The Journal of Education in Perioperative Medicine

Original Research

Constructing a Validity Argument and Exploring Implications for the American Board of Anesthesiology's Basic Examination

Riley S. Carpenter Lide, MD, MHPE *Rachel Moquin, EdD* *Erin Green, MD, MHPE*

Introduction

Physician board examinations require robust validity arguments to justify their use in certifying that diplomates possess the necessary knowledge and skills for safe patient care.[1] Although most medical specialties conduct board exams post-residency, anesthesiology is among the few that administers a high-stakes exam during residency. The American Board of Anesthesiology (ABA) introduced the Basic Exam in 2014 to promote early study habits and foundational knowledge.[2] ABA policy requires serial extensions of training after the third failed attempt at the exam but leaves the number of attempts allowed to the individual training programs. The exam is positioned after the first year of clinical anesthesia (CA-1) training and successful completion is a graduation requirement. Contrastingly, other exams in the specialty's board certification series occur post-residency, making the Basic Exam unique in its potential to prevent graduation.[3] The first-time pass rate has hovered around 90%, suggesting that between 1500 and 2200 trainees have experienced at least 1 failure in the past decade.

The process of assessment validation involves gathering relevant evidence to justify decisions made on behalf of exam results. However, published validity evidence for the Basic Exam primarily addresses its intended purpose, relying on performance metrics from other standardized tests to illustrate enhanced knowledge acquisition.[2,4] Editorial responses highlight the necessity for a more comprehensive validity argument, noting the studies' small effect sizes and questioning the exam's broader impacts on burnout, resident selection, and the balance between clinical and academic responsibilities.[5-7] A decade after implementation, a thorough validity argument remains undeveloped because validation efforts have concentrated on the exam itself rather than on the decisions made based on its behalf.

Rigorous validation is strengthened by explicitly articulating how assessment results will be used, applying formal validation frameworks, and collecting evidence regarding the implications of the assessment.[8] One such framework is Kane's argument-based approach,[9] which has been used in medical education to structure validity arguments for the Mini-Clinical Evaluation Exercise (Mini-CEX)[10] and Objective Structured Assessment of Technical Skills (OSATS).[11] Kane's framework begins with a statement of the intended use of assessment results, followed by an outline of the 4 inferential links necessary to transition from test administration to making a decision about a learner. These links include *scoring*, which translates test performance into scores; *generalization*, which pertains to the reliability of a learner's performance in exam contexts; *extrapolation*, which correlates exam performance with real-world performance; and *implications*, which addresses the consequences of decisions made based on assessment results. Each link functions as a testable hypothesis in developing a comprehensive validity argument, enabling a judgment of whether adequate validity evidence supports the stated use. Kane places particular importance on implications and highlights that a decision-making procedure may be abandoned if it fails to achieve its goals or does so at too high a cost.[9]

This study evaluates how Basic Exam results are used within individual training programs, representing a critical initial step in applying a validation framework to structure a comprehensive validity argument. Concurrently, we explore implications of the exam to substantiate decisions made on its behalf. Kane's argument-based approach to validity and best practices in assessment, as outlined by the Standards for Educational and Psychological Testing (The Standards), provides a conceptual framework for interpreting the results of this study.

Research Questions

Adopting a constructivist worldview, this study uses a phenomenological approach to thematic analysis to explore the uses and implications of the Basic Exam through the lens of anesthesiology program directors (PDs):

How are Basic Exam results used at individual training programs?

continued on next page

continued from previous page

What are the implications of programmatic usage of the Basic Exam?

Materials and Methods

Reflexivity

Beyond addressing personal reflexivity, this study also addresses interpersonal reflexivity concerning power dynamics between researchers and participants, and methodological reflexivity related to the selection of the methodological approach and supporting quotations. This article is based on R.S.C.L.'s thesis for a Master's degree in Health Professions Education (MHPE) from New York University Grossman School of Medicine and Maastricht University. This study was designed with a sense that the validity argument for high-stakes usage of Basic Exam results, as with many high-stakes exams in medicine, was underdeveloped. R.S.C.L. took and passed the Basic Exam during its infancy in 2015 and this perspective informed the initial interview guide, providing background and context for in-depth exploration of the exam's implications (Supplemental Online Material, Appendices A and B). At the time of this study, R.S.C.L. was not yet a PD, which positioned participants in a senior role in the interview context. We believe that the participants' authority in this dynamic positively contributed to the depth and quality of the insights gathered.

To ensure a well-rounded research team, we included RM, who holds a doctorate in education and serves as Associate Vice-Chair in Anesthesiology, and EG, an emergency medicine physician who serves as Assistant Dean in Undergraduate Medical Education. All researchers identify as cisgender white females. At the time of writing, both R.M. and R.S.C.L. worked in programs with a 3-attempt limit on the Basic Exam. Regarding methodological reflexivity, the interview guide permitted participants to suggest revisions to the board examination series. We used this information to approximate participant support for the Basic Exam (Supplemental Online Material, Appendix C) to ensure that perspectives from both ends of the spectrum were considered in our thematic analysis.

Ethics

This study, approved by the Institutional Review Board at the University of Arkansas for Medical Sciences (#274932) and deemed exempt for minimal risk, received a waiver of written consent. Participants were informed about the study's purpose, data collection, and handling procedures, and could opt out of any interview questions, with responses considered ongoing consent. Identifiable information was excluded before data sharing or analysis.

Participants and Setting

The participants were current or former leaders of residency programs at anesthesiology residency programs accredited by the Accreditation Council for Graduate Medical Education (ACGME).

Sampling and Recruitment

This study used a combination of convenience and purposive sampling to investigate the phenomenon of interest. In the initial phase of convenience sampling, we contacted all current anesthesiology PDs at 166 ACGME-accredited anesthesiology residencies via emails obtained from the ACGME program search page.[12] Enrollment criteria included experience with at least 1 resident Basic Exam failure during their tenure as PD. After the initial round, we recorded details such as the program's geographic regions, sizes, PD tenures, and affiliations with professional societies.

Subsequent rounds of recruitment used a purposive sampling approach to include programs with characteristics not represented in the initial phase. We selected 20 participants, as this sample size aligns with methodological norms in qualitative research, which typically requires 9 to 17 interviews to achieve data saturation.[13]

Data Collection

Semistructured interviews were conducted through a web-based platform, using an interview guide designed to address the study's research questions and iteratively developed to explore emerging concepts. The principal investigator (PI) conducted each interview to ensure consistency.[14,15] Audio recordings were transcribed using a third-party service. Each interview lasted approximately 50 minutes, resulting in a final dataset of 20 transcripts, each capturing the entirety of the interview.

Analysis

The PI selected 3 transcripts that best represented the dataset, which were coded independently by each of the 3 team members. Over a series of 3 meetings, the team reviewed these transcripts to create a common codebook, initially containing 48 codes (Supplemental Online Material, Appendix D). Using Atlas Software, the team refined the codebook by re-coding the first 3 transcripts, then applied it to transcripts 4 to 10, adding 5 new codes. Only 3 new codes were added when coding transcripts 11 to 20, indicating coding saturation. Discrepancies were resolved through group meetings, resulting in a final codebook of 56 codes, with 10 intentionally broad codes sub-coded by pairs of researchers. A total of 1941 codes were applied across 20 interviews. To enhance data trustworthiness, preliminary thematic analysis results were shared with participants for member checking, with 4 of 20 responding affirmatively. In reporting data, quotes were edited for clarity, with omissions indicated by 3 dots (...) and added words in square brackets.

Demographics

The dataset represented 20 of 166 ACGME-accredited anesthesiology residency programs.[12] Participants included 18 current PDs, 1 vice chair of education who served as a PD during the implementation of the exam, and 1 associate PD, each representing distinct programs. Among the participants, 7 (35%) were women and 14 (70%) were men, reflecting national PD demographics. Racial and ethnic backgrounds were not disclosed. Participants represented 13 large- (> 50 positions) and 7 medium-sized programs, spanning 7 of 9 US Census Bureau–defined geographical regions.[16] Three participants (15%) had taken the exam in a standard-setting study, and 2 (10%) had taken it as residents. Three participants possessed formal education training and 4 served as PDs when the exam was introduced. Six participants had experience as question writers or examiners in the ABA-staged exam series. Participants held leadership positions with the ACGME, the American

continued on next page

continued from previous page

Society of Anesthesiologists, the Society for Education in Anesthesia, the Association of Anesthesiology Core Program Directors, and the ABA.

