views		phase design, different types of informants, different sites; examination of previous research to frame findings
Confirmability	Findings can be confirmed by other researchers	Objectivity	Triangulation reduces effect of investigator bias; in- depth description of methods to allow scrutiny; recognition of limitations
Transferability	Using rich descriptions, results can be transferred to other contexts or settings with other respondents	External validity/ generalizability	Detailed description of demographics, study context, and constructs
Dependability	Participants in the study confirm	Reliability	Detailed description of study design

Note. Data for definitions adapted from Korstjens and Moser (2018), and Shenton (2004).

Limitations

The Delphi technique presents a limitation to this study as there are no formal guidelines or techniques established for this methodology: this imparts questioning as to the method's rigor. To attest to the trustworthiness of this design, the four criteria of trustworthiness have been applied to the methodology, as noted in Table 12. Also, the writing of this chapter and the results chapter conform to the "Consolidated Criteria for Reporting Qualitative Studies (COREQ): 32-Item Checklist", a validated checklist to assure quality in reporting (Kisely & Kendall, 2011; Shenton, 2004; Tong et al., 2007).

The researcher was solely responsible for all content analysis methodologies performed throughout phase one of this study and at its culmination, with the generation of items (representing phase two of this design). This presents potential infliction of bias into the GDTA results. Returning verbatim responses within each round of phase one minimized bias of this

individual undertaking. Also, by using the triangulation of different methods, different informants, and different settings, this potential issue is minimized. Disseminating the final list of items to the larger, geographically diverse expert sample further minimized the influence of one coder, providing a greater degree of validity to the results.

The expert panels of this methodology pose considerable bias to the study results. Recruiting only those individuals the researcher deems "expert" does not allow for random selection. Also, those that agree to participate are more likely to have an interest in the research subject, adding additional bias. Triangulation as previously discussed will minimize these potential biases as well.

Conducting this research electronically, or by "e-Delphi" methods, presents limitations and potential risk. Internet is necessary to provide and collect data, presenting a limitation for subjects with limited internet access. Additionally, the true identity of the respondent cannot be confirmed. Consent and surveys were sent to individual email accounts and respondents indicated agreement to consent and terms, however attaining confirmation of true participant identity is nearly impossible.

Further challenges to this technique include the iterative process of the study methods. As the sample for phase one was questioned over three rounds, it is nearly impossible to maintain a subject's complete anonymity to the researcher. As all data collection occurred via email, the identity of the participants was kept from each other, but was potentially accessible to the researcher. Having the research assistant collect all data and de-identifying respondents in REDCap mitigated this limitation.

The primary challenges to adaptation of the SAGAT tool and final project achievement lie in instituting in the expert panel a sense of responsibility and urgency to complete the

requisite questionnaires. A participation incentive provided to the phase two subjects may have increased responsiveness, however the amount or what type of incentive they would receive was not disclosed until they completed the survey. Three attempts at email recruitment solicitation were given to remind and prompt the sample to fully complete and return all questionnaires. Two participants in phase one failed to complete the second round of surveys.

Also, as responses came from all regions of the United States, data collection and analysis can prove challenging. REDCap data capture tools used to create and distribute the online surveys provided a barrier to this limitation. Data collected from the online surveys was exported into IBM SPSS 24.0 for data analysis.

Conclusion

This chapter describes in detail the methodology of this research study aimed at effectively adapting and validating the SAGAT for the SRNA subset as a means to quantify their SA during the simulation of the induction of general anesthesia with associated oral endotracheal tube placement. The Delphi technique, sample selection, the study's three distinct phases, the four criteria of trustworthiness in qualitative research as they apply to this study, and potential limitations with strategies employed to address each were presented.

The remaining chapters present the actual results of this research and a discussion of the findings in terms of feasibility and generalizability. Limitations of the study are presented in chapter five. Finally, related topics to explore further research stemming from this study are also presented. The results section details complete data analysis including results of each round of surveys, details of content analysis at each phase, and explicit description of the quantitative methodologies applied. Content validity index computations, scale context validity computations, probabilities of chance agreement and the factor analysis results are presented.

The final chapter of related discussion includes important points regarding implications of this study, future research relevant to this study, and plans for instrument utilization in the near future.

Chapter Four: Results

The chapters presented thus far underlined the importance of SA as a critical construct of proficient decision-making, with an emphasis on its significance in the anesthesia provider. Situation awareness is a key driver of decision-making, the most important determinant of an anesthesia provider's competence. Historically, health care education assessed competence with written examination of knowledge and clinical skills (task) demonstration. However, true provider competence lies in proficient decision- making, and so understanding the basis of decision-making by quantifying SA becomes crucial.

Evidence demonstrates the SAGAT, the only available direct and objective SA measurement tool, as the ideal measure of SA in event-driven domains. The environment within which anesthesia is delivered is dynamic, with constantly changing circumstances: thus, safe and effective anesthesia care is characterized by the proficient provider response to potential and actual changing patient and operating conditions. Accordingly, the SAGAT is an ideal direct and objective SA measurement tool for this domain. The aim of this study was to adapt and validate this instrument for nurse anesthesia graduate students, and with the results presented within this chapter, the research question, "Can Endsley's SAGAT be adapted and validated to quantify the SA of nurse anesthesia graduate students, or SRNAs, during the simulation of a specific anesthesia event?", is answered affirmatively.

The study followed a sequential exploratory mixed methods design, utilizing qualitative methods to perform GDTA according to Endsley's recommendations (Endsley, 2000). Content analysis of these results exposed SAGAT items specific to the induction of general anesthesia with the associated oral endotracheal tube placement. According to study design, a larger, geographically diverse sample of experts ranked the items according to relevancy with respect to

a provided specific induction of general anesthesia simulation scenario. Quantitative methods performed in data analysis, including item content validity indices, scale content validity indices, and factor analysis, provided further evidence of the items' validity. This chapter presents these respective results, demonstrating the resolution of the research question with an adapted and validated SAGAT to quantify the level of SA in nurse anesthesia graduate students during the induction of general anesthesia with the associated placement of an oral endotracheal tube.

Hypotheses

Several hypotheses underlie the study's research question, all relevant to the adaption and validation of SAGAT items to the nurse anesthesia graduate student population. The SAGAT quantifies SA by testing with items specific to SA level, as per Endsley's definitions (Level 1, 2, and 3 SA): perception of data/information, comprehension of its meaning, and the ability to project what may happen based on this information, all within an evolving, dynamic event., respectively. This study tests the following hypotheses, all identified as "H":

H1a: Level 1 SA requirements related to the induction of general anesthesia with oral endotracheal tube placement will be identified through the GDTA process.

H1b: Level 1 SA requirements related to the induction of general anesthesia with oral endotracheal tube placement will be validated through the GDTA process.

H2a: Level 2 SA requirements related to the induction of general anesthesia with oral endotracheal tube placement will be identified through the GDTA process.

H2b: Level 2 SA requirements related to the induction of general anesthesia with oral endotracheal tube placement will be validated through the GDTA process.

H3a: Level 3 SA requirements related to the induction of general anesthesia with

oral endotracheal tube placement will be identified through the GDTA process.

H3b: Level 3 SA requirements related to the induction of general anesthesia with oral endotracheal tube placement will be validated through the GDTA process.

All hypotheses were tested using qualitative methods, with quantitative methods applied to the data analysis at the end of the study for further validation.

Qualitative Results

Adhering to Endsley's (2000) recommendations, the first phase of study design aimed at performing GDTA for the induction of general anesthesia with the associated oral endotracheal tube placement. A sample of seven CRNA, nurse anesthesia educators completed the GDTA and validated the group consensus results in three rounds. All subjects completed each round, except during the second round in which two subjects failed to respond to the survey. These subjects did, however, complete the third round. In each round, some survey questions were left blank without rationale.

Table 11 delineates subject identification of sub-goals for round one, which consisted of 4 open-ended questions asking for the identification of sub-goals, delineation of key decisions to meet the sub-goal, and the SA required to make those decisions. Table 12 provides the number of missing responses by each subject for subsequent rounds two and three of this Delphi process. Detailed descriptions of the subjects involved in this study are provided in the previous chapter.

Round one. After obtaining consent, an open-ended questionnaire asked participants to establish the sub-goals (termed "goals" in the questionnaire) of the induction of general anesthesia with the associated placement of an oral endotracheal tube. The questionnaire was created in and distributed by REDCap electronic data collection tools. Subjects were not limited in the number

Table 11

Subject Response Rate for GDTA Open-Ended Questions Round One of Delphi

Subject	Number of Sub-Goals Identified
1	2
2	6
3	1
4	1
5	9
6	4
7	2

Note: All participants completed the complete GDTA for each sub-goal identified. GDTA tree includes sub-goal, key decisions to obtain sub-goal, requisite SA for key decisions

Table 12

Subject Unanswered Questions for Rounds One and Two of Delphi Process

Participant	Round 2 Quest. Unanswered/ 181 Total Q	Round 3 Quest. Unanswered/ 185 Total Q
1	5	1
2	10	1
3	2	2
4	0	3
5	10	1
6	*	1
7	*	1

Note. Quest. and Q = questions; * = did not respond to survey request

of sub-goals they could include. For each sub-goal a subject provided, additional open-ended questions asked for the key decisions related to achieving the sub-goal. Endsley's definitions for SA Level 1, 2, and 3 were provided and subjects were asked to then list the SA requirements for each key decision by corresponding SA level. No limits were given as to how many responses subjects could submit. Figure 7 depicts a section of the questionnaire asking for a "goal", related key decisions, and SA requirements. After each GDTA, subjects were asked if they could identify another goal. If they respond yes, the survey continues repeating the same series of open-ended questions. This questioning process is repeated until the respondent answers "no", indicating they have no additional goals to include.

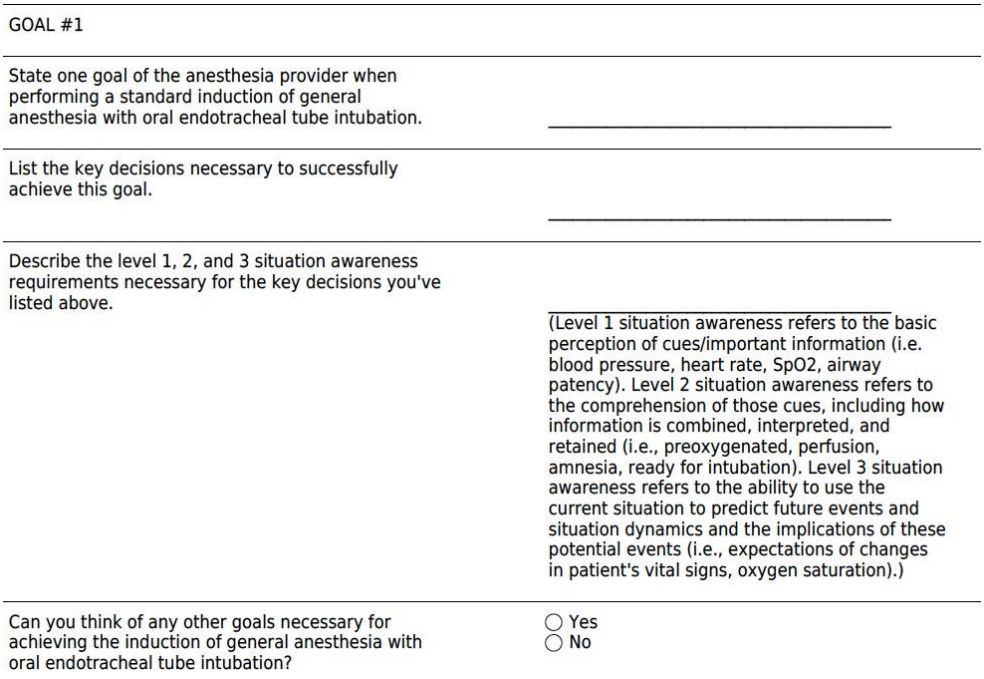


Figure 7. Depiction of open-ended questions for round one of GDTA survey.

