Rapid Decision-making Processes in Anesthesiologists

Jingyi (Janet) Liu

Shilo Anders, PhD.

PSY 3981 Honors Thesis Spring 2022

Abstract

The purpose of this study is to understand differences in rapid decision-making processes between anesthesiologists with excellent, medium, and poor critical event management skills. A panel of expert clinicians and educators rated our recruited anesthesiologists' critical event management skills, categorizing them into high, medium, and low performers based on their ratings. We asked our participants to perform in a simulated scenario, and then we conducted cognitive interviews to extract specific information about their decision-making processes. Next, we conducted content analyses based on the interviews and then carried out an exploratory data analysis based on our decision-making model developed in our lab. We found that high performers were able to give clearer reasonings and evidence behind their differentials and actions, and they dealt with complex interpersonal situations and time pressure better than medium and low performers. Our findings will help educators develop more targeted training for anesthesiologists in the future.

Acknowledgment

This study is a part of an ongoing larger project in my group, IMPACTS coding team, led by my knowledgeable and caring mentor Dr. Shilo Anders who offered me tremendous guidance and support on this project. My teammates include Amanda Burden, Carrie Reale, Christen Sushereba, David Gaba, Janelle Faiman, Jason Slagle, Julie Diiulio, Laura Militello, and Megan Salwei. They all contributed to the larger project on a weekly basis. Among them, Julie is my coding partner, and we meet twice a week to reach a consensus on the content analyses for our cases. My thesis will not be possible without her help in coding our participants' transcripts. Thank you all!

Rapid Decision-making Processes in Anesthesiologists

Medicine is not a linear field. You bounce back and forth.

Anesthesiologists often need to make life-affecting decisions in a very limited amount of time. Errors in their decisions can result in Failure to Rescue (FTR) and other disastrous outcomes, including brain damage and massive hemorrhage (Schulz et al., 2017; Dutton et al., 2014). Moreover, anesthesiologists are expected to engage in critical event management: to rapidly recognize a medical event and to take effective actions accordingly (Gaba et al., 2010). While one might assume that anesthesiologists are well-trained in making rapid decisions during medical crises, prior simulation studies show that even for experienced medical professionals, 20-30% of them fail to adequately manage clinical crises (Henrich et al., 2009). To help anesthesiologists achieve higher performance in crisis management, it is crucial to understand how they perceive patients’ situations and act upon their knowledge under time pressure (Rudolph et al., 2009).

While there is extensive literature on decision-making in non-emergent situations, currently there is a lack of sophisticated understanding of the mechanism behind fast dynamic decision-making processes under time pressure (Rudolph et. al, 2009). Existing decision-making models that focus on uncertain, dynamic, and time-sensitive situations were developed primarily outside of healthcare in other industries including aviation and firefighting. (Orasanu, 1993; Klein et al., 1986). For example, researcher Orasanu proposed a two-part decision-making model based on her investigation of flight crews: an individual first assesses his or her situation and then chooses a course of action (Orasanu, 1993). Another more widely cited model is the Recognition-Primed Decision model, which is based on studies of firefighters (Klein et al., 1986). According to this model, firefighters first experience a situation, and if the situation is familiar to them, they will pick up relevant cues and set expectancies and goals, and finally take actions (Klein et al., 1986). However, both models are linear and fail to capture the dynamic situation in the real world, especially those crisis events anesthesiologists face on a regular basis.

Existing Models

In the healthcare field, prior studies have examined anesthesiologists' actions and problem-solving processes. For example, Rudolph and her colleagues came up with an action-oriented problem-solving model (Figure 1). This model includes three interacting stages of problem-solving: acting, interpreting, and cultivating diagnoses (Rudolph et al., 2009). When an unexpected situation occurs, such as abnormal breathing in a patient, an anesthesiologist will form an initial diagnosis by examining the signs and symptoms as well as the patient's medical history. These actions will help the anesthesiologist develop cues, which will either support or deny his or her current leading diagnosis. As the anesthesiologist picks up more cues during the crisis, he or she will update his or her interpretation of the situation, leading to the cultivation of a new diagnosis (Rudolph et al., 2009).

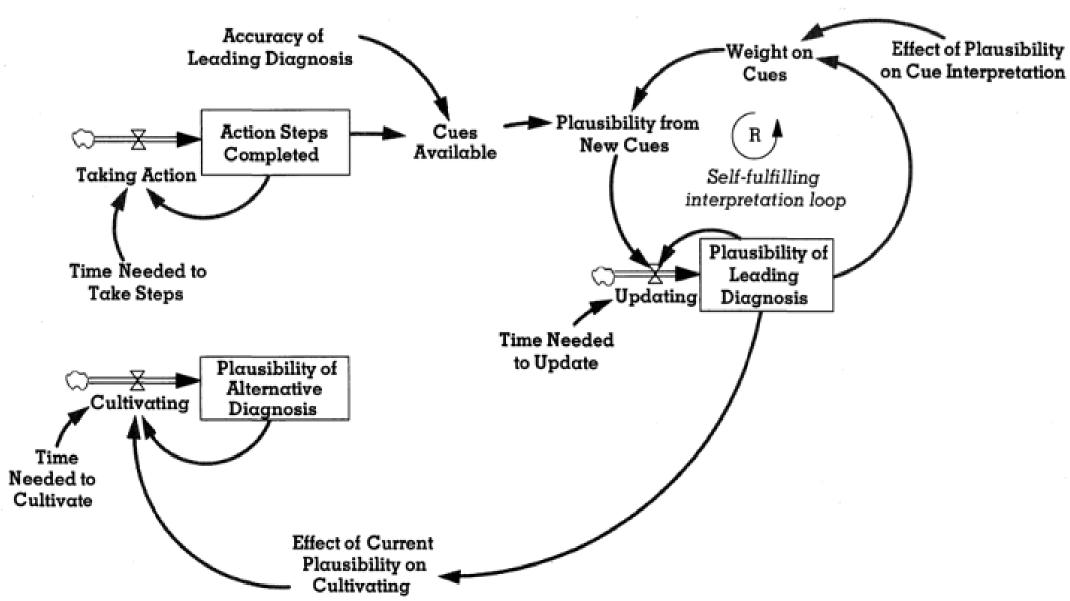


Figure 1. Interactions between acting, sensemaking, and cultivating alternatives (Rudolph et al., 200)

The benefit of Rudolph and his colleagues' model is that it is very dynamic because it captures the anesthesiologist's process of updating his or her diagnosis based on external cues. However, there are several limitations of this model. First, it is not generalizable to doctors of various training levels because this study only examined doctors with intermediate levels of training (Rudolph et al., 2009). Second, while this model seems beneficial to our understanding of how anesthesiologists manage critical events, it has not been effectively operationalized.