Results

A 56-item codebook was inductively developed and used to analyze 20 transcripts, resulting in 1941 coded segments of text. Thematic analysis resulted in 7 themes and subthemes answering our 2 research questions. Table 1 provides a brief description of each, along with illustrative quotes. Theme 1 highlights programmatic policies with an emphasis on dismissal (1b) and informs the interpretive-use argument. Theme 2 demonstrates differences in participants' perceived purposes of the exam: to "weed out" residents who are unlikely to achieve board certification (2a), to provide a data point to support remediation (2b), and to distinguish physicians from other anesthesia providers (2c). Implications begin with theme 3, which captures the Basic's impact on recruitment (3a), operations (3b), and curricula (3c). Theme 4 highlights that residents are studying for the exam, with an emphasis on targeted test preparation (4a). Theme 5 alludes to resident implications, including stress (5a) and distraction from clinical care (5b). Theme 6 describes the implications of failure. Theme 7 synthesizes equity concerns.

Discussion

The findings of this study illustrate that the Basic Exam performance is not only used in graduation, but also in dismissal decisions. In theme 1, we identify 11 programmatic policies allowing exam performance to serve as the sole criterion for dismissing residents. Among 20 programs, we examined 7 instances of resident dismissal, 3 of which cited Basic Exam performance as the only deficiency. The language used in these dismissals often implies that the resident voluntarily resigned from the program, a nuance that complicates a comprehensive examination of this phenomenon and warrants attention in future validation efforts.

Given the stakes involved, we apply Kane's framework to evaluate the validity argument concerning dismissal. To assess this argument, we must first outline the interpretive statement supporting dismissal as a use of exam results. This statement posits that the Basic Exam's format, content outline, and cut score can reliably predict knowledge deficits among CA-1 residents in the testing environment (generalization), which translates to performance deficits in the clinical environment (extrapolation). Such deficits could jeopardize patient safety after residency, thereby justifying a resident's dismissal from anesthesiology (implications). Table 2 presents a comprehensive outline using Kane's framework and provides hypothetical examples for validation evidence at each inference. Table 3 maps the available published evidence to each inference, revealing a critically underdeveloped validity argument supporting dismissal decisions based on exam performance.

In theme 2, participants overwhelmingly perceive that the exam's purpose is to "weed out" residents who are unlikely to pass subsequent written certification exams. This perception appears rooted in the ABA's policy linking Basic Exam performance to graduation. Some view the exam as a means to distinguish physicians from other anesthesia providers, or, in line with the ABA's stated purpose, as a data point supporting remediation. Best practice guidelines highlight the importance of interpreting and using results as intended and note that the intended interpretations and uses of test scores (to develop study habits and foundational knowledge) may fail because of a lack of alignment with their actual use (dismissal).[24] The Standards specify that it is the responsibility of the test publisher, the ABA, to clarify appropriate uses and caution against misuse of assessment results.[25]

Theme 3 outlines the exam's influence on resident selection, curricula, and operations. The term "elbow grease" is frequently used to describe the additional input required to support applicants with historically low performance on standardized exams. "I felt like I wanted to take those risks as an educator… to have a training program that was reflective of my patient population. I couldn't just take people that were good at taking tests… but that meant more elbow grease." Others view this risk as unfair to the applicant, asserting, "I would rather not take somebody that we don't think is going to pass, than to take them and have to remove them from our program because they couldn't pass a test." Operational concessions vary widely; clinical hours, call schedules, and days off for preparation may introduce construct-irrelevant variance. For example, if one program provides a month off to study while another does not, residents' exam performance may reflect their preparation time more than their knowledge. Furthermore, most programs offer dedicated Basic Exam curricula structured around the content outline. Many incorporate practice testing programs, and some overtly teach test-taking skills. The Standards caution that high-stakes assessments can distort the construct they were intended to measure. When curricula targeting exam preparation replace sessions aimed at developing broader knowledge domains, such narrowing tends to undermine the validity of the assessment for any purpose.[25]

Studying is an intended consequence of the exam, as supported by theme 4, with an important caveat: participants distinguish that residents prepare for the exam in a manner that mirrors its format. They repeatedly note that residents' use of question banks has supplanted case-based reading and review of empirical literature. Washback describes the impact of high-stakes assessments on the teaching and learning environment. The literature on washback provides context for this theme, noting that intrinsic and extrinsic performance goals promote different learning strategies. Although intrinsic goals tend to foster deep approaches such as elaboration and critical thinking, extrinsic goals encourage surface learning strategies, such as rote memorization.[26]

Participants report that the exam is a significant source of stress and distraction. They attribute this unintended consequence to the exam's stakes and its placement in the first year of clinical anesthesia training. One participant remarked, "It seems like the worst time… to put a high-stakes test on somebody who's just learning how to be an anesthesiologist." Another noted, "They're feeling like their patient care is suffering

continued on next page

continued from previous page

because they're stressed about the exam." Financial stress also emerges as a concern, with one participant asking, "Where's the $1000 coming from? ... They barely make ends meet, these kids." In addition, participants expressed that the exam's molecular content, although important, is unlikely to be encountered in the primary learning activity of residency—clinical care. "[There's] definitely more of a focus on 'I need to be able to go home and study' and less of a focus on, 'I need to get as much experience as I can to be a…clinically competent anesthesiologist.'" The Standards consider issues with content alignment a source of construct-underrepresentation and a threat to the valid interpretation of score results.[25]

Theme 6 summarizes the consequences of failing the Basic Exam. The stakes associated with this exam suggest that passing is essential for practicing clinical anesthesia. Residents reflect on their failures through this lens. One PD recalled, "She was clinically superb, and it only takes one failure, and she just became a shell of herself." Processing failure consumes valuable time in the brief interval between the score release in July and the retake opportunity in November. Rotational changes are often made to delay subspecialty rotations, allowing for adequate study time. These changes may impact the clinical rotations of both affected residents and their peers, who cover the additional clinical workload. "If this resident wants to move [rotations], somebody else needs to take that person's place. So, [one failure] affects everybody." Similar cycles have been described as negative consequences of high-stakes, summative testing programs among different learner populations.[27] Participants also noted the impact on the fellowship application process for those who experience failure. They described a direct impact, as passing the exam takes priority over fellowship applications, and an indirect impact, alluding to fellowship programs' usage of In-Training Examination (ITE) results: "If you fail the Basic Exam, you cannot really spend as much bandwidth prepping for the ITE in the way that you would like for the advanced content… it may impact your fellowship progression."

Finally, theme 7 synthesizes implications for fairness in testing. One participant without a dismissal policy remarked, "Across town, my current resident… would've already been fired, no questions asked." A participant with a 2-failure dismissal policy describes the importance of transparency with applicants, "Learners deserve to know, and they deserve to have the choice of whether they want to be part of that kind of program or not." When asked what inferences can be made on behalf of Basic Exam performance, one participant replied, "A resident who didn't pass this year has two kids under two, a full-time job with us, and her husband is also a physician and is working in out-of-state, I don't think that says anything about her skill as an anesthesiologist… Who are you weeding out?" Another framed the need for robust validity arguments as we consider time-variable training, "On average, underrepresented residents do worse on standardized tests. So, can you imagine a situation where we have mostly white residents graduating early and mostly underrepresented residents not?" The Standards highlight that differential opportunities to learn carry significant legal implications when access to learning opportunities is unequal, yet individual residents are still held accountable for their assessment results.[25]

Several limitations exist in interpreting our results. First, this study is exploratory rather than confirmatory. Our findings aim to inform future validation efforts, which are most comprehensively addressed using a validation framework and a multifaceted methodological approach. Given our low response rate to member-checking efforts, a follow-up survey may provide additional context for the results of this thematic analysis. Our research questions focused on the implications inference as the most important and often least supported aspect of a validity argument according to Kane. However, the Basic Exam's extrapolatory argument, which correlates exam performance with clinical performance, is also undersupported (Table 3). In this study, robust data emerged at this inference, including 189 segments of text coded as "correlation to clinical performance," and a secondary analysis of our dataset may be informative. Moreover, each separate use of Basic Exam results necessitates its own validity argument. Because usage varied widely among programs, we selected the highest-stakes aspect to inform the validity argument for this study. Although we conclude that the argument for this use is critically undersupported, this does not imply that the Basic Exam's use is invalid for all purposes. Next, our perspectives as researchers interact with our qualitative data, and although we have been intentional about reflexivity and positionality, this may still pose a limitation.