Subjects in this initial round provided valuable data, each with similar responses of "goals", and detailed key actions for each goal. Table 13 depicts the identified sub- goals with

respective number of key decisions and SA requirements by level as identified by respondents. Each subject's responses were manually transcribed verbatim, using the exact words they provided in the response free text sections. All of the respondents' sub- goals with the respective GDTA "tree" for each sub-goal identified were then compared manually. Content analysis of each GDTA revealed common themes among a consensus of respondents. Figure 8 illustrates a sample of the verbatim transcription of responses collected from round one as well as the content analysis for sub-goal themes. All goals and their corresponding GDTA "tree" with identified key decisions and related SA requirements by level were manually entered into REDCap for the second round of survey.

Table 13

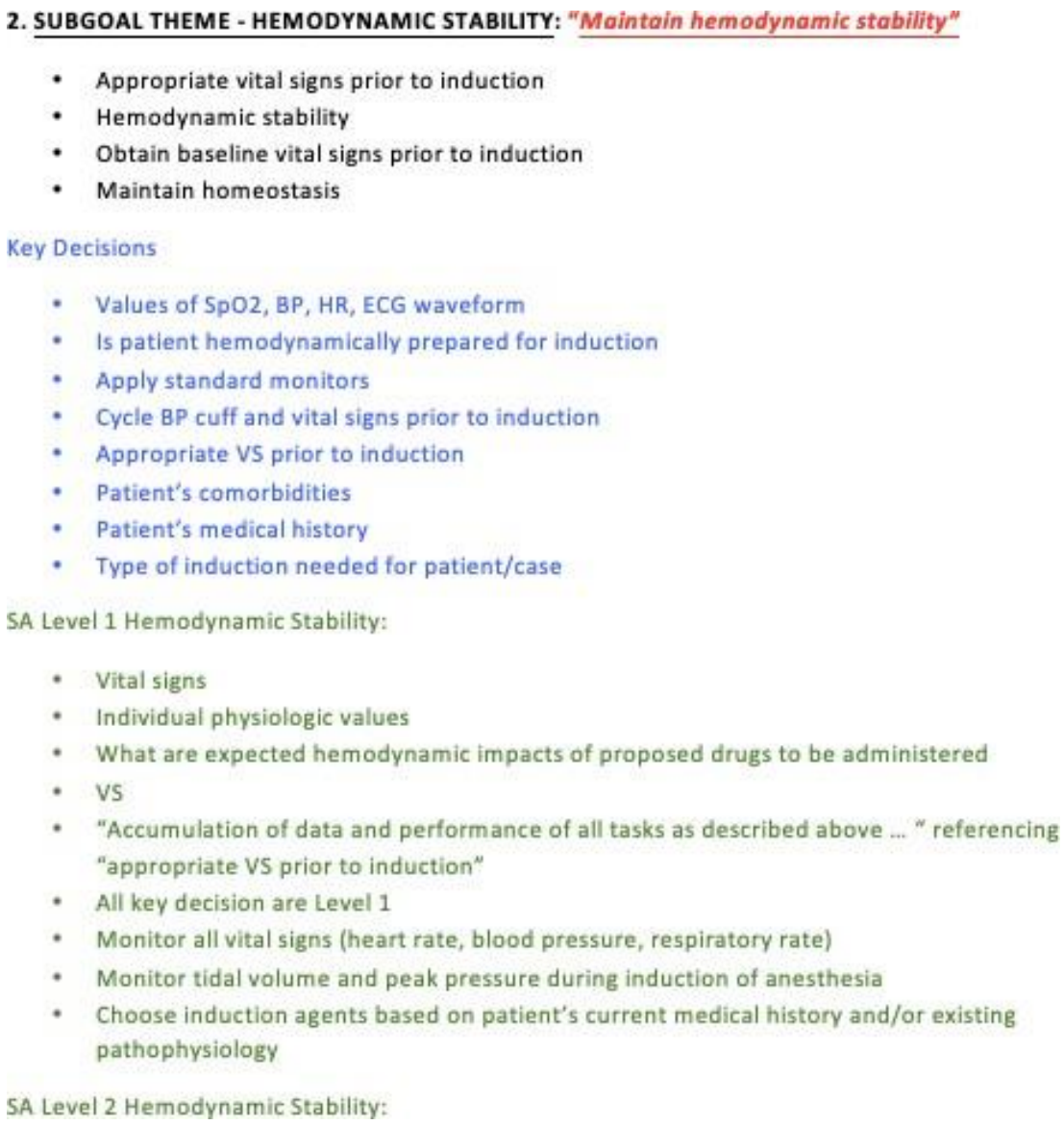
Round 2 Identified Sub-Goals, Key Decisions, and SA by Level.

Sub-Goal	Key Decisions	SA Level 1	SA Level 2	SA Level 3
Equipment/drug preparation	4	0	0	3
Hemodynamic stability	5	6	5	5
Induce anesthesia	7	0	4	7
Respiration/ Ventilation	7	7	12	4
Anesthesia/ induction plan-assess patient	7	13	2	1

Note. Sub-goals, key decisions, and SA requirements by level for this table and utilized in round two survey are direct results from round one survey.

Respondents' demonstrated initial difficulty identifying SA requirements, as evidenced by the lack of responses for some SA level needs for each sub-goal (refer to Table 13). As

respondents were exposed to the GDTA process in the next round, it became evident that revealing the other panelists' responses helped formulate a better understanding of SA, evident by the richness of data in subsequent rounds.



2. SUBGOAL THEME - HEMODYNAMIC STABILITY: *"Maintain hemodynamic stability"*

- Appropriate vital signs prior to induction
- Hemodynamic stability
- Obtain baseline vital signs prior to induction
- Maintain homeostasis

Key Decisions

- Values of SpO2, BP, HR, ECG waveform
- Is patient hemodynamically prepared for induction
- Apply standard monitors
- Cycle BP cuff and vital signs prior to induction
- Appropriate VS prior to induction
- Patient's comorbidities
- Patient's medical history
- Type of induction needed for patient/case

SA Level 1 Hemodynamic Stability:

- Vital signs
- Individual physiologic values
- What are expected hemodynamic impacts of proposed drugs to be administered
- VS
- "Accumulation of data and performance of all tasks as described above … " referencing "appropriate VS prior to induction"
- All key decision are Level 1
- Monitor all vital signs (heart rate, blood pressure, respiratory rate)
- Monitor tidal volume and peak pressure during induction of anesthesia
- Choose induction agents based on patient's current medical history and/or existing pathophysiology

SA Level 2 Hemodynamic Stability:

Figure 8. Depiction of transcription of round one responses collected and the content analysis for sub-goal theme (response sub-goal titles provided below sub-goal theme title).

Round two. REDCap electronic data collection tools were again utilized to administer the second round of surveys. Content analysis of sub-goal derived themes and each respondent's given goal responses were collapsed into a consensus theme. For example, any sub-goal provided related to maintaining hemodynamic stability during induction was collapsed into the sub-goal theme, "Maintain Hemodynamic Stability". Expert panel sub-goals were delineated in this questionnaire round, with each sub-goal followed by a list of related key decisions elicited by all participants. Figure 9 illustrates a sub-goal and its identified key decisions with relevant agreement/disagreement response selection choices as presented in this questionnaire round. Subjects were asked to agree or disagree with each key decision and were given the opportunity to provide comments with each response. At the end of the list of provided key-decisions, subjects were given the chance to add more key decisions after viewing the entire list of responses. Figure10 demonstrates the open-text response made available for additional key decisions identified.

"Maintain hemodynamic stability" is the second sub-goal you have identified to achieve the induction of general anesthesia. Listed below are the key decisions you have identified to successfully achieve this sub-goal. Do you agree with these key decisions?	
Values of SpO2, BP, HR, ECG waveform	◯ Yes ◯ No
If no, please comment	_____
Is patient hemodynamically prepared for induction	◯ Yes ◯ No
If no, please comment	_____
Apply standard monitors	◯ Yes ◯ No
If no, please comment	_____
Cycle BP cuff and vital signs prior to induction	◯ Yes ◯ No
If no, please comment	_____

Figure 9. Questionnaire format for identified sub-goal, requisite key decisions, and relevant agreement/disagreement response selection choices.

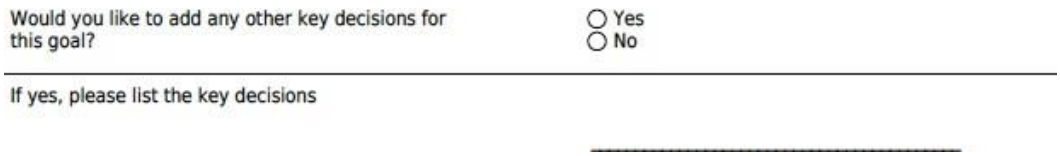


Figure 10. Open-text response section for any additional key decisions identified.

This second round of data received from REDCAp electronic data collection again was analyzed manually. Each response requiring agreement or disagreement was tallied. Two participants did not complete this round. There was no rationale provided by these subjects to the researcher or research assistant. Additionally, episodes of missing data occurred: some respondents did not provide answers for every question of this survey. Agreement on responses was manually tallied in the same manner previously described, with the total respondents provided as the denominator. Any free text responses were transcribed verbatim. Content analysis performed on all responses identified similar responses using different wording and these answers were collapsed into a common content theme.

The same methods described for data analysis of responses related to the key decisions identified were applied to the responses related to levels of SA required for each decision. All responses were hand tallied for consensus and all free text responses were transcribed verbatim. The subjects were given the opportunity to add more Level 1, 2, and 3 SA requirements if they wished. Figure 11 depicts the questionnaire items for levels of SA, response choices, and area for adding more SA requirements.

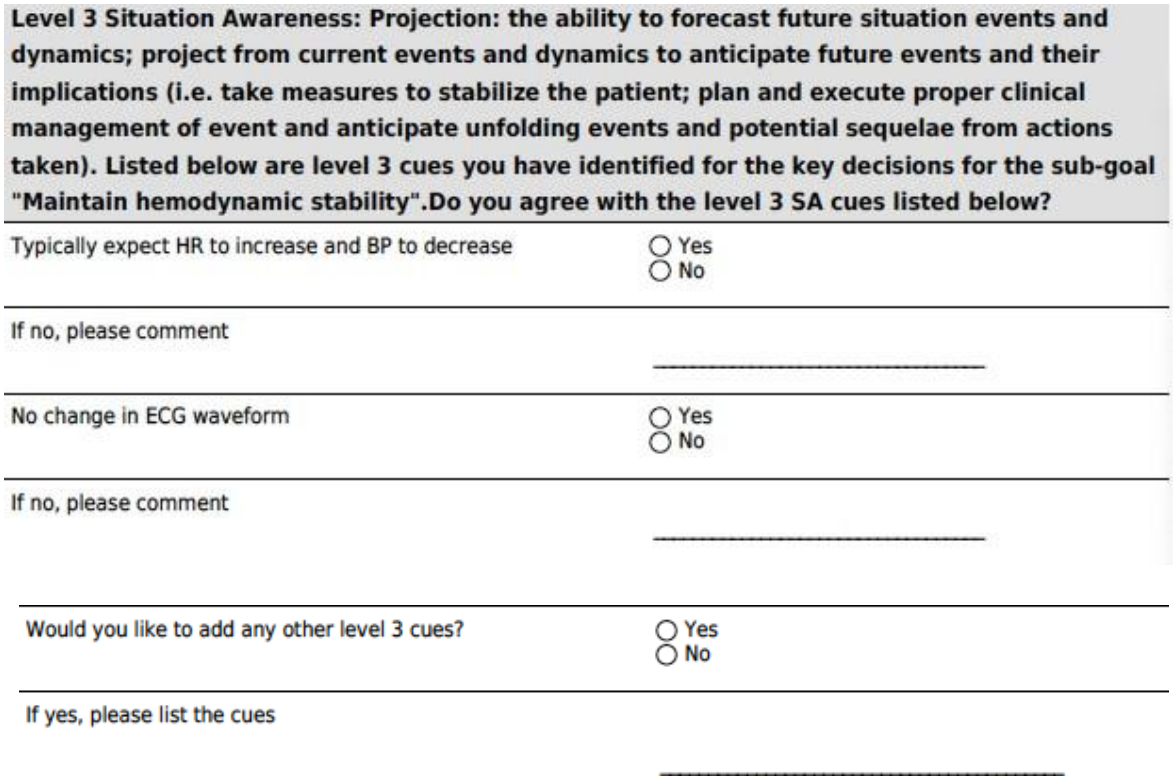


Figure 11. Illustration of questionnaire seeking consensus on identified SA requirements according to level of SA and open-text box for any identified additional SA requirements.