An anesthesiologist and collaborator of our lab, David Gaba, proposed a more complicated cognitive process model of anesthesiologists' problem solving (Gaba et al., 1995; Gaba & DeAnda, 1989). This model portrays five main levels of problem-solving: 1) resource-management, 2) supervisory, 3) sensory-motor, 4) procedural, and 5) abstract reasoning (Figure 2). One inspiring aspect of this model is that it emphasizes metacognition: the resource-management level and supervisory level. At the resource-management level, an anesthesiologist will distribute workload and communicate with other colleagues. At the supervisory level, he or she will allocate attention to different tasks and prioritize them (Gaba et al., 1995).

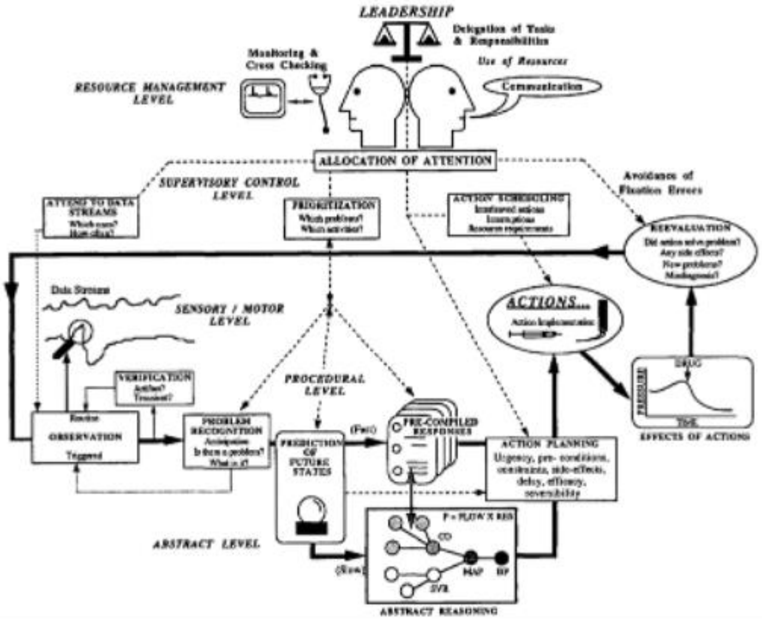


Figure 2. The cognitive process model of the anesthesiologist's problem-solving behavior (Gaba et al., 1995).

Knowing the theories behind anesthesiologists' problem-solving processes is beneficial for us to understand how problem-solving processes connect with our existing understanding of decision-making. To help connect sensemaking to decision-making and close the current knowledge gap on rapid decision-making processes in healthcare, my lab developed a circular multi-step decision-making process based on both Klein and his colleagues' Recognition-Decision Model and Gaba and his colleagues' cognitive process model of anesthesiologists' problem solving (Klein et al., 1986; Gaba et al., 1995).

Our Model and Project Design

The model my lab developed consists of five stages. At the first stage, an anesthesiologist will try to recognize the current situation which involves the patient, the physician, and the overall environment. An event cue that deviates from normal problem recognition will drive the anesthesiologist to conduct a situational assessment. During the assessment at the second stage, he or she will pick up relevant cues and form goals and expectancies, like Klein's Recognition-Primed Decision Model describes (Klein et al., 1986). Then in the third stage, the anesthesiologist will form a mental simulation with his or her leading diagnosis and evaluate the situation. At the fourth stage, he or she will take actions based on his or her diagnosis. Finally, he or she will continue monitoring the patient and update the current situation based on the patient's change in medical conditions.

During the whole process, anesthesiologists are not only subject to internal pressure to define and redefine their differentials, but also external factors like time limits and complex interpersonal dynamics. Understanding how anesthesiologists make and adjust their leading diagnoses and how they perform under both internal and external pressures is vital to improved outcomes for the benefit of both patients and physicians in the future.

Our study aimed to foster this understanding through simulated scenarios of close-to real world community hospital settings. We examined the possible differences between low-performing and high-performing anesthesiologists' decision-making processes, as operationalized by their overall gestalts. Specifically, we conducted an exploratory data analysis through which we looked at anesthesiologists' assessments, cues, and actions during their five-stage decision-making processes and how they utilize different mental tools such as analogical reasoning, rule in/rule out, and mental simulations. We used the simulated scenarios to initiate participants' rapid decision-making processes, and we used cognitive interviews to help both participants and us as researchers unpack their decision-making processes. These incident-based interviews have established validity, and they enable researchers to generate insights from participants' cognitive processes (Anders & Militello, 2019).

Analyzing superior outcomes through simulations can help reveal cognitive mechanisms of high-performers, and this knowledge will significantly contribute to medical training on how to acquire an expert-performance approach (Ericsson, 2015). In our study, we hypothesized that participants who receive higher performance ratings will have a clearer understanding of cues and make quicker actions. They will also be less affected by time pressure and complex interpersonal dynamics. Whether or data support our hypothesis or not, our in-depth understanding of decision-making processes will provide valuable findings that showcase how high-performing participants differ in their decisions from low-performing participants and the reasons behind their different actions and differential diagnoses.