The Basic Exam has been in use for nearly a decade, positioned by ABA policy as a gatekeeper to the field of anesthesiology. However, there remains an underdeveloped validity argument for its use in dismissal decisions, particularly regarding the implications inference. Beginning in January 2025, the ABA will remove most policy mandates related to the Basic Exam—including those concerning graduation and extended training—leaving programs with full discretion over policy development. This shift places an increased responsibility on programs to ensure sufficient validity evidence supports their use of Basic Exam results, underscoring the timeliness of our study. Our findings highlight significant consequences for both programs and residents, with potential implications for fairness in testing, offering relevant insights as programs navigate these policy changes. Future steps should include exploring the themes in this study more broadly with a follow-up survey to PDs, directly evaluating implications with residents, and assessing the characteristics of those who experience failure.

Acknowledgments

The Master of Health Professions Education faculty at New York University and Maastricht University were formative in this work, and their mentorship was invaluable throughout the process. We would like to express our sincere gratitude to all participants who generously shared their time and valuable insights in this qualitative study. Your thoughtful contributions were crucial to the depth and richness of our findings.

continued on next page

continued from previous page

References

1. Dijkstra J, Galbraith R, Hodges BD, et al. Expert validation of fit-for-purpose guidelines for designing programmes of assessment. *BMC Med Educ*. 2012;12(1):20.
2. Zhou Y, Sun H, Lien CA, et al. Effect of the BASIC examination on knowledge acquisition during anesthesiology residency. *Anesthesiology*. 2018;128(4):813-20.
3. American Board of Anesthesiology. ABA Policy Book. https://theaba.org/pdfs/Policy_Book.pdf. Accessed June 30, 2022.
4. Zhou Y, Sun H, Macario A, et al. The American Board of Anesthesiology's staged examination system and performance on the written certification examination after residency. *Anesth Analg*. 2019;129(5):62.
5. Pivalizza EG, Nwokolo OO, Ghebremichael SJ, et al. Does the American Board of Anesthesiology BASIC Examination really affect anesthesiology resident knowledge acquisition? *Anesthesiology*. 2018;129(6):1189-90.
6. Murray DJ, Boulet JR. Anesthesiology board certification changes: a real-time example of "Assessment Drives Learning." *Anesthesiology*. 2018;128(4):704-6.
7. Berman DJ. In-training exams, performance, and exam fatigue. *Anesthesiology*. 2018;129(6):1189.
8. Cook DA, Hatala R. Validation of educational assessments: a primer for simulation and beyond. *Adv Simul (Lond)*. 2016;1:31.
9. Cook DA, Brydges R, Ginsburg S, Hatala R. A contemporary approach to validity arguments: a practical guide to Kane's framework. *Med Educ*. 2015;49(6):560-75.
10. Hawkins RE, Margolis MJ, Durning SJ, Norcini JJ. Constructing a validity argument for the mini-Clinical Evaluation Exercise: a review of the research. *Acad Med*. 2010;85(9):1453-61.
11. Hatala R, Cook DA, Brydges R, Hawkins R. Constructing a validity argument for the Objective Structured Assessment of Technical Skills (OSATS): a systematic review of validity evidence. *Adv Health Sci Educ*. 2015;20:1149-75.
12. ACGME. ACGME Program Search. https://apps.acgme.org/ads/Public/Programs/Search. Accessed June 30, 2022.
13. Hennink M, Kaiser BN. Sample sizes for saturation in qualitative research: a systematic review of empirical tests. *Soc Sci Med*. 2022;292:114523.
14. DiCicco-Bloom B, Crabtree BF. The qualitative research interview. *Med Educ*. 2006;40:314-21.
15. Phillippi J, Lauderdale J. A guide to field notes for qualitative research: context and conversation. *Qual Health Res*. 2018;28(3):381-8.
16. Kenevan MR, Stewart TM, Warner MA, et al. The changing characteristics of anesthesiology program directors. *J Educ Perioperative Med*. 2018;20(3):E625.
17. Markham TH, de Haan JB, Guzman-Reyes S, et al. Anesthesiology resident performance on the US Medical Licensing Examination predicts success on the American Board of Anesthesiology BASIC Staged Examination: an observational study. *J Educ Perioper Med*. 2020;22(3):E646.
18. Zhou Y, Sun H, Culley DJ, et al. Effectiveness of written and oral specialty certification examinations to predict actions against the medical licenses of anesthesiologists. *Anesthesiology*. 2017;126(6):1171-79.
19. Baker K, Sun H, Harman A, et al. Clinical performance scores are independently associated with the American Board of Anesthesiology certification examination scores. *Anesth Analg*. 2016;122(6):1992-99.
20. Hughes FP, Hug CC, Longnecker DE, Saidman LJ. A demonstration of validity for certification by the American Board of Anesthesiology. *Acad Med*. 1994;69(9):740-6. https://doi.org/10.1097/00001888-199409000-00020
21. Murray DJ, Boulet JR. Anesthesiology board certification changes. *Anesthesiology*. 2018;128(4):704-6.
22. Markham TH, de Haan JB, Guzman-Reyes S, et al. Potential harm of elimination of score reporting for the United States Medical Licensing Examination Step 1 Examination to anesthesiology residency selection. *Anesth Analg*. 2021;132(1):275-79.
23. Chen YK, Lekowski RW, Beutler SS, et al. Education based on publicly available keyword data is associated with decreased stress and improved trajectory of in-training exam performance. *J Clin Anesth*. 2022;77:110615.
24. O'Leary TM, Hattie JA, Griffin P. Actual interpretations and use of scores as aspects of validity. *Educ Meas Issues Pract*. 2017;36(2):16-23.
25. American Educational Research Association, American Psychological Association, National Council on Measurement in Education (Eds.). *Standards for Educational and Psychological Testing*. Washington, DC: American Educational Research Association; 2014.
26. Liu X, Yu J. Relationships between learning motivations and practices as influenced by a high-stakes language test: the mechanism of washback on learning. *Studies in Educational Evaluation*. 2021;68:100967.
27. Harrison CJ, Könings KD, Schuwirth L, Wass V, Van der Vleuten C. Barriers to the uptake and use of feedback in the context of summative assessment. *Adv Health Sci Educ*. 2015;20:229-40.

***Riley S. Carpenter Lide** is an Associate Professor and Residency Program Director in the Department of Anesthesiology, University of Arkansas for Medical Sciences, Little Rock, AR. **Rachel Moquin** is an Associate Professor and Associate Vice Chair for Faculty and Educator Development, Department of Anesthesiology, Washington University School of Medicine, St. Louis, MO. **Erin Green** is an Assistant Professor of Emergency Medicine and Assistant Dean for Clinical Learning, Medical College of Wisconsin–Green Bay, Green Bay, WI.*

***Corresponding author:** Riley S. Carpenter Lide, Department of Anesthesiology, University of Arkansas for Medical Sciences, 4301 West Markham, #515, Little Rock, AR 72205. Telephone: (501) 686-6119, Fax: (501) 686-7059*

***Email address:** Riley S. Carpenter Lide: rlide@uams.edu*

***Funding:** Not applicable.*

***Conflicts of interest:** The authors declare no competing interests.*

Abstract

Background: In 2014, The American Board of Anesthesiology introduced the Basic Examination as a graduation requirement for second-year anesthesiology trainees. The exam's validity has been supported by evidence demonstrating enhanced performance on other standardized exams; however, an assessment's validity is inseparable from decisions made on its behalf. This study aimed to understand the usage and implications of the Basic Exam within training programs to construct a comprehensive validity argument.

Methods: Semistructured interviews were conducted with a sample of 20 program directors from Accreditation Council for Graduate Medical Education–accredited anesthesiology training programs. Thematic analysis was performed by a 3-member team.