Round three. For the third and final round of this initial phase of the study, all manually tallied responses were fed-back to participants. All seven participants completed this final round. Tallied responses were provided in different font color to make them more easily identifiable. Missing responses were identified by tally of "no response yes or no given for this item". Figure 12 illustrates a tally of non-responders to an answer choice.

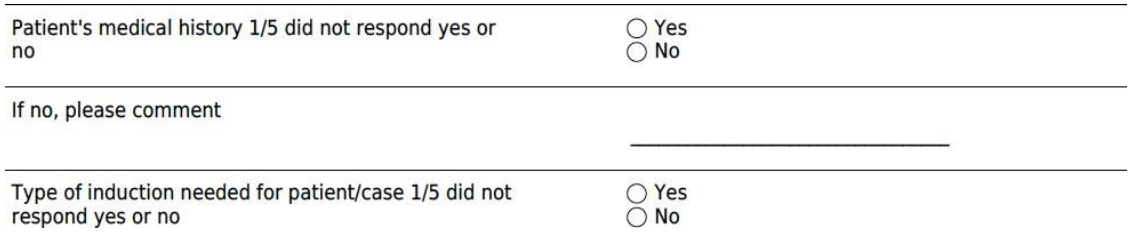


Figure 12. Demonstration of manual tally of non-responders to question in feed-back to subjects.

If some disagreed with an item, a tally was given with the number of subjects in

disagreement relative to the total number of respondents. Subjects again were asked to agree or disagree with responses, in the same format as round two, but this round allowed them to see the number of dissenters, any free text rationales provided, and any responses added from the previous round. Figure 13 depicts a third round question item demonstrating number of dissenters and any provided rationale.

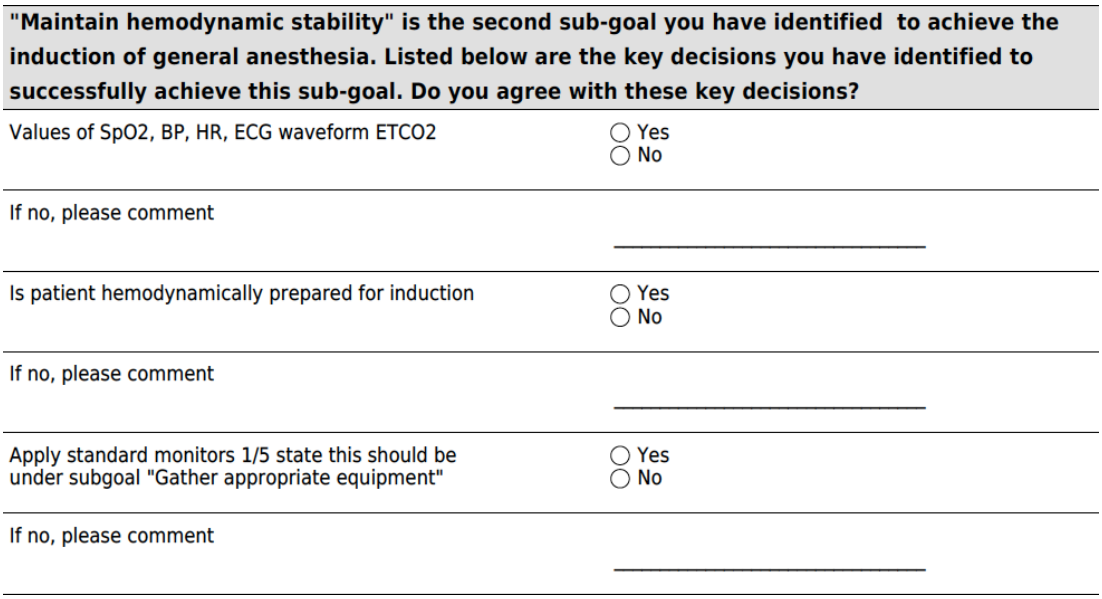
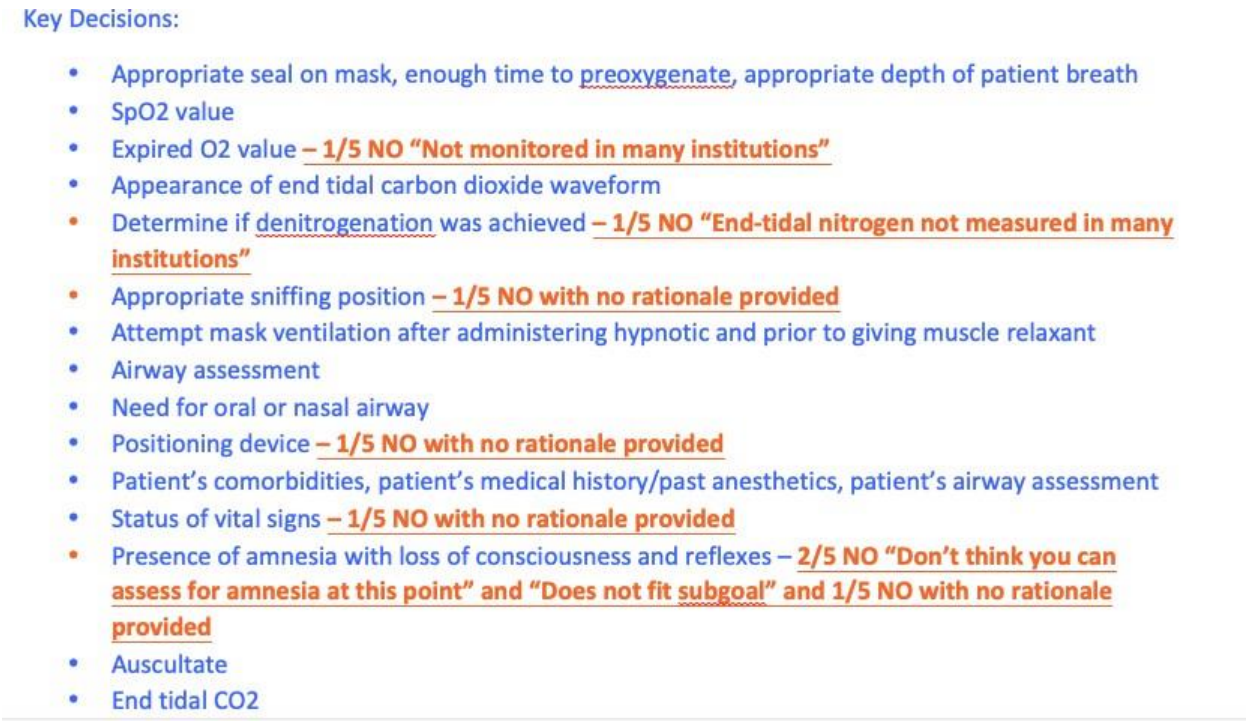


Figure 13. Third round manual tally of dissented items. The tally on this particular item reveals that only five subjects answered this item in round two, and one person disagreeing with the key decision identified.

The third round responses were documented in REDCap electronic data collection tools and manually analyzed. Consensus was determined on each response and those meeting a minimum of 70% consensus were manually transcribed verbatim. This manual transcription used color-coding for sub-goals (black), key decisions (blue), and SA requirements (green). This aided in the visual representation of the GDTA tree. Each transcribed response demonstrating disagreement depicted the tally of dissenters in red. This same technique was used if there was missing data: a tally of no response was indicated in red next to the item. Figure 14 illustrates the

data from the manual tally of responses used for the questionnaire in round three.


Key Decisions:

- Appropriate seal on mask, enough time to preoxygenate, appropriate depth of patient breath
- SpO2 value
- Expired O2 value – 1/5 NO "Not monitored in many institutions"
- Appearance of end tidal carbon dioxide waveform
- Determine if denitrogenation was achieved – 1/5 NO "End-tidal nitrogen not measured in many institutions"
- Appropriate sniffing position – 1/5 NO with no rationale provided
- Attempt mask ventilation after administering hypnotic and prior to giving muscle relaxant
- Airway assessment
- Need for oral or nasal airway
- Positioning device – 1/5 NO with no rationale provided
- Patient's comorbidities, patient's medical history/past anesthetics, patient's airway assessment
- Status of vital signs – 1/5 NO with no rationale provided
- Presence of amnesia with loss of consciousness and reflexes – 2/5 NO "Don't think you can assess for amnesia at this point" and "Does not fit subgoal" and 1/5 NO with no rationale provided
- Auscultate
- End tidal CO2

Figure 14. Illustration of manual data tally techniques utilized from round two data collection. These manual tallied results used as basis for round three survey questionnaire.

Creation of query items. The conclusion of round three represented the end of data collection for phase one of this research study design. A manual tally of results revealed five sub-goals each with a GDTA tree all identified by expert consensus. Each individual data response was analyzed for consensus. Consensus was determined at a minimum of 70% agreement. The identified SA Level 1, 2, and 3 requirements were split from the rest of the GDTA tree for each sub-goal. This list of requirements became the basis for SAGAT query development. The items developed utilized specific language from the SA requirements as determined by group consensus. Figure 15

depicts a sample of items by level of SA developed from the output of the round three survey.

SAGAT queries for Sub Goal "Maintain Hemodynamic Stability"

Level 1 (perception):

- What is the patient's current heart rate?
- What is the patient's current blood pressure?
- What is the patient's current respiratory rate?
- What medications are you giving for induction?

Level 2 (comprehension):

- Is the patient's heart rate responding adequately to the drug you have just given?
- Is the patient's blood pressure responding adequately to the drug you have just given?
- Does the patient's heart rate and blood pressure indicate a need for additional induction agent dosing?
- Does the patient's response to induction require any other intervention?

Level 3 (projection):

- Do you expect any change in the patient's heart rate in the next minute?
- Do you expect the patient's blood pressure to increase, decrease or remain the same in the next minute?
- How do you think the patient's heart rate will respond to intubation?
- How do you think the patient's blood pressure will respond to intubation?

Figure 15. Sample of items, in order by SA level according to survey results, and developed by content analysis of round 3 survey results.

A preliminary list was developed of 113 items from the five agreed upon sub-goals. Separation of items by level of SA (SA Level 1, 2, and 3) proceeded directly from the round three survey results. Two of the identified sub-goals were set aside and not used in the final list of items. These sub-goals, one for the preparation (or gathering) of equipment and preparation for induction and the second for preanesthetic patient assessment, fell out of the scope of this SAGAT testing target: inducing general anesthesia and the associated placement of an oral endotracheal tube. Though proper preparation of the patient and the room is vital to a safe induction of anesthesia, this specific SAGAT aims to test SA during the actual period of induction of anesthesia and associated oral endotracheal tube placement. Therefore, only those sub-goals and relative SA Level 1, 2, and 3 requirements specific to the action of induction of

general anesthesia with associated oral endotracheal intubation were utilized.