Method

Participants

We recruited 26 anesthesiologists from training programs at our 5 study sites: Vanderbilt University Medical Center (VUMC) in Nashville, TN, University of Pennsylvania-Hershey in Hershey, PA, Cooper Medical School of Rowan University in Camden, NJ, WISER/University of Pittsburgh in Pittsburgh, PA, and the University of New Mexico in Albuquerque, NM. Among these anesthesiologists, 6 were junior resident anesthesiologists with 1-2 years of training and 24 were practicing anesthesiologists who are board-certified or equivalent (3-15 years of post-residency practice), including both community providers and attending anesthesiologists. Additionally, 5 panels of expert clinicians and educators who collaborated with us from our five study sites rated all participants’ performance on the five study sites respectively.

While we aimed to analyze all recruited anesthesiologists in a single simulated scenario for further analysis, we only managed to conduct complete analyses on 7 participants due to limited transcription and coding resources. Among these participants, 6 of them were community providers and 1 of them was a junior resident. Regarding their demographic information, 3 of them were females and 4 of them were males; 1 of them was Asian, 1 of them was African American, 4 of them were African Americans, and 1 of them declined to answer their race/ethnicity information.

All participants received market-level compensation. In addition, participants who were board-certified received 8 hours of Continuing Medical Education (CME) credits and were qualified for American Board of Anesthesiologists (ABA) credits.

Materials

Simulated Scenario. The study took place in a community hospital setting. Participants entered 4 simulated scenarios at their study sites with trained actors who played the roles of patients, nurses, and doctors. The four scenarios included 1) a 47-year-old female patient who has chest pain, 2) a 51-year-old male patient with hypotension, 3) a 55-year-old male patient who is short of breath, and 4) a previously healthy adult patient with altered mental status. All scenarios lasted for approximately 15 minutes. Scenario deliveries were standardized across our five study sites.

In this study, we were specifically interested in analyzing anesthesiologists' response to the first scenario: a 47-year-old female with type 2 diabetes, high blood pressure (hypertension), acid reflux (gastroesophageal reflux disease), and anxiety/depression presenting for total thyroidectomy for toxic goiter on methimazole. In the preoperative holding area during pre-anesthetic evaluation, she became progressively more anxious, tachypneic, and uncomfortable. This progressed to obvious chest pain (with electrocardiogram changes) with associated hemodynamic instability with a nadir at 15 minutes if untreated of tachycardia (134), hypotension (78/48), desaturation (82%), and wheezing.

We provided the actors with scripts and offer time-sensitive cues about physical signs. No guidance was provided to participants, and they learned about what happened to the patients by themselves at their study sites.

Study Personnel. The study personnel included our actors and our study team. There were two kinds of patients in our study based on the scenario. The first kind of patients were mannequins dubbed by an expert clinician in the observation room. The second kind of patients were standardized; they were highly trained actors who portray patients. Similarly, the roles of nurses and doctors were played by highly trained actors in all scenarios. The study team included 1) simulation researchers and expert anesthesiologists, 2) cognitive interviewers, and 3) simulation technicians.

Rating. To evaluate anesthesiologists' critical event management, we assessed both their technical and behavioral performance (Gaba et al., 1998). Our lab prepared scoring rubrics for independent expert clinicians and educators at the participants' sites to assess the two criteria. For each assessment, two raters assigned a participant's performance to one of 3 categories: poor, medium, and excellent (Figure 3). Then they chose one of 3 numbered levels that represent low, medium, or high performance within that category. Hence, a participant with poor performance would receive a score between 1 and 3; a participant with medium performance would receive a score between 4 and 6, and a participant with excellent performance would receive a score between 7 and 9.

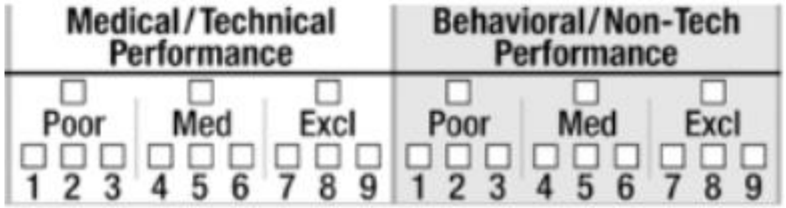
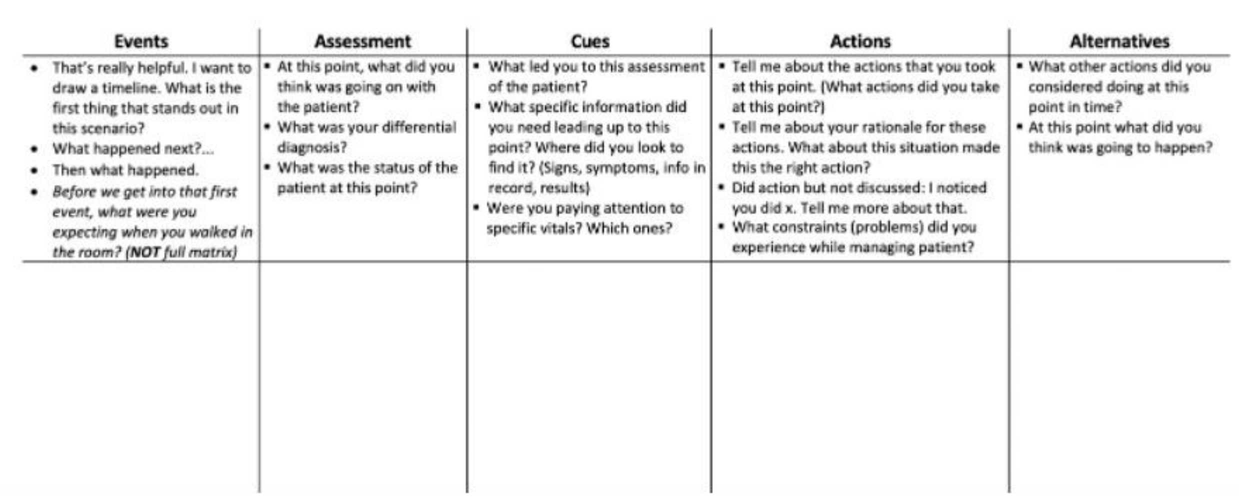


Figure 3. The rating rubric developed in my lab

Note. "Med" refers to "Medium" and "Excl" refers to "Excellent".