Results: A 56-item codebook was developed and applied to the 20 transcripts, yielding 1941 coded segments organized into 7 themes. Theme 1 highlights varied programmatic policies, including dismissal (1a). Theme 2 addresses the perceived purposes of the exam: as a tool to "weed out" residents unlikely to achieve board certification (2a), a data point supporting remediation (2b), and a distinguishing accomplishment of physician anesthesiologists (2c). Theme 3 captures programmatic implications for recruitment (3a), operations (3b), and curricula (3c). Theme 4 confirms that residents are studying for the exam, emphasizing targeted test preparation (4a). Theme 5 discusses resident implications, including stress (5a) and clinical distraction (5b). Themes 6 and 7 explore the implications of failure and equity concerns, respectively.

Conclusions: This study identifies a significantly underdeveloped validity argument supporting dismissal based on Basic Exam results and explores implications to guide future validation efforts.

Keywords: Assessment validation, high-stakes testing, anesthesiology/education, graduate medical education/standards, equity

continued on next page

continued from previous page

Tables

Table 1. *Results of Thematic Analysis and Supporting Quotations*

Theme 1: Programmatic Policies	
Subtheme 1a: Number of Attempts	
Programs exhibit significant variability in the number of attempts allowed for the Basic Exam. One program director mentioned a historical policy of immediate dismissal after a first failed attempt, which was rescinded in 2020. Three programs maintain policies advocating for dismissal after a second failed attempt and 8 programs support dismissal after a third failed attempt. Much of the reluctance to permit additional attempts stems from the American Board of Anesthesiology (ABA) policy, which stipulates that a resident cannot graduate without successfully passing the Basic Exam. These policies aim to avoid serial extensions of training, which are mandated after the third failure.	
Supporting Quotations	*Participant 2: After a third failure, you then must leave the program. So, you could have 2 full years of training that counts for nothing that you'd have to then give up.*
	Participant 9: If they fail twice, which […] haven't […] yet, then our policy is that they will be let go from the program.
	Participant 10: There used to be [a limit] ... 1 fail, they're out. I changed all that. We have no formal policy.
	Participant 12: Given that the pass rate was in the mid-90s and in aggregate with the presentation that [historical ABA Board Member] made on the topic, it seemed like a logical thing to have […] [a] 2 strikes and you're out policy with a strong emphasis on remediation.
	Participant 13: Our policy is that if [there are] 3 fails, we dismiss from the program to avoid the mandatory extension of training that comes along with a third fail.
	Participant 15: The policy says that, since failing 3 times automatically extends your training, you are probably not allowed to fail 3 times because the GME may have something to do with your ability to continue in the program [extension of training].
	Participant 20: We don't have a formal cap… after the second fail, I think that's a fair question to ask a resident, "Look, we want you to succeed. We want you to fulfill your dream of being an anesthesiologist because that's what you want to do, but you have to, at least in the current system, pass these major exams, and there are 2 more [exams] looming after you pass the basics [Basic Exam]." So that's a key point. If they struggle mightily and they just can't pass them, we don't want them to languish in our program year after year and not be able to pass, because there's no value to residency training if completing it if you're not eligible to sit for the boards.
	Participant 19: [No policy]. If I help this trainee […] through these fourth attempts or fifth attempt or sixth attempts, and finally they pass, but they're phenomenal clinicians.
	Participant 17: I've had one resident fail 3 times and we've never threatened termination. I've just said, "No, we're going to give you the time." I've given residents weeks off to prepare, get their game on, no stress.
	Participant 3: So, we don't have an official policy in terms of maximum number of attempts, which is something that I just talked with my chair about, about possibly revising that standing now that we're in a situation where we have a CA3 has going to have to extend by 6 months.

continued on next page

continued from previous page

Tables *continued*

Subtheme 1b: Resignation as a Surrogate for Dismissal	
This latent theme emerged as participants responded to the prompt, "Tell me a story about a resident who was affected by your program's Basic Exam policies." The term dismissal was rarely used. Instead, participants described residents facing an ultimatum of dismissal or resignation, or "counseling themselves out" of the specialty due to repeated exam struggles.	
Supporting Quotations	*Participant 6: Before this most recent Exam, I only had 1 resident who failed it more than once, and that person was struggling clinically, was struggling with [United States Medical Licensing Exam, USMLE] Step 3, was struggling with multiple things on their initial remediation plan, was struggling on their initial probation plan. And then eventually after their fourth fail, we said that we were not going to renew the probation, and the resident resigned.*
	Participant 8: We did have 1 resident who failed the Basic Exam. He also was struggling clinically. It turned out that he had [a] new [medical condition], which hadn't been really diagnosed, and […] a lot of other health issues. And then he just never strongly indicated that he wanted to continue, and certainly would've needed to pass the Basic Exam to continue.
	Participant 12: There was 1 resident that failed and then sort of opted to reevaluate their performance within the program and then counsel themselves out and said they just weren't happy, and we were able to successfully place them in a completely different program.
	Participant 14: It got to the point that this person would have needed to extend residency for a year. We absolutely supported this person in doing so, but this person had just had enough. This person just felt beaten down and this person just... They just couldn't take it anymore. What that means is this person did not finish a residency in spite of spending more than 4 years in residency. I think this person spent 4 and a half years in residency because they did have to extend for 6 months. Again, the Basic was the hurdle.
	Participant 16: I think dismissed is also a very soft word because sometimes residents are counseled that it's in your best interest to not try to keep going…So I think our current policy is written as, "At the third attempt, you will not progress to the next stage of training," and that is open. And when we have had that happen in the past, that individual has chosen to leave.
Theme 2: Discrepancies in Stated and Perceived Purpose of the Basic Exam	
The stated purposes of the Basic Exam, as outlined by the ABA, are as follows: (1) to incentivize residents to develop positive study habits early in training, (2) to encourage a focus on foundational content areas for subsequent training, and (3) to enable program directors to identify residents needing additional support early in their training. However, participants' interpretations of the exam's purpose vary, which is significant, as purpose informs the utilization of exam results. Participants' understanding of the purpose is elaborated on in the subthemes below.	

continued on next page

continued from previous page

Tables *continued*

Subtheme 2a: "Weed-Out" Tool for Those Unlikely to Achieve Board Certification	
Participants cross-reference Accreditation Council for Graduate Medical Education (ACGME) and ABA policies to deduce that the exam's purpose is to protect programs and the specialty from graduating residents unlikely to achieve board certification.	
Supporting Quotations	*Participant 10: The real goal, remember, was the ACGME. So the answer is gone are the citations for failure to pass tests, because you kind of weaned out the person who's not going to pass the Advanced because they didn't pass the Basic [Exam]… You just had an In-training Exam score and you tried to act on it, but…That's just intended to measure a resident's current knowledge. How do you intervene on that? Do you not graduate somebody because of an In-training Exam? It made it hard, right? So, along came the Basic [Exam].*
	Participant 8: It has the potential to either weed out or bring up to speed residents who, through no fault of their own, may have been at medical schools where the basic sciences weren't as well taught as at other places. So, you're helping establish more of a level playing field for going forward in residency, potentially. Again, the weeding-out factor, so that if somebody is truly unlikely to pass their boards, maybe they don't spend years and years doing this, and then discover that they can't. Or if they're not even really suited to anesthesia, it can be a way to weed them out. Or you can be using it as a way to weed them out, even if they could eventually pass.
	Participant 12: He said [that], based on their understanding of the Exam performance, if somebody was unable to successfully pass the Basic Exam after 2 attempts, the probability that they would go on to pass then the advanced Exam, the oral Exam, and then eventually the OSCE [Objective Structured Clinical Exam], which came into being all in sequencing under the 7-year timeframe after graduation, was slim to none. So that might enter into the program's decision-making on the number of time attempts that they allow a candidate to have.
	Participant 13: Yes, if you don't pass this exam, we dismiss you. It doesn't matter how much I like you, how good of a citizen you are.
Subtheme 2b: The Basic Exam as a Data Point for Supporting Remediation or Dismissal When Other Aspects of Resident Performance Are Unsatisfactory	
Supporting Quotations	*Participant 6: Before that one resident that I had, I would've said really nothing. But I think that that was helpful for us to say, you know what, anesthesia isn't right for you. You got to get your life together. You've got to pass [USMLE] Step 3. You've got to learn how to pass this Exam and do various other things that that person had to do in their lives before becoming an anesthesiologist. And that was helpful for us to say that, you know what, you failed it 4 times. We're not going to continue on with your probation. It's time to move on.*
	Participant 11: I'm going to say, maybe the Basic [Exam] was helpful there in that it allowed me to turn up the heat temporarily. But we've still had ongoing conversations around stuff that they need to work on.
	Participant 14: I think it actually got this person's attention in a way that might not otherwise have happened… The policy says you have to fail it 3 times before you extend. I don't think it's unreasonable, frankly. I think it gives that resident and that faculty a little bit of extra time to just make sure that clinically, they're doing fine.
	Participant 15: If there's one thing that's valuable, is that it's another data point.