The list of 39 SAGAT items derived from this process became the questionnaire content for the next phase of the study. A sample selection of created items and the resultant survey is illustrated in Figure 16 of the quantitative results. It is our plan to utilize the remaining list of SA requirements specific to the sub-goals not considered for this SAGAT in future SAGAT development: one specific to patient preoperative assessment, and the other for anesthesia machine check and the preparation of medications and equipment.

Quantitative Results

A large geographically diverse sample of CRNA, nurse educators from the United States ranked the 39 items by degree of relevancy based on a provided specific induction of general anesthesia simulation scenario. Forty-nine of those recruited agreed to participate in this phase of the study, with 40 of those participating in the item ranking survey instrument. Nine subjects only completed the demographic survey, with no rationale provided for not attempting the accompanying item relevancy survey instrument. Figure 16 illustrates a sample of these items as presented in the survey. This survey utilized Likert scale rankings from “not relevant” (coded as “0”) to “highly relevant” (coded as “3”). Content validity index calculations were performed on the results of this survey for each of the 39 items. Calculations of scale content validity indices were also performed. Additional quantitative methods to further validate items included an exploratory factor analysis, which will be further described in this chapter.

Item content validity index. Calculations performed for each individual item provided a validity index (I- CVI), which is a measure of inter-rater agreement that is used in a large majority of nursing research to establish content validity. Following guidelines provided by Polit, Beck, and Owen (2007), content validity is established with a minimum I-CVI of 0.83 when using eight or more experts.

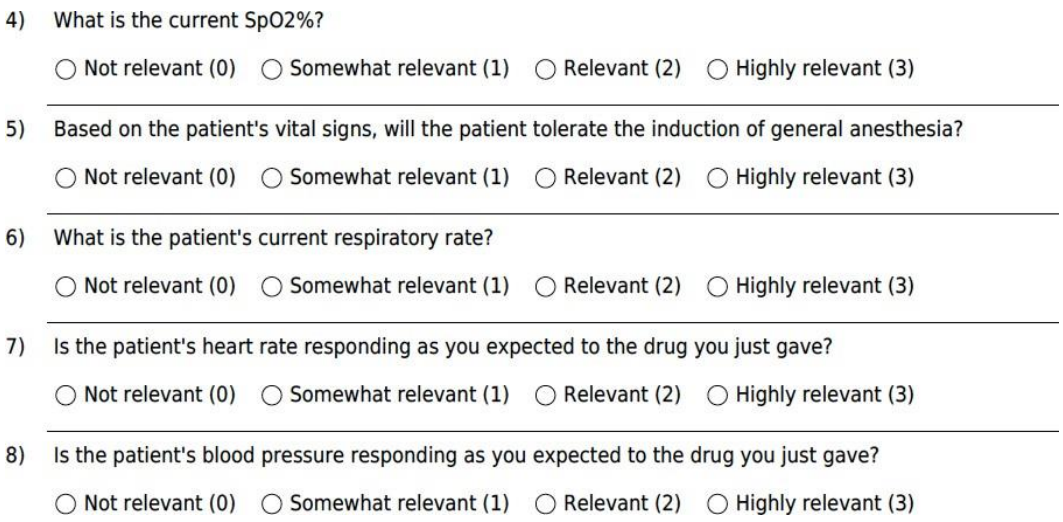


Figure 16. Sample of items for survey assessing item relevancy for research study phase two.

content validity is established with a minimum I-CVI of 0.83 when using eight or more experts. Six of the 39 items were ranked by 39 of the 40 expert subjects, with the remaining 33 items ranked by all 40 participants. Of the 39 items, two resulted in an I-CVI less than the minimal threshold, with calculated I-CVIs of .74, and .80 respectively. Table 14 provides the calculated I-CVI for all 39 items.

Table 14

I-CVI calculations for all items

Item No.	No. of Experts	No. of "High Relevancy"* Ratings	I-CVI
1	40	37	0.93
2	40	39	0.98
3	40	40	1.00
4	40	40	1.00
5	40	37	0.93
6	40	34	0.85
7	39	37	0.95
8	40	39	0.98
9	39	38	0.97
10	40	39	0.98
11	39	39	1.00
12	40	35	0.88
13	40	33	0.83
14	40	36	0.90
15	40	35	0.90
16	40	36	0.90
17	40	38	0.95
18	40	40	1.00
19	40	34	0.85
20	40	36	0.90
21	40	40	1.00
22	40	34	0.85
23	40	36	0.90
24	40	39	0.98
25	39	38	0.97
26	40	34	0.85
27	40	40	1.00
28	40	34	0.85
29	40	38	0.95
30	40	40	1.00
31	40	37	0.93
32	40	39	0.98
33	40	35	0.88
34	40	38	0.95
35	39	37	0.95
36	40	40	1.00
37	39	29	0.74**
38	40	32	0.80**
39	40	34	0.85

Note. No. = number; "High Relevancy"* = ranked as relevant (2) and highly relevant (3);
** I-CVI < minimum 0.83

Scale content validity (S-CVI) calculations utilized the S-CVI/Ave, which averages all I-CVIs (Polit et al., 2007). The measure of S-CVI/Ave for all items was calculated at 0.92, meeting criteria for excellent content validity (O.90). Table 15 depicts the S-CVI/Ave for the 39 items.

Table 15

SCVI/Ave

No. of Items	No. of Experts*	S-CVI/Ave Calculated
39	40	0.92

Note. No. = number; 6 items had 39 experts respond but 40 experts used for this calculation; S-CVI/Ave > 90 = excellent scale content validity

Exploratory factor analysis. Thirty-four out of 40 subjects completed the entire survey of 39 questions. Table 16 presents the item mean scores in rank order based on responses from all 40 participants in the sample. Mean rankings ranged from 2.0 to 2.93, indicating items even at the lowest end of the ranking list were somewhat relevant to this sample.

Table 16

Rank Order of All 39 Items by Mean Score

Item	Rank	*N*	Mean	Standard Deviation
10	1	40	2.93	.350
4	2	40	2.92	.267
11	3	39	2.85	.366
27	4	40	2.82	.385
32	5	40	2.80	.464
5	6	40	2.78	.577
9	7	39	2.77	.485
21	8	40	2.75	.439
8	9	40	2.72	.506
3	10	40	2.65	.483
2	11	40	2.65	.533

17	12	40	2.65	.580
34	13	40	2.63	.586
18	14	40	2.62	.490
25	15	39	2.62	.544
31	16	40	2.60	.632
30	17	40	2.58	.501
7	18	39	2.54	.600
20	19	40	2.53	.679
24	20	40	2.50	.555
12	21	40	2.50	.784
1	22	40	2.48	.716
29	23	40	2.45	.597
36	24	40	2.43	.501
19	25	40	2.43	.747
33	26	40	2.40	.709
13	27	40	2.40	1.057
35	28	39	2.38	.590
14	29	40	2.38	.667
16	30	40	2.37	.667
22	31	40	2.35	.975
15	32	40	2.32	.694
28	33	40	2.30	.791
6	34	40	2.20	.758
26	35	40	2.20	.758
23	36	40	2.17	.675
39	37	40	2.13	.723
38	38	40	2.05	.749
37	39	39	2.00	.918

Exploratory factor analysis with principal components was conducted to determine if the items on the survey combine into thematic factors. The initial EFA indicated that the correlation matrix was not positive and therefore inappropriate for analysis. Six items (numbers 1, 11, 19, 21, 31, and 32) were found to have low item-total correlations (< .30) with the other items on the survey and were excluded from further analysis. After excluding these six items, the correlation matrix was found to be adequate for EFA with a Bartlett's Test of Sphericity p value < .001. Of note, the Kaiser-Meyer-Olkin (KMO) Measure of Sampling Adequacy was .204, less than the threshold.

The scree plot indicated two factors explaining 46.5% of the total variance, one with 36.8% of the variance and the other 9.7% (refer to the scree plot illustrated in Figure 17). The two factors were correlated with an r = .502, leading to the use of an oblique promax rotation to interpret the factors. The rotated pattern matrix with the factor loadings revealed five additional items to exclude based on the criterion that only items with factor loadings ≥ .40 are sufficiently associated with a factor (Stevens, 2002). The pattern matrix is shown in Table 17. Eighteen items were found to load on the first factor, goal-driven processing, and 10 items on the second, data-driven processing. Five items revealed no substantial loading on either factor and were subsequently extracted. Cronbach's alpha of the final 28 items was .937, which is considered excellent reliability for this group of queries. The results of this EFA, in addition to the content validity indices, demonstrate positive confirmation of the study's hypotheses.

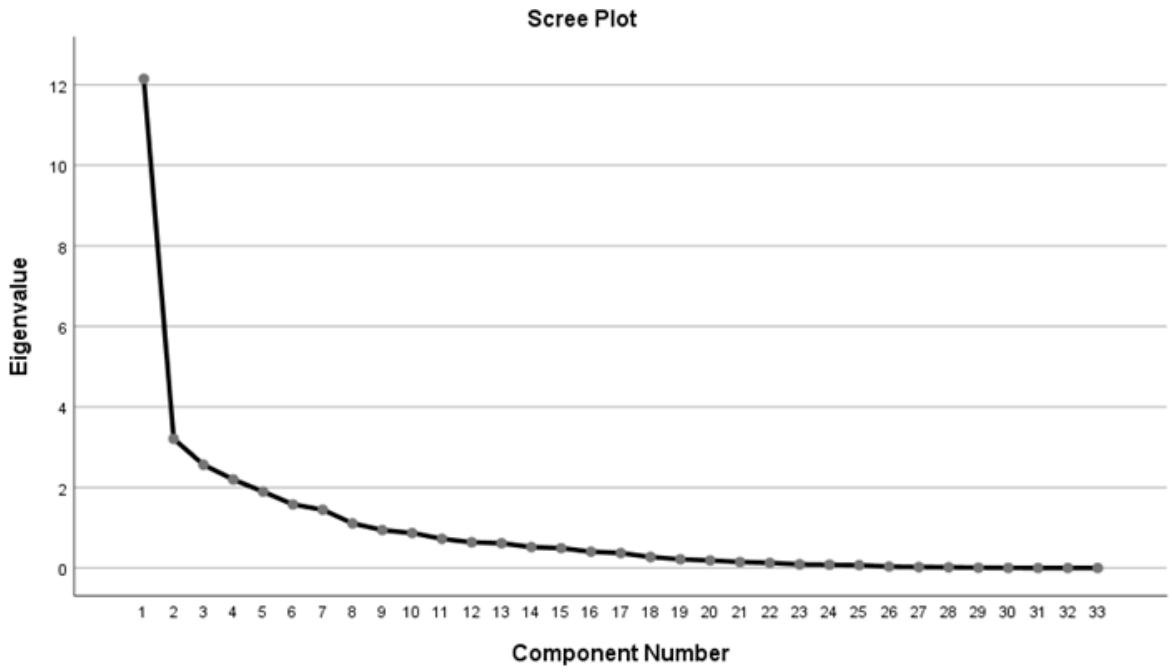


Figure 17. Scree test indicating survey comprised of two factors with 46.5% of variance.