Cognitive Interview Guide. We used a Cognitive Interview Guide developed in our lab to standardize the interview procedure. A previously established Applied Cognitive Task Analysis (ACTA) provides an example of a simulation table that emphasizes 5 key areas during an interview: events, actions, assessment, critical cues, and potential errors (Militello & Hutton, 1998). Similarly, the Cognitive Interview Guide we used focused on questions in these 5 categories: events, assessment, cues, actions, and alternatives (Figure 4).



Events	Assessment	Cues	Actions	Alternatives
• That's really helpful. I want to draw a timeline. What is the first thing that stands out in this scenario? • What happened next?... • Then what happened. • *Before we get into that first event, what were you expecting when you walked in the room? (**NOT** full matrix)*	• At this point, what did you think was going on with the patient? • What was your differential diagnosis? • What was the status of the patient at this point?	• What led you to this assessment of the patient? • What specific information did you need leading up to this point? Where did you look to find it? (Signs, symptoms, info in record, results) • Were you paying attention to specific vitals? Which ones?	• Tell me about the actions that you took at this point. (What actions did you take at this point?) • Tell me about your rationale for these actions. What about this situation made this the right action? • Did action but not discussed: I noticed you did x. Tell me more about that. • What constraints (problems) did you experience while managing patient?	• What other actions did you considered doing at this point in time? • At this point what did you think was going to happen?

Figure 4. The layout and questions of the cognitive interview guide developed in my lab

Coding. All Cognitive Interview transcripts were coded using a codebook shared in our lab (Figure 7, Appendix A). During the coding process, each coder individually coded the transcript with DeDoose. Each transcript was coded by two coders, and the coders discussed their results during a subsequent meeting. Then, the two coders individually coded the Content Analysis sheet (Figure 5) and reached their consensus during a second meeting.

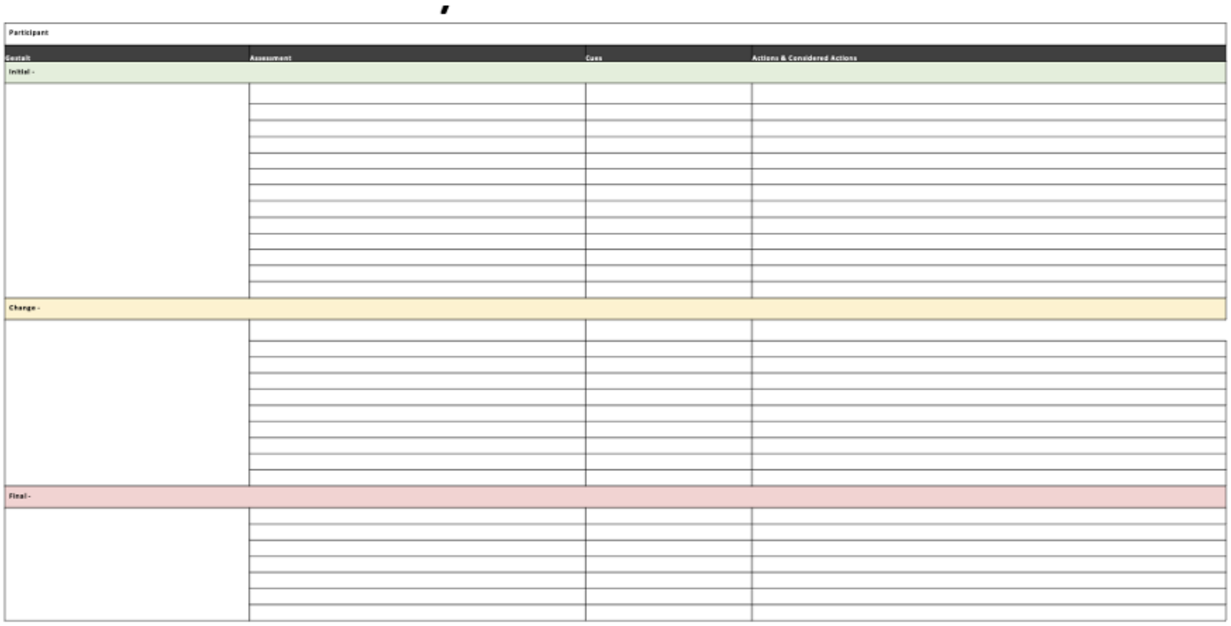


Figure 5. The Content Analysis sheet developed in my lab

Standardized Questions. By the end of each interview, we asked three standardized questions to all our participants:

1) Have you seen a case like this before?
2) On a scale of one to ten, with one being very easy and ten being very difficult, how challenging was this case?
3) On a scale of one to ten, with one being very unrealistic and ten being very realistic, how realistic was this case?

Procedure

Before the study day, we provided study overviews and consent forms for all participants. On the study day, participants arrived at one of our study sites at 8 am. They first participated in a 60-minute study session which included 15-minute didactics, 30-minute hands-on practice in the simulated environment, and then they performed a 15-minute scenario-equivalent "warm-up" scenario. After getting familiar with the simulation room setup through the "warm-up" scenario, they completed one of the four 12-20 minutes simulated scenarios followed by a 40-minute cognitive interview with a cognitive interviewer. The interviews were conducted through Zoom. Participants first wrote down a list of events that happened in their simulation on their own to create a brief timeline. Then, interviewers shared the cognitive interview guide on an excel worksheet with participants to create common ground via Zoom screen sharing. All interviews were audio and video-taped and transcribed for later analysis.