continued on next page

continued from previous page

Tables *continued*

Subtheme 2c: Understood Purpose 3: A Rigorous Board Examination Series as a Factor Distinguishing ABA-Certified Anesthesiologists from Non-physician Practitioners	
Supporting Quotations	*Participant 18: That's what I heard at the [annual anesthesiology program director meeting] that we have to hold ourselves to certain standards, and this is perhaps one of the standards, standardized tests, and stuff. But for the trainees, I would say it adds very little, especially in the current format. Now, if it's in a more clinical format, maybe, but in the current format, it adds very little or no value in my opinion.*
	Participant 14: I do think it's important that we distinguish ourselves from our nurse anesthetists and AA [Anesthesiology Assistant] colleagues. This is one of the ways that we distinguish ourselves. Our nurse anesthetists know that if they push succinylcholine, […] there'll be some fasciculations and then the patient will be paralyzed. But part of what we think distinguishes us as physicians is that we, at some point, knew that there was an alpha 2 subunit.
	Participant 19: They also wanted to show that anesthesiology is not like a nursing profession. We are above that. Our goal is to make sure that our residents understand the basics of anesthesiology. Having that knowledge, then they apply that knowledge in their clinical factors. And with that, hopefully, the goal was to achieve certain milestones, but I have not seen anything from ABA [to support that].
	Participant 2: The thing that gets brought up are things like, well we need some sort of metric to differentiate us from CRNAs [Certified Registered Nurse Anesthetists] or to make sure the public knows that it's rigorous to become an anesthesiologist. And I think that surviving residency and being able to pass a rigorous oral board and OSCE [objective structured clinical exam] is pretty significant.
Programmatic Implications	
The Basic Exam has broad programmatic implications, which can be categorized into 3 areas: resident selection, curriculum, and operational resources. Within each domain, there is considerable variability among programs.	
Theme 3: Programs Are Dedicating Vast Resources to Support Preparation for the Basic Exam	
Subtheme 3a: Resident Selection	
Beginning with resident recruitment, participants discussed the implications of ranking applicants whose test-taking abilities may be challenged by the Basic Exam. Others expressed a willingness to invest effort, encapsulated by the term "elbow grease," reflecting the anticipated early intervention and resources required for selecting applicants with marginal standardized test-taking abilities.	
Supporting Quotations	*Participant 9: This is going to sound really selfish, but I would rather not take somebody that we don't think is going to pass, than to take them and have to remove them from our program because they couldn't pass a test… when you talk to folks in other specialties and they're like, "Oh, it's great, [USMLE] Step 1 score is going away. This is all great and all this," it's like maybe, but our residents need to pass a test in the middle of their training, and if they can't pass a test it doesn't matter how good they are in everything else, we can't graduate them from residency.*
	Participant 17: It [USMLE step scores] does carry still a lot of weight because it is directly proportional to exam taking... It's hard. It does take resources and it is debilitating to the resident who's struggling through these things.
	Participant 3: But the question I have that I can get conflicted about is how many of those people can I take and invest the fact that there's going to be a lot of elbow grease that has to go into getting them to then pass these American Board of Anesthesiology exams to get them through.
	Participant 12: And with the Basic Exam, we did take a little bit of a pause on that and think about it more. I always felt strongly that we could get everybody through, and I was largely correct with that one notable exception. I felt like I wanted to take those risks as an educator and as somebody [who] wanted to have a training program that was reflective of my patient population and would be the best, that I couldn't just take people [who] were really good at taking tests. I wanted to train the best people in the world; they weren't all just going to be numerologists and wizards in the library. They would come from different backgrounds and have different skills, but that meant more elbow grease.

continued on next page

continued from previous page

Tables *continued*

Subtheme 3b: Operational Resources	
Many programs cite logistical challenges that limit operational support for Clinical Anesthesia Year 1 residents (CA1s) as they approach the Basic Exam. Most programs cannot implement significant adjustments to alleviate clinical workload, call burden, and night-to-day scheduling fluctuations before the exam. However, a few programs have made operational changes to reduce clinical work hours for CA1 residents before their first exam attempt.	
Supporting Quotations	*Participant 2: For example, we can't take all our CA1s off of challenging rotations for the month before the Basic Exam. So, some of them are going to be in the ICU. And I've had residents who tell me, look, I don't think I'm going to do well on this exam…. I'm working every third night, 28 hours… what am I supposed to do? I got nobody else to put in there except another CA1.*
	Participant 15: It just takes them out of the OR for an extra day and creates a whole big stress, and also it impacts how I structure the rotations and everything else because it's a priority. Certainly, no question about it, okay, distorting and compressing the timeline for real clinical learning, no doubt about it.
	Participant 7: Our residents previously used to work till 7:00, 8:00 PM, fairly regularly in the ORs. And once the Basic Exam was instituted, we realized that we had to because there's such a limited amount of time for them to prepare... the goal was to have our residents relieved from the ORs by 5:00 PM in order to protect some academic time in the evening.
	Participant 10: But we don't get them out… [Another program at which this participant was Program Director] used to get them out for a month before.
Subtheme 3c: Curricular Resources	
Most participants have developed a dedicated Basic Exam curriculum based on the content outline, with some programs explicitly teaching test-taking strategies. Many reported improvements in CA1 feedback when the didactic curriculum is aligned with the exam. Other programs noted the limitations of a didactic curriculum focused on test preparation and have shifted away from this approach, sometimes at the expense of resident satisfaction. Some participants indicated they would allocate didactic time differently if the Basic Exam did not exist.	
Supporting Quotations	*Participant 4: A lot of [residents], when they failed, at least brought up the fact that they felt like we didn't have a structured lecture system and that it made it a little bit more difficult, especially with the Basic [Exam], because it's more science stuff. It's not as much fun to study as it is for clinically related stuff. Now, all the CA1s on the first Wednesday of the month get pulled out for 3 hours in the morning. They do grand rounds and then have 2 to 2 and a half hours of dedicated lecture time before they ever go to their clinical site.*
	Participant 1: Our residents have a lecture series every Monday and Wednesday from 4:00 to 6:00 PM, and these are mandatory lectures and the lecture topics are strictly pretty much focused on the Basic Exam. We take the Basic Exam keyword list, and pretty much we break it down.
	Participant 14: [Lecture] is run by a group of faculty who talk through concepts and test-taking strategies.
	Participant 20: I guess one has to ask, do we want to have a curriculum that's based on learning to take an exam and pass it? …And some of our residents sometimes complain, "This isn't going to help me on an exam." And my response is, "I know. That's the point." Right? That's the point. So, I would counter the argument that a curriculum needs to be based on exam prep by saying I think it misses the mark of a good curriculum.
	Participant 18: If the Basic Exam didn't exist, there are so many other things that we could do… unfortunately with the amount of time and space we have, we need to help the trainees to the best. And that's why our didactics have a heavy [Basic Exam] keyword component... is this the right thing to do? Yes, if the Basic [Exam] exists. If there is no Basic [Exam], we would switch back to full clinical and more advanced forms of lectures.