Table 17

Pattern Matrix

Item Number	Data Driven	Goal Driven
29	.931	-.421
25	.865	-.344
24	.858	-.046
39	.809	.204
26	.781	.049
10	.772	-.328
27	.713	-.494
38	.696	.268
15	.690	.156
28	.679	.065
35	.653	.190
14	.650	.183
7	.619	.337
16	.539	.321
8	.536	.310
2	.490	.114
30	.449	.268
36	.437	.353
34*	.344	.085
13*	.301	.003
17*	.278	.137
5	-.333	.910
4	-.155	.831
9	-.207	.686
6	.248	.655
23	-.052	.623
3	.078	.534
33	.341	.528
37	.323	.468
22	-.064	.448
12	.010	.425
20*	.133	.325
18*	.193	.280

Note. * = did not reach or exceed .400

Conclusion

This chapter presented the results of both the qualitative and quantitative techniques used to positively answer the research question: "Can Endsley's SAGAT be adapted and validated to quantify the SA of nurse anesthesia graduate students, or SRNAs, during the simulation of a specific anesthesia event?" The mixed methods approach utilized in conducting this research provides rich results for both the validity and reliability of results as well as the replication of study design. The qualitative methods ensured face and content validity of results, while the applied quantitative methods gave additional support to their veracity.

The qualitative methods including Delphi expert surveys and ensuing content analysis generated validated items for an anesthesia specific adapted SAGAT. The quantitative methods providing item content validity and scale content validity indices demonstrated excellent content validity of 37 of the 39 items, and good content validity for the other two of three items. The entire grouping of items demonstrated excellent content validity. An EFA further confirmed content validity despite a small sample size. These findings support the use of these items in a SAGAT for quantifying the level of SA in SRNAs during the simulation of the induction of general anesthesia with associated oral endotracheal tube placement. The next and final chapter presents these results in relation to the research study's objectives, the limitations of the study and findings, and offers recommendations for future research.

Chapter Five: Discussion

Thus far, the background underlying this research study, a review of the literature supporting its necessity, and the methods utilized to conduct the research were presented. Chapter's one and two laid the foundation for this study's need and timeliness in patient safety efforts, particularly in the anesthesia (and nurse anesthesia subset) specialty. Educational and training efforts' current focus on provider competence and its weight over the ability to perform tasks and rote knowledge in ensuring safe, effective, and efficient health care is underscored throughout the preliminary chapters of this manuscript. The thorough literature review reveals how the importance of provider competency grew to its current focus in human factors and patient safety realms, creating the need for valid, direct measures of such aptitude. These chapters also provided the indisputable link between provider competence and clinical decision making, with the construct of situation awareness placed at the vanguard.

The methods used and the study's results with supporting documentation were included in the discussion of the previous two chapters. The richness and depth of description of the qualitative methods, which predominate this design, bear evidence to the validity and reliability of outcomes. The multi-modal quantitative approach conveys further evidence as to validity of results. This chapter now presents the study's conclusion in terms of the hypotheses delineated in preceding chapters, with a discussion of these results relevant to the research question. This chapter concludes with an account of the limitations and suggestions for future research.

Background of the Problem

Emerging research points to SA as a critical cognitive construct for patient safety and positive outcomes (Jones et al., 2018). The ability to constantly interpret key information and make accurate projections from a continuously changing situation is the hallmark of high level

SA, according to Endsley, the author of the most widely accepted SA theory, "Theory of Situation Awareness" (Endsley, 1995a). Situation awareness' importance as a construct for the anesthesia provider originates with David Gaba and his 1995 publication with Howard and Small, "Situation Awareness in Anesthesia" (Gaba et al., 1995). In this pivotal work on SA, Gaba et al. conclude that in dynamic situations, such as those occurring in anesthesia, the ability to perceive and read clues in an ever- changing environment, the essence of good SA, is crucial to decision-making. This work presents the link between anesthesia and aviation and the need for the emulation of aviation's simulation and crisis management training to ultimately realize an improvement in patient outcomes.

Situation awareness develops with experience and training, and this is particularly true of anesthesia providers. Endsley underscores the relationship between the ability to rapidly attain and maintain high level SA, and experience and training (Endsley, 1995a). This relationship accentuates the importance of simulation in nurse anesthesia education. For more than a decade, simulation's role in anesthesia education and training has been increasingly aimed at providing practice in patient management without patient harm. Beginning with Gaba's work in 1995, the goal of incorporating simulation training has evolved to its current focus on the education and training of human factors' constructs in relation to medical error and patient safety.

Based on the growing evidence of simulation training's positive impact on patient safety and optimal clinical management, the Council on Accreditation of Nurse Anesthesia Programs (COA) stipulated in the "Standards of Accreditation of Nurse Anesthesia Programs – Practice Doctorate", Curriculum Standard 11, that, "Simulated clinical experiences are incorporated into the curriculum" (Council on Accreditation of Nurse Anesthesia Educational Programs, 2018, p. 22). To support the incorporation of these changes in nurse anesthesia curricula, valid measures

are needed to support the impact of simulation on training and performance. As the realization of the importance of health care providers' SA grows, particularly for those in the anesthesia specialty, and is now highlighted in current education and training, a means to measure this construct is of paramount importance.

As important as SA is as a human factors' construct, so is the ability to assess the SA of operators who function in complex systems. Measuring SA has become the focus of researchers since its importance in safety was established in the literature from the late 1980s through 1995. Situation awareness measurement has been attempted by both direct and indirect approaches. Direct measures refer to those methods that purport to quantitatively assess an individual's SA, and indirect methods measure behavioral and performance outcomes, thereby making inference to an individual's SA. Finding instruments that quantify performance by computing level of SA (direct measures) during simulation is important but difficult (Wright, 2004). Currently available measurement tools are subjective, relying on expert rater observations of behavior to make assessments, or they utilize retrospective self-report. Adding to the drawbacks of these tools, the use of expert raters introduce high costs on measurement, as there is a price for expert rater observation and his or her time to perform an assessment.

To date, the only direct and objective SA measurement tool is the "Situation Awareness Global Assessment Technique", or SAGAT, developed by Mica R. Endsley. This tool was originally created for the measurement of SA in military fighter pilots. A literature search revealed that this instrument has been adapted and validated in many domains, including aviation, nuclear power control, automotive, and most recently in health care specialties such as trauma, emergency medicine, nursing, and for use with medical residents.

Purpose of Study

An exhaustive review of the literature reveals that there is no SAGAT adapted and validated for the anesthesia domain. The purpose of this study was to adapt and validate the SAGAT for the anesthesia domain, and specifically to measure the level of SA in the SRNA subset. Using Endsley's "Theory of Situation Awareness" as its underpinning, this study followed her recommendations in performing a goal-directed task analysis to adapt SAGAT items and then validated these items by mixed methods approaches.

Theoretical Foundation

Endsley first described the "Theory of Situation Awareness" in her 1995 work "Toward a Theory of Situation Awareness in Dynamic Systems" (Endsley, 1995a). In this publication, Endsley describes three levels of SA that impact appropriate decision- making and performance in dynamic environments: perception of information; comprehension of its meaning; and based on the current situation, projection of possible future events (Endsley, 1995a). She further outlines the destructive effect of inaccurate SA on outcomes and the impact SA has on effective team cohesion and dynamics. Since its debut, "The Theory of Situation Awareness" has been heralded in human factors' research as the backbone of safety and effective situation management (Wickens, 2008).

Situation awareness is a component of a feedback system in which decisions are made based on observed data and predicted outcomes of events (Endsley, 1995a, 2015). As decisions are made and executed, the outcomes further impact data and continued decision-making such that every taken action has an impact on the situation and observable data, and subsequently the next planned action (Endsley, 2015; Schulz et al., 2013). Decisions are based on past experience and training, and the execution of these decisions are based on abilities, stress, and workload

(Schulz et al., 2013).

Applying Endsley's theory to the health care domain, situation awareness is a critical determining factor of clinical judgment (Lavoie et al., 2016). Clinical decision- making can be related to the levels of SA described by Endsley (Wright et al., 2004). As an extension of this process to the anesthesia domain, and specifically the SRNA subset, sound patient management decisions can be found in higher order SA (level 3). Higher order SA is crucial to effectively manage patients in a dynamic physiological state, such as when anesthetized, within a dynamic environment, for example the operating room.

This study was designed to adapt and validate an instrument that will for the first time quantify the level of SA in SRNAs, which can then be used to appropriately guide their education and training in an effort to enhance SA and improve clinical decision- making. The research question of this study was: "Can Endsley's SAGAT be adapted and validated to quantify the SA of nurse anesthesia graduate students, or SRNAs, during the simulation of a specific anesthesia event?"

Methods

A sequential exploratory mixed methods design, utilizing qualitative methods to perform GDTA according to Endsley's recommendations (Endsley, 2000), was used to adapt and validate the SAGAT instrument for nurse anesthesia graduate students. Content analysis of these results exposed SAGAT items specific to the induction of general anesthesia with the associated placement of an oral endotracheal tube. According to study design, a larger, geographically diverse sample of experts ranked the items in terms of their relevancy to a specific simulation scenario. Quantitative methods performed in data analysis, including item content validity indices, scale content validity indices, and factor analysis, provided further evidence of the

items' validity.

Review of Results

After IRB human subjects' deliberation and receipt of exempt review status, the first phase of the study utilized Delphi methods and an expert panel of seven subjects to complete the GDTA. Content analysis performed on these results led to the creation of SAGAT items that were then rank ordered for relevancy by the larger expert panel of 40 participants. Content validity indices and scale content validity calculations indicated excellent content validity for 37 of the 39 items, and for all items considered as one scale. Aside from the two items that did not meet the content validity index (I-CVI) threshold of 0.83, I-CVI's ranged from 0.83 to 1.0. Scale content validity, calculated as the average of all I-CVI's (S-CVI/Ave) was 0.92, exceeding the threshold for excellent validity of 0.90.

Exploratory factor analysis (EFA) was then performed, and after extraction of six items with very low correlation to the rest of the grouping, EFA with principal components extraction produced a Bartlett's Test of Sphericity that was significant at 0.00, indicating adequacy of items for the EFA. A scree test indicated two factors representing 46.5% of the shared variance, 36.8% and 9.7% respectively. An orthogonal rotation revealed a pattern matrix with items loading on two factors, accounting for the 46.5% variability. Eighteen items loaded on the factor goal-driven processing, and 10 items loaded on the factor data-driven processing. Five more items were excluded, as they failed to show substantial loading on either factor. The final 28 items demonstrated excellent reliability with a Cronbach's alpha of .937.

Study Findings

Study findings are presented in relation to the hypotheses and the research objectives with respect to the research question. Study results reinforce the literature review previously

presented and are presented herein. Study results suggest that the SAGAT instrument is indeed adaptable to the anesthesia domain and the SRNA subset of this population, and its validation is evident by qualitative and quantitative methods.

Hypotheses. All six hypotheses presented in this study were verified. Hypotheses are grouped and discussed by identification of the Level 1, 2, and 3 SA requirements (H1a, H2a, and H3a), and by validation of these Level 1, 2, and 3 SA requirements (H1b, H2b, and H3b), all of which are in relation to the induction of general anesthesia with associated oral endotracheal tube placement:

H1a: Level 1 SA requirements related to the induction of general anesthesia with oral endotracheal tube placement will be identified through the GDTA process.

H2a: Level 2 SA requirements related to the induction of general anesthesia with oral endotracheal tube placement will be identified through the GDTA process.

H3a: Level 3 SA requirements related to the induction of general anesthesia with oral endotracheal tube placement will be identified through the GDTA process.

Examination of the GDTA process by Delphi methods and the tabulations and content analysis of these results demonstrate the identification by expert panel of Level 1, Level 2, and Level 3 SA requirements for this event. Study results of Delphi methods are evaluated based on qualitative criteria of trustworthiness: credibility, defined as detailed identification of the researcher, his or her decisions, all actions, and any potential biases in the study; dependability, including strict adherence to the methodology with meticulous documentation of all records

including data collection, data analysis, tabulations, and content analysis methods; transferability, referring to the audit trail allowing another expert in the same field to recognize the results; and confirmability, meaning other researchers following the same methods and resultant data will have the same results (McPherson et al., 2018).