This process was repeated a total of 4 times with the 4 different scenarios. During the process, the raters recorded participants' performance in an observation room. Participants took a lunch break after 2 sessions. The scenario presentation order was randomized for each participant. At the end of each scenario, two independent raters in my lab assessed all participants' preliminary performance for that scenario. In this study, we only examined the results of the chosen scenario on the 47-year-old female patient with chest pain because it was the most complicated and participants have the most interaction with patients and colleagues. At the end of the study day, participants had a debrief with expert clinicians to discuss their performance and learnings.

Results

Performance Ratings

Our participants received a wide range of performance ratings from our expert clinicians and educators: 2 participants received excellent ratings (*mean = 6.5*), 2 participants received medium ratings (*mean = 5*), and 3 participants received poor ratings (*mean = 4*). The only junior resident anesthesiologist among our participants received a low rating while the other participants who were community anesthesiologists received a wide range of ratings from poor to excellent.

Difficulty and Realism Ratings

Figure 6 presents the pattern of means for high, medium, and low performing participants for difficulty and realism ratings for the scenario.

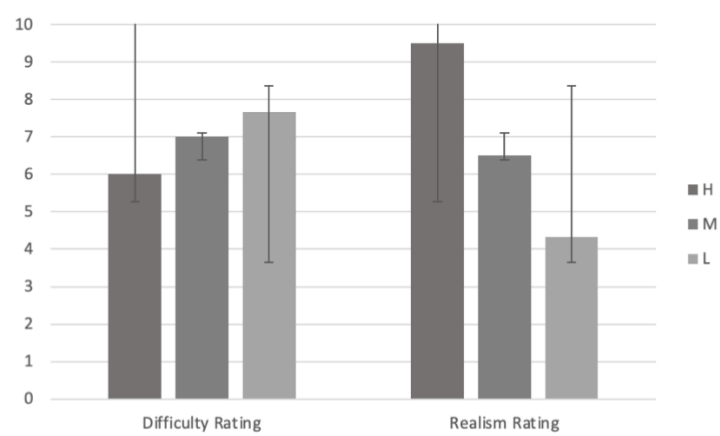


Figure 6. Means of difficulty and realism ratings for high, medium, and low performers

Similar Case Experience

A total of 3 participants said they had seen a case like this before: 1 medium performer and 2 low performers. Among the 4 participants who said they had not seen a case like this before, 2 high performers mentioned although they had experience with patients who were not optimized for surgeries, they did not experience extreme pushbacks from surgeons as presented in this scenario; 1 medium performer said they never saw a patient having mass problem in a case *"where there was pressure when the patient wasn't medically optimized";* 1 low performer said he/she did not have experience with patients with airway issues.

Stages of Our Mental Model

First Stage. During this problem-recognition stage, participants got familiar with the scenario by developing initial hypotheses and plans (Priming) and forming initial assessments (Problem-detection). They also utilized analogical reasoning while they were recognizing the current situation. For example, a participant who had treated goiter issues before reported that *"Whenever someone has a goiter in their history, I want to make sure that there were no airway problems that I might anticipate."*

I. Priming. All participants except a low performer reported that they were forming an initial idea of the situation before walking into the pre-operation room. High performers' expectations were more specific and problem-oriented as compared to low performers'. For example, a high performer reported *"I was expecting to see a patient ready for surgery, but the information was incomplete, that I was going to be finishing the preoperative evaluation and the exam. And I was expecting the patient to be a little anxious, and the surgeon was anxious to get started, impatient."* Another high performer mentioned the patient's vital signs: *"When I was looking at her pre-op heartrate and blood pressure, I saw a lot of, you know, things that, on most patients, would be concerning, her uncontrolled hypertension, her tachycardia."* Meanwhile, an example of a low performer's expectation statement was *"[We] got this patient coming in for a thyroidectomy, she's really nervous. We're running late. Surgery's ready to go. Nobody's really, nobody's talked to her. She hasn't had any workup at least we felt. Just walking into a minefield."*

II. Problem Detection. All participants formed their initial assessments based on given information about the patient. However, their certainty and evidence varied. Both high performers immediately listed out key concerns they had, including inadequate vital signs, pronounced hypertension, and a history of anxiety. Medium and low performers, however, failed to give sufficient evidence and were more unsure about the situation. For example, one of them reported *"I thought it was all related to her mass that she had, that maybe she was having a thyroid storm. But at one point, I think the surgeon said something about the patient being hypothyroid, which confused me because I thought that she would probably be hyperthyroid. And with her being tachycardic, it didn't make sense to me that she would be hypothyroid."*

Second Stage. At this stage, participants picked up cues through information seeking, and they formed goals and expectancies. Their cues, goals, and expectancies would give them opportunities to refine their initial differentials and rule out previous hypotheses.

I. Information Seeking. All participants sought information multiple times during their decision-making processes. The cues they specifically attended to include the patient's oxygen level, drug history, blood pressure, blood sugar, heart rate, and EKG patterns. This information also helped participants to set goals that would turn into actions at a later stage. For example, when a participant noticed the patient's heart muscle started to not pump as well, he/she wanted to increase the patient's blood pressure.

II. Rule-in/Rule-out and Refining the Differential. With specific cues, participants were able to rule out their previous hypotheses and refine their differentials. For example, a high-performing participant mentioned that he/she ordered a blood sugar because they noticed the patient was diabetic. The blood sugar test helped the participant to rule out the possibility that the patient was hypoglycemic. While medium and low performers used the same tactic, high performers knew how to prioritize their hypotheses: one high performer said he/she *"rule[d] out the dangerous things first"* such as panic attacks. This participant first ruled out this possibility because he/she learned from the patient that *"she had a panic attack before, and it was not similar to this."*

Third Stage. In the third stage, participants carried their leading diagnosis forward and formed mental simulations, picturing their potential course of action and consequences.