continued on next page

continued from previous page

Tables *continued*

Resident Implications	
Theme 4: Residents Are Studying for the Basic Exam	
Some participants recognize intrinsic value in early studying and knowledge reinforcement during residency, behaviors that the Basic Exam supports and that are considered an intended consequence of the exam.	
Supporting Quotations	*Participant 5: I know that the board points to improvements, and the residents' performance and things like In-training exams (ITEs), and they credit that part probably to more intensive study in the middle of residency for the Basic [Exam]... And again, so that's not the end-all, but I think it does help the residents to develop study habits or to figure out how, as [...] residents, they need to be reading and studying earlier in their residency than they might have to do otherwise when it's just one exam at the end of the residency.*
	Participant 10: So, we wrote the first Basic [Exam]...and it really was intended to hopefully get people studying. So, you see the literature that came out from the board that showed that in-training scores went up, and again, no surprise. When you introduce a high-stakes test, imagine that people start studying earlier.
	Participant 11: That said, I do think that every residency program has some residents who don't have as much of the intrinsic motivation to study. And so these external things put a little bit more pressure to motivate them to study. I don't know that it's that much more than what the In-Training Exam already offered.
Subtheme 4a: Targeted Test Preparation	
Many participants noted that residents are studying in a manner that mirrors the exam format, observing that the use of question banks has supplanted traditional reading.	
Supporting Quotations	*Participant 11: And I think true north is where you want to head and magnetic north is studying for tests. If you're just starting off, it probably doesn't matter because they're more or less in the same direction. But if you follow it to its end, you're somewhere like 200 miles away from where you want to be. So, the idea [is] that there's a type of studying that you do to be able to recognize enough to be able to select a multiple-choice answer that's correct. And I think that's different from the type of thing that you need to do to be a good clinician.*
	Participant 20: I think most residents, right or wrong, I'm not sure, probably don't do a lot of general reading. And that has been substituted, in my opinion, by doing questions... But I do find that what troubles me is that some of our residents can get the answer on an exam, great. But if you really probe their knowledge further, they don't really understand or know, in an in-depth level, anything more than the surface of what the question is in front of them on an exam.

continued on next page

continued from previous page

Tables *continued*

Theme 5: Consequences for First-Time Test-Takers	
Consequences for first-time test-takers include stress and clinical distraction, both attributed to the high stakes of the Basic Exam.	
Subtheme 5a: Stress	
The code "stress" was applied only when participants discussed residents who had not failed the Basic Exam. This broad code was later subcoded to better understand its manifestations among residents. Subcodes included various sources of stress: the exam's placement in the CA1 year, its impact on wellness, stress arising from its high stakes, financial stress, disproportionate stress for value-added performance, and exam fatigue.	
Supporting Quotations	*Participant 4: You finish your first year of just learning how [to get] over that fear of not killing someone, and then you're like, "Well, bam, take this exam."*
	Participant 16: The Clinical Anesthesia- 1 (CA1) year…seems like the worst time, in my opinion, to put a high-stakes test on somebody who's just learning how to be an anesthesiologist, which is very, very different from literally everything else they've ever done. And I don't know that [placing] that test after 1 year is the best time for it when they're unsure of themselves in the OR already.
	Participant 15: It's shameful for the ABA to be asking them to pay $900 for this exam ... Where's the $1,000 coming from? ... They barely make ends meet, these kids. Did you know that in my county here they qualify for affordable housing?
	Participant 2: So, not only are they stressed about the exam, but they're feeling like their patient care is suffering because they're stressed about the exam. It's impacting their ability to sleep even when they have a chance to sleep and they're more tired and more stressed at work. So, I think there's no question that the stress of the exam is impacting the wellness [and] the stress level of our CA1s.
Subtheme 5b: Clinical Distraction	
Participants noted the impact of the exam on clinical learning, observing a shift in resident priorities that favored exam preparation over clinical duties. There is a perceived need to relieve residents of clinical responsibilities to allow for study time or to provide stable sleep schedules leading up to the exam. In addition, residents engaged in test preparation activities, such as using question banks, while providing direct clinical care.	
Supporting Quotations	*Participant 3: So, there's certainly some residents [whose] primary goal is to get out of the operating room (OR) as soon as they can in the afternoon, not to leave the hospital or anything like that. And all of my faculty are saying, this person has such terrible professionalism issues, all they want to do is leave…However, I also have this reality of being able to have to sit for this high-stakes exam. So, I need to put in the necessary footwork for that. So it's not sufficient to sit in the OR, it's something I have to go sit in the library for...From a clinical perspective, what I'll say is there's definitely more of a focus on "I need to be able to go home and study" and less of a focus on "I need to get as much experience as I can to be able to be a truly good competent clinical anesthesiologist."*
	Participant 5: I do think, and I agree with the board, the ABA's assessment is that it has probably caused or induced residents to do more serious reading or study, for lack of a better word, during the middle of their residency, rather than just pushing all of the things to the end, but it has certainly pushed some undesirable behavior…the expectation that we cover residents and that we relieve them and we don't have them on call.
	Participant 12: Even during their time in the operating room, sometimes, they would have it up on the computer or something like that or on their iPads or things. I never knew how to feel about that. I'm somebody [who] can't concentrate very well on taking care of patients and doing questions and stuff. We had our CA1s in major index cases by October, in their CA1 years... if you're doing major casework and you don't read about it, you're not going to be able to take care of the patients.

continued on next page

continued from previous page

Tables *continued*

Theme 6: Consequences After a Basic Exam Failure	
The cycle of failure emerges as a latent theme, with participants collectively describing the effects of failing the exam on residents.	
Subtheme 6a: Emotional Impact and Personal Sacrifices	
Participants described the emotional impact of failure on residents. Recovering from this emotional toll consumes valuable time between the release of results in July/August and the retake exam in November/December. Furthermore, it appears to negatively affect the clinical performance of residents who previously excelled. Participants noted that residents remain affected by a failing result even after a successful retake. Personal sacrifices made by residents to prepare for the Basic Exam retake include using vacation time for study and facing challenges in maintaining a work/life balance.	
Supporting Quotations	*Participant 10: So, the biggest thing, they think it defines them as an anesthesiologist. We spend a long time just trying to build them up from that, right? But it's hard. You get told you failed, and if you fail the second time, you get a letter from the board saying you're unsatisfactory. It's pretty significant for their mental status.*
	Participant 18: They feel they are a failure. They feel that they're not worth it. I haven't had anybody who's actually suicidal or anything to that extent. But again, I don't know, they wouldn't tell me...I think this is a very unfortunate sort of barrier that we created for our own trainees... Especially if it's really affecting their wellness.
	Participant 20: So, before the impetus to study starts, I think there's that self-reflection and self-deprecation, just, am I meant to be doing this?
	Participant 4: I think, actually, it destroyed her confidence... This person was doing really, really well, and it took a good 4 weeks of being like, "This doesn't define you," and that support. They ultimately passed. I still think it sticks in their mind of failure.
	Participant 17: I've watched several residents completely knocked off their confidence game and decimated so they could be clinically superb, and they fail that board, and they have a hard time recoiling their confidence level, and it's a huge distraction... She was clinically superb and it only [took] the 1 failure, and she just became a shell of herself.
	Participant 20: So it changed his life probably for the worst in the short term because all of a sudden, he's like, "I need to find more time in my day to study." And that's hard because of a lot of family commitments. And so I think his family suffered a bit, sure. Right? There's a limit to time in the week for all of us residents, working hard clinically.
Subtheme 6b: Confidentiality and Isolation	
Concerns about confidentiality arise when scheduling concessions are made to accommodate a retake. Such concerns can further isolate residents who have experienced failure on the Basic Exam. Although peer support is valuable, it often comes at the expense of confidentiality.	
Supporting Quotations	*Participant 5: I mean, they're just very concerned about the stigma and being labeled... And it does cause a lot of soul-searching for them. Some of them are amazingly willing to take whatever input or help or guidance that's willing to be offered, some of them want to handle it very privately and really don't want any outside help. And we have to be very careful with confidentiality and privacy.*
	Participant 7: We really work to protect the anonymity of our residents who don't pass, which is hard too, I think, because they feel even more isolated.
	Participant 12: Again, working with their peers if they were comfortable, I thought was really powerful because they were close to the Exam to figure out what would work... Typically, people didn't want to admit they failed to their peers, but some people were pretty comfortable with that.