H1b: Level 1 SA requirements related to the induction of general anesthesia with oral endotracheal tube placement will be validated through the GDTA process.

H2b: Level 2 SA requirements related to the induction of general anesthesia with oral endotracheal tube placement will be validated through the GDTA process.

H3b: Level 3 SA requirements related to the induction of general anesthesia with oral endotracheal tube placement will be validated through the GDTA process.

Qualitative and quantitative results confirm the validation of the Level 1, 2, and 3 SA requirements through the GDTA process. Qualitative results from the Delphi process validate the Level 1, 2, and 3 SA requirements identified by the expert panel. According to Polit and Beck (2010), conservative recommendations for consensus are considered at 70% agreement, and yield validation of items (Polit & Beck, 2010). Results from the three rounds of expert panel surveys to develop the GDTA used this threshold as a minimum, with the majority of consensus coming at a higher threshold.

Quantitative methods applied to the study results revealed item content validity indices for 37 of the 39 items reaching the minimum I-CVI of 0.83 to establish excellent validity (Polit & Beck, 2006). The I-CVIs for the 37 items ranged from 0.83 to 1.00. The scale content validity

index (S-CVI) for the entire set of 39 items based on the average I- CVIs of all items (S-CVI/Ave) was calculated at 0.92, above the 0.90 threshold indicating excellent validity (Polit & Beck, 2006). Exploratory factor analysis methods resulted in confirmation of these hypotheses, with a Cronbach's alpha of 0.937 demonstrating excellent reliability (DeVellis, 2003).

Application to the Literature

The literature review provided both the underpinnings for the importance of this study as well as its theoretical basis, and methodology. Recruitment of experts and their completion of the requirements in both phases of the design demonstrate the importance of SA in the anesthesia domain, as well as anesthesia providers' (CRNAs) recognition of and determination to mitigate the effects of human factors on patient safety. This is supported by findings of Gaba (1998, 2000a, and 2000b), Gaba, Howard, and Small (1995), and Weinger and Gaba (2014), which all bear evidence to the link between human factors with an emphasis on SA, anesthesia related medical errors, and patient safety (Gaba, 1998, 2000a, 2000b; Gaba, Howard, & Small, 1995; Weinger & Gaba, 2014).

The successful use of GDTA in this study, based on its positive results in both qualitative and quantitative methods, underscores Endsley's "Theory of Situation Awareness" attestations. Goal-directed task analysis by the expert panel clearly identified the relevant sub-goals, underlying key decisions, and respective SA Level 1, 2, and 3 requirements according to Endsley's writings and recommendations (Endsley, 1995, 2000; Schulz et al., 2013; Wright et al., 2004). The EFA pattern matrix revealed 18 items loading to the factor data driven processing, and ten items loading to goal driven processing, affirming Endsley's assertion that SA is a feedback mechanism fueled by both data driven, or bottom-up, processing and goal driven, or top down, processing (Endsley, 2000, 2015). Despite the small sample size, the

adequacy of the data for EFA was indicated by a Cronbach's alpa of .937 of the 28 items remaining after extraction of further items that did not demonstrate substantial loading on either of the two factors, as evidenced by a scree test: data driven processing and goal driven processing (Pituch & Stevens, 2015).

Practical Implications

The results of this study have practical implications for nurse anesthesia educators, as well as educators of all anesthesia trainees and anesthesia providers (for example, continuing education instructors). Objectively quantifying SA in anesthesia trainees (or providers) bears useful assessment from both a formative and summative perspective. As SA can be enhanced with training and education, understanding the underlying SA level of students gives educators a foundation upon which teaching and learning exercises can be directed.

In formative assessment, evaluation is used to guide education and training. This assessment can take the form of an assignment grade within a course; however, its outcome is not used as the sole determinant for passing the entire course, thereby removing high stakes factors (and undo stress) on the outcome. Formative assessment is utilized as a method of teaching, with results appropriated to guide further instruction or to adjust methods of teaching. In summative assessment the measurement is used to determine progression of a student. This is considered high-stakes evaluation. The research team does not recommend this adapted and validated SAGAT for use in this regard, though it can be used in such a manner. The intended use of the SAGAT resulting from this study is aimed at the enhancement of SRNA SA and ultimately his or her clinical decision making abilities. The instrument can certainly also be used across anesthesia programs to assess SRNAs' level of SA, with results utilized to guide curriculum modifications or additions.

As presented in the introduction and literature review for this study, COA has recently modified its standards for nurse anesthesia educational programs offering a clinical doctorate to incorporate simulation into their curriculum. This SAGAT can be used to assess the impact of simulation incorporated into the curriculum for SRNAs in terms of their clinical decision making proficiency. Results of testing with this SAGAT can influence nurse anesthesia program accreditation policies and standards.

Limitations

The limitations of this study relate to study design and statistical analyses. Threats to internal validity and external validity represent its major limitations and are recognized within this section.

Internal validity. In qualitative methods, internal validity is discussed in terms of credibility. Credibility refers to establishing plausibility of the study results, meaning the findings represent content drawn from participants' original data and correctly portray their views (Korstjens & Moser, 2018; Shenton, 2004). The threats to internal validity or credibility in this study arise from the Delphi process, the transcription of collected responses, and the content analysis. This study attempted to account for these threats so that their influence is minimized.

Participation in an expert panel holds potential participant response bias, as individuals may have concerns regarding how others perceive his or her response. Using e-Delphi techniques, in which all data collection occurred through electronic data collection tools, gave the most assurance possible that the experts and their responses were blinded from each other. Though the researcher was blinded to responses through this same mechanism, the researcher's familiarity with the first seven-expert sample presented a risk to each member's anonymity. The larger 49 expert sample used during the second phase of the study was blinded to the researcher,

as only the participation pool from which the subjects were solicited was known, not the identity of those that actually participated.

The researcher as sole transcriber and coder for content analysis also presents threats to the credibility of this study. Maintaining meticulous records of actual subject responses and providing subjects with verbatim feedback of these transcribed responses minimizes the risk of researcher/coder bias inflicted on results. The subjects were asked within each survey to confirm responses they provided in preceding rounds. Additionally, the Delphi technique in and of itself imparts content validity to results. The criteria of credibility in qualitative research highlights that content validity of instrument items can be assured by use of those with item content specific knowledge in their creation (Shenton, 2004).

Fidelity of the content analysis was confirmed by the transcription and feedback of verbatim data provided to subjects. The subjects were given opportunity to provide free text comments with every survey question. Content analysis to generate items for the second phase of the study also was conducted by one researcher. These items were taken directly from the third round responses from phase one of this study, with content and terms used in the items taken verbatim from the Level 1, 2, and 3 SA identified by the expert panel. Meticulous records were kept of the processes taken to develop these items. Additionally, by using triangulation of different methods, different informants, and different settings, this limitation is further minimized. Disseminating the final list of items to the larger, geographically diverse expert sample provided further protection from researcher bias, and provided a greater degree of validity to the results.

Attrition also presented a limitation to this study's results. The Delphi process can be tedious and lengthy, making it difficult to retain subjects from one round to the next. In phase

one of this design, two of the seven experts failed to return the round two survey. Though they did participate in the final third round, the smaller second round sample (five rather than seven), limits the ability to confer consensus with certainty on results. Allowing the two subjects to provide expert opinion in the third round, as well as the ability to provide additional responses in free text format minimized the impact of this limitation.

Though neither subject offered rationale for nonparticipation, the time required for thoughtful response certainly is a plausible factor. Subjects were asked to self-report the time spent completing each round of the survey, and six of the seven subjects anonymously delivered this report. The mean time for the completion of round one was reported as 29 minutes, round two as 21 minutes, and round three as 27 minutes. Performing accurate and complete GDTA for a sub-goal requires much thought and deliberation: participants, despite the less than 30 minute reported mean completion time, could consider the process tedious. Providing the ability to administer surveys and collect data electronically minimized this limitation as subjects could answer at any time as long as internet services were available.

The quantitative methods applied to the data analysis create threats to internal validity, specifically the EFA. Exploratory factor analysis is characteristically used in large sample sizes, with recommendations of a minimum of five cases per factor to achieve adequacy of results. Such large sample sizes are not the norm of qualitative designs, particularly those related to behavioral or health care studies. Literature does, however, support the use of results for a sample based on the Bartlett's Test of Sphericity reaching significance (< .05) (Pituch & Stevens, 2015). For the results of this study, the Bartlett's Test of Sphericity was significant at .000. Cronbach's alpha demonstrates reliability and after initial extraction methods, these results received a Cronbach's alpha of .928 for the list of 33 items. With further extraction of five items

based on the pattern matrix, a Cronbach's alpha of .937 resulted.

Before EFA was applied, content validity indices for each item and the entire grouping of items were calculated and both indicated excellent validity. Item content validity indices measure the number of experts rating an item as relevant or highly relevant as a proportion of the total number of experts responding to an item (Polit et al., 2007). The study results revealed item content validity indices for 37 of 39 items reaching the minimum I-CVI of 0.83 to establish excellent content validity (Polit & Beck, 2006). The I-CVIs for the 37 items ranged rom 0.83 to 1.00. The scale content validity index (s-CVI) for the entire set of 39 items based on the average I-CVIs of all items (S- CVI/Ave) was calculated at 0.92, above the 0.90 threshold indicating excellent validity (Polit & Beck, 2006).

External validity. External validity speaks to the generalizability of the results, and in qualitative methods, is referred to as transferability. For qualitative measures, using rich descriptions enables the transferability to other contexts or settings with other respondents (Korstjens & Moser, 2018; Shenton, 2004). The qualitative method utilized in this study, namely the Delphi technique, is a lengthy process that may pose difficulty in reproduction. For qualitative designs, providing "thick" description supports transferability of results (Polit & Beck, 2010). This study included rich descriptions of the research setting, the study participants, all transcriptions related to data collection, and all processes undertaken including content analysis. Rich descriptions are also provided of SA as it relates to anesthesia providers, and specifically SRNAs. Rich descriptions allow the reader to judge the "proximal similarity" of the study with their own situations, fostering generalizability or transferability (Polit & Beck, 2010).

Replication of sampling is another way to minimize the risk of transferability (Polit & Beck, 2010). In qualitative designs, "… various purposive sampling strategies that involve

deliberate replication…" advance both, "… analytic generalization and transferability." (Polit & Beck, 2010, p. 1454) The sample of expert subjects providing opinion in the first phase of the design was replicated with a larger, geographically diverse sample in the third phase of the design. This part of the second phase added replication as well as maximal variation of the sample. Maximal variation applies as this sample group is culturally heterogeneous based on their geographic diversity (Palinkas et al., 2015). Though the characteristics of being a CRNA and nurse educator is the same as the first phase sample in this study, the second phase sample comes from various regions across the United States, adding an element of cultural diversity to the group. This sample strengthens the transferability of the study's results (Polit & Beck, 2010).

Conclusions and Recommendations for Future Research

This research study resulted in the adaptation and validation of the SAGAT, the only direct and objective measurement tool to quantify SA in operators of complex systems. Using the Delphi technique, a Goal Directed Task Analysis (GDTA) was performed by content experts, resulting in a final list of 28 items for use as a SA measurement tool in the SRNA subset during the simulation of the induction of general anesthesia with the placement of an oral endotracheal tube. The validity of these items was further confirmed by quantitative methods applied to the study results. Item Content validity indices resulted in values exceeding the threshold for excellent validity, as did scale content validity calculations applied to the list of items in their entirety (S-CVI/Ave = 0.92). Additionally, an EFA was performed demonstrating a Bartlett's Test of Sphericity significance of 0.00, and a Cronbach's alpha of 0.937.