I. Mental Simulation. All participants utilized mental simulations during this stage to assess whether they should perform certain actions or make certain decisions. One of the most important decisions in this scenario was whether the participant should cancel the patient's surgery. A high performer's mental simulation for this decision went *"[I]f we do proceed, if we proceed safely, there's always a rush to sort of get things done. And most of the time, I'm very much like that, like let's not delay. But in this case, things didn't feel right. She wasn't in a nice regular, optimized state. There were some things going on that made me concerned about proceeding with surgery too fast."* A low performer's mental simulation went *"I was thinking like, oh, if this thyroid is that out of control, it's probably worth cancelling the case."*

Fourth Stage. At this stage, participants performed actions such as giving medicines and seeking out more information as they were trying to stabilize the patient, finalize their diagnosis, and decide on whether to cancel the case.

I. Actions. To stabilize the patient, participants performed various actions such as giving her Versed to control her anxiety, using Esmolol to manage her heart rate, and talking to her to calm her down.

II. Temporizing. Among our 7 participants, 2 high performers, 1 medium performer, and 1 low performer engaged in temporizing behaviors, namely buying time for the patient. While the high and medium performers had clear reasons for such behaviors, low performers did not. For example, a high performer stated that *"[S]he wasn't in the state of being ready for that type of surgery at that time. And so my actions were to buy some time to see how she turns out by also letting the surgeon know that she may not be ready… I thought about ways to sort of temporize, giving her some antianxiety medicine, finishing the rest of my questions, trying to get to the heart of, trying to understand why she was so anxious, and why she seemed like she was having difficulty breathing. You know, what medications had she taken that morning? Had she had a history of this in the past? And so the alternatives were, do we proceed, or we, do we hold off, and how I could temporize or manage, if something was temporary, how I could fix that so that she was optimized and calmed down and to a state where we could proceed."* Low performers, however, were only considering buying time to make sure the patient was *"stable enough for a general anesthetic"*.

Fifth Stage. At this stage, most participants made their final differentials and cancelled the cases. Among our 7 participants, only 3 failed to cancel the case: 1 medium performer and 2 low performers. While the medium performer wanted to cancel the case but did not have enough time to do so, the 2 low performers were too slow on their information-seeking steps and failed to come up with the conclusion that they should cancel the case.

I. Confirming/Disconfirming. All participants except a low performer actively confirmed/disconfirmed their hypotheses. High performers and medium performers were able to give detailed reasoning to their confirmation/disconfirmation, whereas low performers seemed to be at loss. While disconfirming demand ischemia, a high performer reasoned *"the fact that she, like, felt so suddenly ill all of a sudden, it made me think that probably wasn't just that her heart was beating a little fast, and her - you know, was causing a little of ST depression."* With this reasoning, this participant was able to narrow his/her differential down to acute coronary syndrome. Similarly, the other high performer was able to narrow his/her differential down to acute myocardial infarction (MI) with troponin information: *"once I saw the troponin come back positive, it kind of confirmed that she was having an MI or that it was elevated. And that she had sort of all the clinical symptoms of some agitation and sort of hard, she called it nausea, she called it dizziness, that kind of correlates well."* An example of a low performer's disconfirmation went: "*I was initially thinking she was having airway issues. And then hemodynamics got bad, and her mental status got bad. And then the doctor came in and said that she didn't think it was airway. So all these things, you know, sort of took, it basically confirmed what I was sort of leaning towards hat it was an airway."*

External Pressure

Interpersonal Dynamics. All participants mentioned challenges with interpersonal dynamics during their interviews. High performers exhibited more confidence while dealing with pushy surgeons, and they dealt with interpersonal situations better. For example, a high performer specifically talked about not being negatively impacted by a pushy surgeon: *"[J]ust not letting yourself get pushed into a bad situation by a pushy surgeon. As always, you know, anesthesia, it's the anesthesia fight that happens often sometimes in certain locations. So, you know, not, like, putting false pressure on yourself just to accommodate the O.R. schedule or somebody else's schedule if there's a legitimate emergency going on."* Low performers, on the other hand, were easily influenced by interpersonal distractions. One of them reported *"It was confusing because the doctor kept poking in and out. So there was like some distractions, and I think that's kind of part of the problem."*

Time Pressure. Among our 7 participants, 4 of them explicitly mentioned time pressure: 2 high performers and 2 low performers. A high performer's reported *"There was the time pressure. There was a patient who was focused on other things and, you know, needed to kind of be corralled and directed at first. There was a surgeon who was upset. So, there was a lot of, you know, outside distractors to the clinical scenario at hand, and she had a lot of things that, you know, a lot of her initial presenting symptoms could be explained by the story at hand."* Low performers, however, tended to attribute their errors to time pressure: *"If I had more time, I could've looked stuff up real quick and see what the initial management would've been. And that's what I was kind of doing at my phone. And said no, I don't have time to read that right now."* They also expressed frustration about being unsure about their diagnoses but still having to make rapid decisions under time pressure: *"[I am] always unsure of my diagnosis. Always. Always worried that I'm missing something. And then not knowing the next step to take immediately. I know where I can find the information. I know people I can call to take – to find the next step, to find out the next step. But do I have enough time to do that."*

Discussion

Our findings support our hypothesis that participants who receive higher performance ratings have a clearer understanding of the case and know how to prioritize and perform actions better based on the cues they picked up. Furthermore, while these high performers were equally affected by external factors like time pressure and complex interpersonal dynamics as the medium and low performers, they were able to acknowledge those pressures and deal with them with confidence. Low performers, however, tended to consider external factors as distractions, and they attributed their errors to these external factors.

Participants' Performance and Standardized Questions

By analyzing high, medium, and low performers' responses to our standardized questions, we can conclude that 1) previous experience with similar cases did not have any influence on participants' performance, 2) the better participants performed, the lower difficulty ratings they gave to the scenario, and 3) the better participants performed, the higher realism ratings they gave to the scenario.