continued on next page

continued from previous page

Tables *continued*

Subtheme 6c: Subspecialty Rotations	
As residents progress to subspecialty rotations following a failure, a misalignment occurs between the Basic Exam content and the teaching and learning activities of clinical training, compounded by time constraints imposed by the clinical workload. This frequently necessitates schedule adjustments to facilitate a successful retake.	
Supporting Quotations	*Participant 9: Then they have to start studying all over again when at the same time they're doing a lot more of their subspecialty rotations. The expectation from those faculty would be that they're reading and learning about the cases that they're doing in these subspecialties, but they really feel like they have to be studying for the Basic [Exam] because now the pressure to pass this exam is even higher, right?... So I think it actually may detract from their clinical learning, and their learning for the advanced exam because they are repeating what they'd done before while in the midst of these higher-level rotations.*
	Participant 12: For many programs where they didn't start subspecialties until [the] CA2 year, that'd be a huge problem because every month is subspecialty month... I couldn't imagine sending somebody to a cardiac or neuro for the first time while they're trying to cram for this Basic Exam.
Subtheme 6d: Operational Impact	
Most participants indicated that they provide lighter rotational or call schedules for residents retaking the exam. In some programs, this operational support occurs at the expense of elective time in the CA3 year. These changes do not occur in isolation; other residents must cover clinical duties to accommodate scheduling changes following a failure.	
Supporting Quotations	*Participant 4: 90% of the time, we're changing schedules around, because people will start their Clinical Anesthesia-2 (CA2) year. If they have cardiac [rotation] where you're not going to have much spare time to study, we'll switch their schedule and push cardiac back. In some respects for some residents, I think it's a good thing.*
	Participant 9: We give them some elective time that we pull them out [of clinical rotations] to give them a couple [of] weeks to study and prepare if they want it, but that time is taken away from elective time from their Clinical Anesthesia-3 (CA3) years. So, instead of having an elective where they are learning how to run the board...they're using that time to study for an exam again.
	Participant 16: So, we have switched people off of rotations right before the Exam that we know are particularly heavy. So, somebody was on an ICU month, and we're like, "Well, you're never going to study on that month," so we moved that month away... We tried to rearrange their call schedules closer to the test date...so that they could potentially have several nights of not flipping from days to night sleep.
	Participant 18:. So we had to switch rotations and it affects the whole program because it affects some other residents. If this resident wants to move, somebody else needs to take that person's place. So it affects everybody.
Subtheme 6e: Fellowship Matching	
Failing the Basic Exam may disadvantage residents when applying for fellowships. This observation is attributed to the proximity of the Basic retake in November/December and the deadlines for fellowship applications. Other contributing factors include the reduced capacity to prepare for the CA2 ITE among residents retaking the Basic Exam and the reliance of fellowship programs on ITE scores for admission decisions. Finally, participants noted that the Basic Exam often takes precedence over other aspects of professional development, such as scholarly activity, which may further affect the competitiveness of fellowship applications for residents retaking the exam.	
Supporting Quotations	*Participant 7: I think maybe that's another big piece of the conversation too, is the impact on fellowship matching…fellowship applications are always looming within this next 6-month horizon for them once they get the results back.*
	Participant 12: If you fail the Basic Exam, you can't really spend as much bandwidth prepping for the ITE in the way that you'd like for the advanced content… it may impact your fellowship progression.
	Participant 6: But I think we put a hold on everything else, like research projects, until they've gotten past that point [of passing the Basic Exam].

continued on next page

continued from previous page

Tables *continued*

Theme 7: Equity	
When asked about inferences that can be drawn from a resident's performance on the Basic Exam, participants expressed concerns regarding fairness, referencing individual characteristics of residents that may limit the validity of such inferences based on exam results.	
Supporting Quotations	*Participant 16: Bias is obviously heavily apparent in every other standardized test that we have that exists. It seems pretty arrogant and unrealistic to assume that we, as anesthesiologists, are somehow immune to that bias. And so, I would like to know how those breakdowns go ... so I think having a breakdown of who's doing well on the test and not, who's writing the questions, who's not… I think that most standardized tests do obviously favor the majority and men. And so, I would worry that, "Are being biased towards residents that already have a really higher likelihood of getting professionalism reports who are subject to disciplinary actions and typically have higher dismissal rates than other residents?" And so, I would hate for us to use this as a tool to just further disparities.*
	Participant 3: I know across town, my current resident that I'm struggling with right now, he would've already been fired, like no questions asked. He would've been terminated already and would be one of the few desperate ones [who]… would've been messaging me and saying, "Listen, I failed in this previous program, got kind of kicked out because of my failure. Can I come join your program? I'm going to do better this time." So, I definitely think there's inequity from that perspective.
	Participant 2: And I think that if we're ever going to move to time variable training in anesthesiology, I think it's going to need to involve a really robust oral board-type OSCE-type examination and not standardized test taking. Because you can imagine, right, again, on average, underrepresented residents do worse on standardized tests. So, can you imagine a situation where we have mostly white residents graduating early and mostly underrepresented residents not? I mean, that would be a complete disaster, but that could happen if we rely on standardized testing.
	Participant 14: I also know that socioeconomic status plays a huge role in that. Because I just think that the resident who has... I had a resident who graduated 2 years ago who had 5 or 6 children. That resident, his ability to... His time to study for an exam was different from somebody who is single.
	Participant 9: I mean, a resident who didn't pass this year has 2 kids under 2, a full-time job with us, and her husband is also a physician and is working in another town, another state, actually, I don't think that says anything about her skill as an anesthesiologist... Maybe it is, and that's great then, but I don't know, maybe it's not, and then who are you weeding out?
	Participant 12: I mean philosophically... Look, if you're going to do something like that, learners deserve to know and they deserve to have the choice of whether they want to be part of that kind of program or not. And in fact, [on] interview day, residents or applicants signed a document for the training program, when I was program director, detailing the 2 strikes and you're out policy…Philosophically, that may help them make a choice between training programs.

continued on next page

continued from previous page

Tables *continued*

Table 2. *Inferences Required to Support the Use of the Basic Exam Results in Graduation Decisions*

Inference	**Description** Cook DA, Brydges R, Ginsburg S, Hatala R. A contemporary approach to validity arguments: a practical guide to Kane's framework. *Med Educat.* 2015;49(6):560–75. https://doi.org/10.1111/medu.12678
	The Standards American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, eds. *Standards for Educational and Psychological Testing*. American Educational Research Association; 2014 Page number is listed individually next to each Standard
Scoring:	A resident's score represents the knowledge the student demonstrated in responding to the items on the Basic Exam.
	Standard 6.1: Score precision might be depicted by error bands or likely score ranges, showing the standard error of measurement. Reports should include discussion of any administrative variations or behavioral observations in clinical settings that may affect results and interpretations (Page 114). **Standard 7.12:** When test scores are used to make predictions about future behavior, the evidence supporting those predictions should be provided to the test user. Comment: The test user should be informed of any cut scores or rules for combining raw or reported scores that are necessary for understanding score interpretations. A description of both the group of judges used in establishing the cut scores and the methods used to derive the cut scores should be provided (Page 129).
Generalization:	The resident's score (based on items answered correctly) is a good estimate of the score they would receive if they had answered all available Basic Exam formats and questions as well as basic content items on the in-training and advanced exams.
	Standard 9.13: The test user corroborates results from testing with additional information from a variety of sources, such as interviews and results from other tests (eg, to address the concept of reliability of performance across time and/or tests). When an inference is based on a single study or based on studies with samples that are not representative of the test-takers, the test user should be more cautious about the inference that is made (Page 145).
Extrapolation:	The resident's score is a good estimate of their knowledge relative to the necessary requirements for practicing clinical anesthesia safely. The implementation of the Basic Exam has resulted in improved patient safety during and after residency. Program leadership feels graduates are providing higher quality care compared with before the Basic Exam and the proportion of graduates who are board certified has increased.
	Standard 5.23: When feasible and appropriate, cut scores defining categories with distinct substantive interpretations should be informed by sound empirical data concerning the relation of test performance to the relevant criteria. In contexts in which distinct interpretations are applied to different score categories, the empirical relation of test to criterion assumes greater importance (Page 108). For assessments used in credentialing, suitable criterion groups (eg, successful vs unsuccessful practitioners) are often unavailable. Nevertheless, when appropriate and feasible, the test developer should investigate and report the relation between test scores and performance in relevant practical settings. **Standard 9.13:** In educational, clinical, and counseling settings, a test taker's score should not be interpreted in isolation; other relevant information that may lead to alternative explanations for the examinee's test performance should be considered. In clinical and counseling settings, the test user should not ignore how well the test taker is functioning in daily life (Page 145).