This study generates further questions regarding the measurement of SA and the use of simulation in the curricula of anesthesia trainees. Simulated clinical events systematically expose trainees, such as SRNAs, to established and effective patient management protocols in a

controlled environment, and demonstrate an improvement in patient outcomes and clinical management. Future use of this SAGAT will provide a measure of the impact this training has in this population, as well as give indication to adjust didactic and training experiences in an effort to improve clinical decision-making and subsequent patient outcomes.

Though the effectiveness of simulation as an educational method for low occurrence, high risk events has been richly explored in recent literature, further examination of its use in routine anesthetic practices to support didactic educational concepts is warranted in the SRNA curricula. High fidelity human patient simulation is a costly undertaking and requires additional training for its proficient use in education. The ability to incorporate simulation throughout nurse anesthesia curricula would be of great benefit to both educators and nurse anesthesia graduate students. The SAGAT can provide a formative assessment tool for routine anesthesia concepts and practices, as well as critical, high-risk, low-occurrence events.

With the culmination of this study and its positive results, the most pressing issue is putting this adapted and validated SAGAT into use. Further testing of this tool in simulation with SRNAs is needed. It is the researchers intent to test the SAGAT simultaneously with an objective structured clinical examination tool (OSCE) measuring the SA of SRNAs during the simulation of the induction of general anesthesia with the associated placement of an oral endotracheal tube. Further exploration of the SAGAT by this means will add to the meaningfulness of this instrument and this study, as well as propel its adaptation and use with other scenarios. It is the researcher's intent to use the additional sub-goals and respective GDTA resulting from phase one of this study that were not within the scope of this project to formulate additional SAGAT instruments: one for preoperative patient assessment; and one for anesthesia machine check and equipment set-up.

The usefulness of SA measurement will grow exponentially as the need for adept health care providers burgeons from the increasing numbers of health care consumers over the next years. It is also evident that SA is increasingly important in the anesthesia domain with the advances and incorporation of more technology and automation of equipment. Anesthesia is a domain fraught with high workload demands, including stressful work environments, increased production pressure, work areas with many distractions, an increasing use of technology, and the constant need to prioritize work actions. This is compounded by anesthesia services being increasingly delivered in areas outside of the controlled operating room environment, such as: interventional radiology, and gastroenterology and urology procedure areas. Successful and safe anesthetic management requires high level SA to meet these (increasing) workload needs. Having a tool to directly and objectively measure SA in anesthesia trainees such as SRNAs is necessary and indeed quite timely.

References

American Association of Nurse Anesthetists. (n.d.). CRNA fact sheet. Retrieved from https://www.aana.com/membership/become-a-crna/crna-fact-sheet

Andreatta, P., Saxton, E., Thompson, M., & Annich, G. (2011). Simulation-based mock codes significantly correlate with improved pediatric patient cardiopulmonary arrest survival rates. *Pediatric Critical Care Medicine, 12*(1), 33-38.

Blandford, A., & Wong, B. W. (2004). Situation awareness in emergency medical dispatch. *International Journal of Human Computer Studies, 61*(4), 421-452.

Bolstad, C. A., & Cuevas, H. M. (2010). Integrating situation awareness assessment into test and evaluation. *ITEA Journal, 31*, 240.

Brown, E. N., Lydic, R., & Schiff, N. D. (2010). General anesthesia, sleep, and coma. *New England Journal of Medicine*, *363*(27), 2638-2650.

Chin, M., & Lagasse, R. S. (2017). Assessment of competence: developing trends and ethical considerations. *Current Opinion in Anaesthesiology*, *30*(2), 236-241.

Cooper, J. B., Newbower, R. S., Long, C. D., & McPeek, B. (1978). Preventable anesthesia mishaps: A study of human factors. *Anesthesiology,49*(6), 399-406.

Cooper, J. B., & Taqueti, V. (2004). A brief history of the development of mannequin simulators for clinical education and training. *BMJ Quality & Safety*, *13*(suppl 1), i11-i18.

Cooper, S., Porter, J., & Peach, L. (2013). Measuring situation awareness in emergency settings: A systematic review of tools and outcomes. *Open Access Emergency Medicine, 6*, 1-7. https://doi.org/10.2147/OAEM.S53679

Council on Accreditation of Nurse Anesthesia Educational Programs. (2018). Accreditation Standards, Policies and Procedures, and Guidelines. Retrieved from

https://www.coacrna.org/accreditation/Pages/Accreditation-Policies-Procedures- and-Standards.aspx

Creswell, J. W., Plano Clark, V. L., Gutmann, M. L., & Hanson, W. E. (2003). Advanced mixed methods research designs. *Handbook of Mixed Methods in Social and Behavioural Research, 209,* 240.

De Villiers, M. R., De Villiers, P. J., & Kent, A. P. (2005). The Delphi technique in health sciences education research. *Medical Teacher, 27*(7), 639-643.

DeVellis, R.F. (2003). Scale development: Theory and applications (applied social research methods). *Social Research, 900,* 14929.

Draycott, T., Sibanda, T., Owen, L., Akande, V., Winter, C., Reading, S., & Whitelaw, A. (2006). Does training in obstetric emergencies improve neonatal outcome? *BJOG: An International Journal of Obstetrics and Gynaecology, 113*(2), 177-182.

Endsley, M. R. (1988, May). Situation awareness global assessment technique (SAGAT). In *Aerospace and Electronics Conference, 1988. NAECON 1988., Proceedings of the IEEE 1988 National* (pp. 789-795). IEEE.

Endsley, M. R. (1995a). Measurement of situation awareness in dynamic-systems. *Human Factors, 37*(1), 65-84.

Endsley, M. R. (1995b). Toward a theory of situation awareness in dynamic systems. *Human Factors, 37*(1), 32-64.

Endsley, M. R. (2015). Situation awareness misconceptions and misunderstandings. *Journal of Cognitive Engineering and Decision Making, 9*(1), 4-32.

Endsley, M. R., & Bolstad, C. A. (1994). Individual differences in pilot situation awareness. *The*

International Journal of Aviation Psychology, 4(3), 241-264.

Endsley, M. R., & Kiris, E. O. (1995). The out-of-the-loop performance problem and level of control in automation. *Human Factors, 37*(2), 381-394.

Endsley, M. R., & Garland, D.J. (Eds). (2000). *Situation awareness analysis and measurement.* Mahwah, NJ: CRC Press.

Fetters, M. D., Curry, L. A., & Creswell, J. W. (2013). Achieving integration in mixed methods designs-principles and practices. *Health Services Research, 48*(6pt2), 2134-2156.

Flin, R., Patey, R., Glavin, R., & Maran, N. (2010). Anaesthetists' non-technical skills. *British Journal of Anaesthesia, 105*(1), 38-44.

Gaba, D. M. (1989). Human error in anesthetic mishaps. *International Anesthesiology Clinics, 27*(3), 137-147.

Gaba, D. M. (2000a). Anaesthesiology as a model for patient safety in health care. *BMJ: British Medical Journal, 320*(7237), 785.

Gaba, D. M. (2000b). Structural and organizational issues in patient safety: a comparison of health care to other high-hazard industries. *California Management Review, 43*(1), 83-102.

Gaba, D. M. (2004). The future vision of simulation in health care. *BMJ Quality & Safety, 13*(suppl 1), i2-i10.

Gaba, D. M., Howard, S. K., Flanagan, B., Smith, B. E., Fish, K. J., & Botney, R. (1998). Assessment of clinical performance during simulated crises using both technical and behavioral ratings. *Anesthesiology: The Journal of the American Society of Anesthesiologists, 89*(1), 8-18.

Gaba, D. M., Howard, S. K., & Small, S. D. (1995). Situation awareness in anesthesiology.

Human Factors, 37(1), 20-31

Gardner, A. K., Kosemund, M., & Martinez, J. (2017). Examining the feasibility and predictive validity of the SAGAT tool to assess situation awareness among medical trainees. *Simulation in Healthcare, 12*(1), 17-21.

Glavin, R., & Flin, R. (2012). The influence of psychology and humanfactors on education in anesthesiology. *Canadian Journal of Anesthesia/Journal Canadien D'anesthese*, *59*(2), 151-158.

Gosbee, J. (2002). Human factors engineering in patient safety. *Quality & Safety in Health Care*, *11*(4), 352-354.

Green, M., Tariq, R., & Green, P. (2016). Improving patient safety through simulation training in anesthesiology: where are we? *Anesthesiology Research and Practice, 2016.*

Hagemann, V., Herbstreit, F., Kehren, C., Chittamadathil, J., Wolfertz, S., Dirkmann, D., … Peters, J. (2017). Does teaching non-technical skills to medical students improve those skills and simulated patient outcome? *International Journal of Medical Education*, *8*, 101.

Haller, G., Laroche, T., & Clergue, F. (2011). Morbidity in anaesthesia: Today and tomorrow. *Best practice & research Clinical Anaesthesiology*, *25*(2), 123-132.

Harris, P. A., Taylor, R., Thielke, R., Payne, J., Gonzalez, N., & Conde, J. G. (2009). Research electronic data capture (REDCap)-A metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of Biomedical Informatics, 42*(2), 377-381.

Hasson, F., Keeney, S., & McKenna, H. (2000). Research guidelines for the Delphi survey technique. *Journal of Advanced Nursing, 32*(4), 1008-1015.

Hogan, M. P., Pace, D. E., Hapgood, J., & Boone, D. C. (2006). Use of human patient simulation and the situation awareness global assessment technique in practical trauma skills assessment. *Journal of Trauma and Acute Care Surgery, 61*(5), 1047-1052.

Holzman, R. S., Cooper, J. B., Gaba, D. M., Philip, J. H., Small, S. D., & Feinstem, D. (1995). Anesthesia crisis resource management: real-life simulation training in operating room crises. *Journal of Clinical Anesthesia, 7*(8), 675-687.

Jones, C. P. L., Fawker-Corbett, J., Groom, P., Morton, B., Lister, C., & Mercer, S. J. (2018). Human factors in preventing complications in anaesthesia: A systematic review. *Anaesthesia, 73*, 12-24.

Jordan, L. M., Quraishi, J. A., & Boust, R. (2015). The AANA Foundation malpractice closed claims study: a descriptive analysis. *AANA Journal, 83*(5), 318-323.

Kaber, D. B., & Endsley, M. R. (2004). The effects of level of automation and adaptive automation on human performance, situation awareness and workload in a dynamic control task. *Theoretical Issues in Ergonomics Science, 5*(2), 113-153.

Keeney, S., Hasson, F., & McKenna, H. P. (2001). A critical review of the Delphi technique as a research methodology for nursing. *International Journal of Nursing Studies, 38*(2), 195-200

Kisely, S., & Kendall, E. (2011). Critically appraising qualitative research: A guide for clinicians more familiar with quantitative techniques. *Australasian Psychiatry, 19*(4), 364-367.

Kohn, K. T. C. J., & Donaldson, M. S. (1999). *To err is human: Building a safer health care system*. Washington, DC: National Academy Press.

Korstjens, I., & Moser, A. (2018). Series: Practical guidance to qualitative research. Part

4: Trustworthiness and publishing. *European Journal of General Practice, 24*(1), 120-124.

Krage, R., & Erwteman, M. (2015). State-of-the-art usage of simulation in anesthesia: skills and teamwork. *Current Opinion in Anaesthesiology, 28*(6), 727-734.

Lampotang, S. (2008). Computer and web-enabled simulations for anesthesiology training and credentialing. *Journal of Critical Care, 23*(2), 173-178.

Lavoie, P., Cossette, S., & Pepin, J. (2016). Testing nursing students' clinical judgment in a patient deterioration simulation scenario: Development of a situation awareness instrument. *Nurse Education Today, 38*, 61-67.