The first conclusion might seem a bit surprising because we would usually think that with similar case experiences, one would perform better. However, it is important to note that it was the combination of the patients' medical situation and external factors that makes our chosen scenario unique and challenging, and external factors change case by case. Hence, it is understandable that even participants who had similar case experiences still failed to perform well in our simulated scenario.

One possible explanation for the second conclusion is that high performers had a better sense of what the patients' diagnoses could be and what they should do to help the patient. Many low performers, on the other hand, were confused during the scenario and failed to make their decision to cancel the case in time. Hence, it is reasonable that the low performers gave higher difficulty ratings than the high performers.

One possible explanation for the third conclusion is that high performers thought time pressure, difficult interpersonal dynamics, and initially confusing cases were common for them during their everyday practice. Although they might not have encountered similar cases, they gained much insight into how to calm down or optimize anxious patients, how to deal with external factors, and when to cancel the case.

Five-stage Decision-making Model

While our five-stage model helped us understand how participants came up with their differentials and why many of them decided to cancel the cases, it is important to note that this process is non-linear. For example, information seeking could take place throughout the entire decision-making process and the cues anesthesiologists picked up from information seeking could be used to make assessments such as refining differentials and to drive actions like ordering EKG and calling consultants.

Furthermore, our findings on the anesthesiologists' mental processes at each stage showed a consistent discrepancy between high performers and low performers in 1) providing reasonings and evidence for their assessments and actions, 2) using specific cues, and 3) forming a clear picture of what is the best for the patient with updated information and differentials.

Strengths and Limitations

Overall, our study has a relatively high external validity because we selected participants from our five different study sites, covering a wide range of geographic locations. Furthermore, the fact that we had more than one independent expert clinicians and educators rate participants' performance ensures a high level of internal validity and inter-rater reliability. However, there are several limitations of our study.

One of the most salient limitations of our study is the small sample size. Because of this, we did not run statistical tests on participants' answers to their standardized questions. The small sample size also limited our ability to perform a random selection of our participants. If we had a larger number of transcribed interview data of all our 26 recruited participants, we would first split all our participants into high-performing, medium-performing, and low-performing bins based on their performance ratings. Next, we would randomly select 5 participants from each bin and analyze their decision-making processes in the same scenario. Hence, in an ideal situation, we would have a total of 15 anesthesiologists of different experience levels.

Another limitation of our study is that almost all our participants were community anesthesiologists with 3-15 years of post-residency practicing experience. Only 1 of our participants was a junior resident who had 1-2 years of training experience. Hence, our participants in our sample study had a skewed distribution towards more practicing experience. In an ideal case, we would have an even number of junior residents and practicing anesthesiologists so that we could compare how anesthesiologists with various experience levels perform differently. Moreover, for practicing anesthesiologists, it would be helpful for us to have both community anesthesiologists and attending anesthesiologists who work at academic medical centers. We think including a more diverse sample of practicing anesthesiologists would be helpful because attending anesthesiologists do both teaching and practicing. Hence, including them in our study would give us a clearer picture of how anesthesiologists who engage in frequent teaching might make their decisions and explain why they make their decisions during our interviews differently.

Finally, it is worth noting that while we only examined the results of one simulated scenario, participants were asked to participate in four scenarios in one day due to the nature of the bigger study this study is a part of. This intense process might not match what the participants experience in their everyday work, leading participants to stress and fatigue that could potentially result in their lower performance in our chosen scenario.

Implications and Future Directions

Our study has the potential to provide insights into anesthesiologists' decision-making processes for educators to develop more targeted training for simulated learning scenarios.

Moreover, by testing out the practicability of the multi-step decision-making model my lab developed, we were able to understand more about the rapid decision-making process. Beyond this project, future researchers can examine other facets of rapid decision-making, including the number of errors anesthesiologists make and factors that influence how anesthesiologists work with physicians in a team-based setting while rescuing patients. With this newfound knowledge of both internal and external factors that influence anesthesiologists' decision-making processes, medical centers can tailor their training programs by highlighting common mistakes made by anesthesiologists and bringing in more training on how to deal with complicated interpersonal dynamics and how to perform well under time pressure.

References

Anders, S., & Militello, L. (2019). Incident-Based Methods for Studying Expertise. In *The Oxford Handbook of Expertise* (1st ed.). Oxford University Press. doi: 10.1093/oxfordhb/9780198795872.013.19

Cohen, I. (2008). Improving Time-Critical Decision Making in Life-Threatening Situations: Observations and Insights. *Decision Analysis*, *5*, 100–110. doi:10.1287/deca.1080.0111

Dutton, R., Lee, L., Stephens, L., Posner, K., Davies, J., & Domino, K. (2014). Massive Hemorrhage: A Report from the Anesthesia Closed Claims Project. *Anesthesiology*, *121*, 450–458. doi: 10.1097/ALN.0000000000000369

Ericsson, K. A. (2015). Acquisition and Maintenance of Medical Expertise: A Perspective From the Expert-Performance Approach With Deliberate Practice. *Academic Medicine*, 90(11), 1471–1486. doi: 10.1097/ACM.0000000000000939.

Gaba, D., & DeAnda, A. (1989). The response of anesthesia trainees to simulated critical incidents. *Anesthesia and Analgesia, 68,* 444–451. doi: 10.1213/00000539-198904000-00004

Gaba, D., Howard, S., & Small, S. (1995). Situation Awareness in Anesthesiology. *Human Factors, 37*, 20–31. doi: 1518/001872095779049435

Gaba, D., Howard, S., Flanagan, B., Smith, B., Fish, K., & Botney, R. (1998). Assessment of clinical performance during simulated crises using Both Technical and behavioral ratings. *Anesthesiology, 89*, 8–18. doi: 10.1097/00000542-199807000-00005

Gaba, D. (2010). Crisis resource management and teamwork training in anesthesia. *British Journal of Anesthesia*, *105,* 3–6. doi: 10.1093/bja/aeq124

Henrichs, B., Avidan, M., Murray, D., Boulet, J., Kras, J., Krause, B., Snider, R., & Evers, A.

(2009). Performance of Certified Registered Nurse Anesthetists and Anesthesiologists in a Simulation-Based Skills Assessment. *Anesthesia and Analgesia*, *108*, 255–262. doi: 10.1213/ane.0b013e31818e3d58

Klein, G., Calderwood, R., & Clinton-Cirocco, A. (1986). Rapid Decision Making on the Fire Ground. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 30*, 576 - 580.

Militello, L. G., & Hutton, R. J. (1998). Applied cognitive task analysis (ACTA): a practitioner's toolkit for understanding cognitive task demands. *Ergonomics, 41*, 1618–1641. doi: 10.1080/001401398186108

Orasanu, J., Dismukes, R. K., & Fischer, U. (1993). Decision errors in the cockpit. Proceedings of the Human Factors and Ergonomics Society 37th Annual Meeting (pp. 363–367). Santa Monica, CA: Human Factors and Ergonomics Society.

Rudolph, J. W., Morrison, J. B., & Carroll. J. S. (2009). The Dynamics of Action-Oriented Problem Solving: Linking Interpretation and Choice. *The Academy of Management Review, 34*, 733–756. doi: 10.5465/AMR.2009.44886170

Schulz, C., Burden, A., Posner, K., Mincer, S., Steadman, R., Wagner, K., & Domino, K. (2017). Frequency and Type of Situational Awareness Errors Contributing to Death and Brain Damage: A Closed Claims Analysis. *Anesthesiology, 127*, 326–337. doi: 10.1097/ALN.0000000000001661

AppendixA

Figure 7. Coding Terms Explanation

Priming	Participant describes planning and/or hypotheses before seeing the patient
Problem detection and framing	Participant articulates a problem, an anomaly, something unexpected that requries scrutiny. Participant articulates a frame for the problem. E.g., this code most likely applied early on where participant first identifies/describes a
Assessing/Recognizing	
Analogical reasoning	Participant indicates they have seen something analogous before and made judgments based on the similar case. E.g., this is analogous comparison, not instinctive recognition of a familiar situation.
Goals	Participant describes goals at a point in time, or participant describes global goals "I wanted to bring BP down."
Mechanistic thinking	Physiologic thinking as doing routine things tagging back to the physiology. E.g., describing the physiology behind an assessment or action.
Pattern matching	Participant describes a global assessment that integrates several sets of data points. E.g., describes a pattern + an assessment (A+B+C = D)
Recognition	Participant describes recognizing the situation or condition as familiar -- knew exactly what to do. E.g., this code is typically used when participant explicitly states that they have seen situation before and knew what to do (instinctive). This code will likely be used sparingly.
Refining the differential	Participant describes multiple diagnostic hypotheses and how he/she is ordering and re-ordering the hypotheses in terms of likelihood. E.g., Not necessarily tied to a specific piece of information.
Sensemaking or storybuilding	Participant describes specific cues and how they fit together. Could include mutliple possible interpretations. E.g., Participant describes what is going on (trying to rationalize) but not necessarily tying it all together (like in pattern matching). This code may be used when another code is not a better fit.
Acting	
Anticipating	Participant describes an action taken or considered in anticiption of patient's future condition
Execution evaluation	Checking for effective implementation of an action
Information seeking	Provider takes an action with the intent of learning more about the patient's condition
Innovation	Participant reports that protocol is not sufficient; they innovate
Rule in/Rule out	Participant describes strategy for ruling in/ruling out hypotheses. E.g., Rationale behind an action (ordered an EKG to make sure she wasn't having a heart attack)
Rule-based behavior	If this, then that; tied to a rule, protocol based; includes use of procedures in decision aids.
Serial vs. concurrent implementation	Participant reports that they have to prioritize, choose which thing(s) to do first. E.g., I need to do x, y, z, but I can't do them all at once. Or, participant describes a series of actions to achieve a certain goal.
Temporizing	Forestalling, buying time, watchful waiting, hedgin; Timing -- how soon to move on? Includes stop-gap measures.
Critiquing	
Confirming/disconfirming	Description of how test results or change in patient condition provided confirming or disconfirming evidence to a potential diagnosis. E.g., tied to a piece of evidence (e.g., test result)
Expectancies	Participant describes expectancies, especially if violated/confirmed. E.g., Expected to see change after administering dantroline; expected pt to be asleep, have difficulty breathing if benzo OD; Expected attending to know what to do.
Mental simulation	Participant describes playing the situation forward; imagining potential courses of action and/or potential outcomes. This includes weighing pros and cons of different actions/frames. Specific to the patient
Correcting	
Frame shift	Updating a frame; choosing a different course of action or plan. E.g., New information, all of a sudden thinking about it differently
Standardized questions	
Case experience	
Difficulty rating	
Realism rating	
Hindsight/Self-identified errors	
Errors of omisison	Code for things we expected to see, but didn't
Fail to detect an error	
Faulty re-evaluation	
Faulty understanding of the differential; insufficient medical knowledge	
Fixation	Mindset that "it has to be this"
Misinterpretation of data/information	
No recognition of the problem	
Select inappropriate action	"I did the wrong thing. Gave her too much"
Environmental factors	
Interpersonal dynamics	Participant mentions any interpersonal dynamics that influenced decision-making and/or actions
Resource constraints	Participant describes not having enough/proper resources (not enough help, lacking the right equipment, drugs, etc.)
Time pressure and stress	Participant describes time pressure and/or stress
Uncertainty	Participant describes sources of uncertainty
Interesting behaviors and comments	
For discussion	Put things here you want to discuss with the group
Good Quotes	
Next Steps	What the participant was planning to do next.
Resignation	Participant describes giving up.
Simulation artifact	Describes a simulation artifact