Lee Chang, A., Dym, A. A., Venegas-Borsellino, C., Bangar, M., Kazzi, M., Lisenenkov, D., … Eisen, L. A. (2017). Comparison between simulation-based training and lecture-based education in teaching situation awareness: A randomized controlled study. *Annals of the American Thoracic Society*, *14*(4), 529-535.

Li, G., Warner, M., Lang, B. H., Huang, L., & Sun, L. S. (2009). Epidemiology of anesthesia-related mortality in the united states, 1999-2005. *Anesthesiology*: *The Journal of the American Society of Anesthesiologists, 110*(4), 759-765.

Makary, M. A., & Daniel, M. (2016). Medical error-the third leading cause of death in the US. *BMJ, 353,* i2139.

Manz, J. A., Hercinger, M., Todd, M., Hawkins, K. S., & Parsons, M. E. (2013). Improving consistency of assessment of student performance during simulated experiences. *Clinical Simulation in Nursing, 9*(7), e229-e233.

Matveevskii, A. S., & Gravenstein, N. (2008). Role of simulators, educational programs, and nontechnical skills in anesthesia resident selection, education, and competency

assessment. *Journal of Critical Care, 23*(2), 167-172.

McKenna, L., Missen, K., Cooper, S., Bogossian, F., Bucknall, T., & Cant, R. (2014). Situation awareness in undergraduate nursing students managing simulated patient deterioration. *Nurse Education Today, 34*(6), e27-e31.

McPherson, S., Reese, C., & Wendler, M. C. (2018). Methodology update: Delphi studies. *Nursing Research, 67*(5), 1. 404-410.

Merriam-Webster. (n.d.). Ergonomics. Retrieved from https://www.merriam-webster.com/dictionary/ergonomics?utm_campaign=sd&utm_medium=serp&utm_source=jsonld

Morgan, P. J., Kurrek, M. M., Bertram, S., LeBlanc, V., & Przybyszewski, T. (2011). Nontechnical skills assessment after simulation-based continuing medical education. *Simulation in Healthcare, 6*(5), 255-259.

Morgan, P., Tregunno, D., Brydges, R., Pittini, R., Tarshis, J., Kurrek, M., … Ryzynski, A. (2015). Using a situational awareness global assessment technique for interprofessional obstetrical team training with high fidelity simulation. *Journal of Interprofessional Care, 29*(1), 13-19.

Najjar, R. H., Docherty, A., & Miehl, N. (2016). Psychometric properties of an objective structured clinical assessment tool. *Clinical Simulation in Nursing, 12*(3), 88-95.

O'Brien, K. S., & O'Hare, D. (2007). Situational awareness ability and cognitive skills training in a complex real-world task. *Ergonomics, 50*(7), 1064-1091.

Orique, S. B., & Despins, L. (2018). Evaluating situation awareness: An Integrative Review. *Western journal of nursing research, 40*(3), 388-424.

Palinkas, L. A., Horwitz, S. M., Green, C. A., Wisdom, J. P., Duan, N., & Hoagwood, K. (2015).

Purposeful sampling for qualitative data collection and analysis in mixed method implementation research. *Administration and Policy in Mental Health and Mental Health Services Research, 42*(5), 533-544.

Petros, P. (2003). Non-linearity in clinical practice. *Journal of Evaluation in Clinical Practice, 9*(2), 171-178.

Pituch, K. A., & Stevens, J. P. (2015). *Applied multivariate statistics for the social sciences: Analyses with SAS and IBM's SPSS*. Routledge.

Polit, D. F., & Beck, C. T. (2006). The content validity index: Are you sure you know what's being reported? Critique and recommendations. *Research in Nursing and Health, 29*(5), 489-497.

Polit, D. F., & Beck, C. T. (2010). Generalization in quantitative and qualitative research: Myths and strategies. *International Journal of Nursing Studies, 47*(11), 1451-1458.

Polit, D. F., & Beck, C. T. (2012). *Nursing research: Generating and assessing evidence for nursing practice*. Lippincott Williams & Wilkins.

Polit, D. F., Beck, C. T., & Owen, S. V. (2007). Focus on research methods is the CVI an acceptable indicator of content validity. *Research in Nursing & Health, 30*, 459-467.

Reason, J. (1995). Understanding adverse events: Human factors. *BMJ Quality & Safety, 4*(2), 80-89.

Reason, J. (2000). Human error: models and management. *BMJ, 320*(7), 768-770.

Reason, J. (2005). Safety in the operating theatre - Part 2: Human error and organisational failure. *BMJ Quality & Safety*, *14*(1), 56-60.

Riley, B. (2015). Using simulation-based medical education to meet the competency requirements for the single accreditation system. *The Journal of the American*

Osteopathic Association, 115(8), 504-508.

Salmon, P. M., Stanton, N. A., Walker, G. H., Jenkins, D., Ladva, D., Rafferty, L., & Young, M. (2009). Measuring situation awareness in complex systems: Comparison of measures study. *International Journal of Industrial Ergonomics, 39*(3), 490-500.

Salmon, P., Stanton, N., Walker, G., & Green, D. (2006). Situation awareness measurement: A review of applicability for C4i environments. *Applied Ergonomics, 37*(2), 225-238.

Schulz, C. M., Endsley, M. R., Kochs, E. F., Gelb, A. W., & Wagner, K. J. (2013). Situation awareness in anesthesia: Concept and research. *Anesthesiology: The Journal of the*

American Society of Anesthesiologists, 118(3), 729-742.

Schulz, C. M., Krautheim, V., Hackemann, A., Kreuzer, M., Kochs, E. F., & Wagner, K. J. (2015). Situation awareness errors in anesthesia and critical care in 200 cases of a critical incident reporting system. *BMC Anesthesiology, 16*(1), 4.

Schulz, C. M., Schneider, E., Kohlbecher, S., Hapfelmeier, A., Heuser, F., Wagner, K. J., … Schneider, G. (2014). The influence of anaesthetists' experience on workload, performance and visual attention during simulated critical incidents. *Journal of Clinical Monitoring and Computing, 28*(5), 475-480.

Shelton, C. L., Kinston, R., Molyneux, A. J., & Ambrose, L. J. (2013). Real-time situation awareness assessment in critical illness management: Adapting the situation present assessment method to clinical simulation. *BMJ Quality & Safety, 22*(2), 163-167.

Shenton, A. K. (2004). Strategies for ensuring trustworthiness in qualitative research projects. *Education for Information, 22*(2), 63-75.

Sollid, S. J., Dieckman, P., Aase, K., Søreide, E., Ringsted, C., & Østergaard, D. (2016). Five

topics

health care simulation can address to improve patient safety: Results from a consensus process. *Journal of Patient Safety, 00*(00), 1-10. doi:10.1097/PTS.0000000000000254

Staender, S. E., & Mahajan, R. P. (2011). Anesthesia and patient safety: Have we reached our limits? *Current Opinion in Anaesthesiology, 24*(3), 349-353.

Stanton, N. A., Chambers, P. R., & Piggott, J. (2001). Situational awareness and safety. *Safety Science, 39*(3), 189-204.

Stanton, N. A., Salmon, P. M., Walker, G. H., Salas, E., & Hancock, P. A. (2017). State- of-science: situation awareness in individuals, teams and systems. *Ergonomics, 60*(4), 449-466.

Tolley, K., Marks-Maran, D., & Burke, L. (2010). The Snapshot tool: A new form of practice assessment. *British Journal of Nursing, 19*(14), 905-911.

Tong, A., Sainsbury, P., & Craig, J. (2007). Consolidated criteria for reporting qualitative research (COREQ): A 32-item checklist for interviews and focus groups. *International Journal for Quality in Health Care, 19*(6), 349-357.

Toronto, C. (2017). Considerations when conducting e-Delphi research: A case study. *Nurse Researcher (2014+), 25*(1), 10.

Van Den Bos, J., Rustagi, K., Gray, T., Halford, M., Ziemkiewicz, E., & Shreve, J. (2011). The $17.1 billion problem: The annual cost of measurable medical errors. *Health Affairs, 30*(4), 596-603.

Vukotich Jr, C. J. (2016). Challenges of T3 and T4 translational research. *Journal of Research Practice,* 12(2).

Wayne, D. B., Didwania, A., Feinglass, J., Fudala, M. J., Barsuk, J. H., & McGaghie, W.

C. (2008). Simulation-based education improves quality of care during cardiac arrest team responses at an academic teaching hospital: A case-control study. *Chest, 133*(1), 56-61.

Weinger, M. B., & Gaba, D. M. (2014). Human factors engineering in patient safety. *Anesthesiology: The Journal of the American Society of Anesthesiologists*, *120*(4), 801-806.

Weinger, M. B., & Slagle, J. (2002). Human factors research in anesthesia patient safety: Techniques to elucidate factors affecting clinical task performance and decision making. *Journal of the American Medical Informatics Association, 9*(Supplement_6), S58-S63.

Wetmore, D., Goldberg, A., Gandhi, N., Spivack, J., McCormick, P., & Demaria, S. (2015). An embedded checklist in the Anesthesia Information Management System improves pre-anaesthetic induction setup: A randomised controlled trial in a simulation setting. *BMJ Qual Saf, 25*(10), 739-746.

Wickens, C. D. (2008). Situation awareness: Review of Mica Endsley's 1995 articles on situation awareness theory and measurement. *Human Factors, 50*(3), 397-403.

Wiener, B., Chacko, S., Brown, T. R., Cron, S. G., & Cohen, M. (2009). Delphi survey of research priorities. *Journal of Nursing Management, 17*(5), 532-538.

Wright, M. C., Taekman, J. M., & Endsley, M. R. (2004). Objective measures of situation awareness in a simulated medical environment. *BMJ Qual & Saf, 13*(suppl 1), i65-i71.

Wright, S. M. (2015). Patient safety in anesthesia: Learning from the culture of high-reliability organizations. *Critical Care Nursing Clinics*, *27*(1), 1-16.

Wright, S. M., & Fallacaro, M. D. (2011). Predictors of situation awareness in student registered nurse anesthetists. *AANA Journal*, *79*(6), 485.

Yule, S., Parker, S. H., Wilkinson, J., McKinley, A., MacDonald, J., Neill, A., & McAdam, T. (2015). Coaching non-technical skills improves surgical residents' performance in a simulated operating room. *Journal of Surgical Education, 72*(6), 1124-1130.

Zausig, Y. A., Grube, C., Boeker-Blum, T., Busch, C. J., Bayer, Y., Sinner, B., … Graf, B. M. (2009). Inefficacy of simulator-based training on anaesthesiologists' non- technical skills. *Acta Anaesthesiologica Scandinavica, 53*(5), 611–619.

Vita

Deniz Dishman (nee Tombuloglu) was born in Boston, Massachusetts on December 21, 1966 and graduated from Massapequa High School, New York in June of 1984. She was awarded a Bachelor of Arts Degree in Social Science from Hofstra University, New York in May of 1987 and a Bachelor of Science in Nursing, Suma Cum Laude, from Molloy College, New York in August of 1994. She subsequently received a Master of Science in Nursing from Columbia University in the city of New York, New York in August of 2000. She went on to practice as a Certified Registered Nurse Anesthetist (CRNA) for 16 years, predominantly at Memorial Sloan Kettering Cancer Center, New York. She received a Doctor of Nurse Anesthesia Practice degree with a departmental award for leadership at Virginia Commonwealth University in 2014. Subsequently, she began a career in nurse anesthesia graduate education as an assistant professor in the Graduate Studies Department at the University of Texas Health Science Center, Cizik School of Nursing in Houston, Texas. During this time she published work titled, "Anesthetic Implications of Superior Vena Cava Syndrome in Interventional Radiology: A Case Report" in the AANA Journal (December, 2017). The AANA Foundation generously awarded her a research grant to fund the study presented in this dissertation manuscript.