continued on next page

continued from previous page

Tables *continued*

Implications:	The implementation of the Basic Exam has not adversely affected subgroups recruited into the specialty. Decisions to delay graduation or dismiss residents who have not passed the Basic Exam contribute to improved patient safety, and remediation following unsuccessful attempts at the Basic Exam leads to objective improvements in patient safety. The cut score is supported by evidence addressing the risk of false positive and false negative classifications. Stakeholders, including residents, training programs, and the public, perceive a benefit in the utilization of the Basic Exam for this purpose. The Basic Exam does not impose an excessive burden on residents or training programs, and its performance is consistent across subpopulations, as evidenced by differential item functioning and overall pass rates. Mean differences between subgroup performance and programmatic performance are not attributed to construct underrepresentation or construct-irrelevant variance. The cost of retaking the Basic Exam, extending residency, and dismissing residents based on Basic Exam performance is acceptable to all stakeholders. Curricular resources dedicated to the Basic Exam have improved the educational experience for residents. Residents nationwide have an equal opportunity to succeed on the exam, including access to preparation resources, content experts, didactic curriculum, study strategies, time to prepare, and the number of attempts.
	Standard 9.0: Test users are responsible for knowing the validity evidence in support of the intended interpretations of scores on tests that they use, from test selection through the use of scores, as well as common positive and negative consequences of test use (Page 142). **Standard 12.1:** When educational testing programs are mandated by school, district, state, or other authorities, the ways in which test results are intended to be used should be clearly described by those who mandate the tests. It is also the responsibility of those who mandate the use of tests to monitor their impact and to identify and minimize potential negative consequences as feasible. Consequences resulting from the uses of the test, both intended and unintended, should also be examined by the test developer and/or user (Page 195).

continued on next page

continued from previous page

Tables *continued*

Table 3. *Validity Framework for Use of the Basic Exam in Dismissal Decisions*[a]

Use: Resident Dismissal Interpretation: A failing result on the Basic Exam indicates that a resident does not have the minimum level of knowledge necessary to practice clinical anesthesiology safely and independently; those who cannot pass the Basic Exam pose a risk to public safety when practicing independently.		
Inference	**Examples of Validation Evidence (derived from Cook et al[9])**	**Published Evidence**
Scoring	A resident's score represents the knowledge the student demonstrated in responding to the items on the Basic Exam.	Test administration and security procedures
Generalization	A resident's score (based on items answered correctly) is a reasonable estimate of the score they would receive if they had answered all available Basic Exam formats and questions. A resident's Basic Exam performance correlates across Basic content items on the in-training and advanced exams.	Zhou et al[2] noted a significant 2-point increase in in-training examination scores 6 months after implementation of the Exam, and in-training scores of second-year clinical anesthesia residents in the staged cohort reached the level of third-year residents in the traditional cohort. Zhou et al[4] used equated items among 2013 and 2014 cohorts and found a 7- to 8-point increase in the scaled score ($P < .05$) in the staged Exam 2014 vs 2013, which the article attributes to the implementation of the Basic Exam. This article also describes test equating procedures. Markam et al[22]: In a multivariable logistical regression, clinical anesthesia in-training examination and US Medical Licensing Examination Step 1 scores were significant predictors of success on the Basic Exam.
Extrapolation	A resident's score is a reasonable estimate of how much they know compared with what they need to know to practice clinical anesthesia safely. Implementation of the Basic Exam has improved patient safety during and after residency. Program leadership feels graduates are providing higher quality care than before the Basic Exam. The proportion of board-certified graduates has increased.	**None directly address the Basic Exam Re: ABA certification:** Zhou et al[18]: Physicians who passed both the written and oral board examinations, but not just the written examinations, had a lower likelihood of disciplinary action against their license. Baker et al[19]: In a multivariate analysis at a single institution in the traditional exam system, clinical performance scores could independently predict written and oral board performance. Hughes et al[20]: Program Directors were asked if they would permit each graduating resident to administer 3 increasingly complex anesthetics. Residents with Program Director confidence in all scenarios had a 75% pass rate on their first attempt on the applied exam.

continued on next page

continued from previous page

Tables *continued*

Implications	Decisions to delay graduation or dismiss residents who have not passed the Basic Exam improve patient safety. Remediation following unsuccessful attempts leads to objective improvement in performance and thus, patient safety. Stakeholders, including residents, training programs, and the public, perceive a benefit. Mean differences between subgroup and programmatic performance are not attributable to construct underrepresentation or construct-irrelevant variance. The cost of retaking the exam, extending residency, and dismissing residents on behalf of Basic Exam performance is acceptable to all stakeholders. Curricular resources dedicated to the Basic Exam have improved the educational experience for residents. Residents nationwide have equal opportunity to succeed, including but not limited to preparation resources, content experts, didactic curriculum, study strategy, time to prepare, and number of attempts.	**None directly address the Basic Exam** Murray and Boulet[21]: Editorial in response to Zhou et al.[2] The editorial addresses unanswered validity questions; they focus on the impact of the study patterns of residents, their clinical experience, the administration of the curriculum, how residents prepare for the Basic Exam, whether this additional study time and stress, and the impact of other learning activities. Pivalizza et al[5]: Clinical academic balance editorial evaluating the impact of unintended consequences in resident exam preparation priorities and resident attrition accountability. Markham et al[22]: As US Medical Licensing Examination Step 1 transitions to p/f, the authors argue for retaining some academic performance scores in anesthesia resident selection to predict Basic Exam and written exam success. Chen et al[23]: They focused on keyword-based didactics associated with an early trend toward improved In-Training Examination percentiles, and residents reported significant benefits to their confidence in knowledge of anesthesia topics, organization of study plans, willingness to educate others, and stress levels.

[a] For inferences in which there is no available evidence directly supporting the Basic Exam, we have included validation efforts for the American Board of Anesthesiology (ABA) certification series as a whole, mostly from publications preceding the Basic Exam, to serve as examples.

continued on next page

continued from previous page

Supplemental Online Material

Supplemental Online Material

Appendix A. Initial Interview Guide

Script for Verbal Informed Consent: This study was approved by the Institutional Review Board. With your permission, the session will be recorded to ensure that we do not miss any important details.

The University of Arkansas for Medical Sciences is interviewing current Anesthesiology Program Directors who have experienced at least one resident with a Basic Exam failure in the last three years to participate in a research study. The purpose of this study was to understand how Basic Exam results are used in individual training programs and to elucidate the implications of the Basic Exam from the perspective of the program director.

Participation in this study involved a one-time commitment of less than one hour for a video conference with the researcher for a semi-structured interview, and optional follow-up participation in a webinar to provide feedback on the themes generated from data analysis.

The researcher conducting your interview will keep your information and that of your institution confidential, but the audio data from the video conference will be transcribed and de-identified before the analysis phase. Participation is voluntary, and your medical care, student status, or employment status, will not change as a result of your decision. If you are uncomfortable being recorded, please let me know now so we can take more extensive notes (pauses for comments).

Would you voluntarily consent to continue with this recorded interview? Providing a response will be part of the ongoing consent process, and you can decline to answer or choose not to proceed at any point during the interview. For this study, there are no right or wrong answers; our goal is to gain a better understanding of the implications of the Basic Examination from your experience as a program director. The data collected in this study will be de-identified and aggregated to protect the identity of the program. I anticipate that this session will be completed within 60 minutes. In the future, you may be asked to provide further commentary on the thematic analysis of today's sessions via email or a webinar.

continued on next page

continued from previous page

Supplemental Online Material *continued*

1. Begin by sharing your name, role in the residency office, and the duration of that role.
2. Describe your anesthesiology residency program experience with the Basic Exam over the last three years.
3. Describe your program's policies around the basic exam.
    a. Can you provide an example of a resident who was affected by those policies?
    b. If not, has your residency experienced Basic Exam failure during your time as a PD?
        i. If not, has your residency experienced basic exam failures in the last five years?
4. Can you describe any past or upcoming changes in the program or curriculum that have resulted from the Basic Exam?
    a. Rotational changes? Curriculum changes? Recruitment changes?
        i. Authors' note: These topics were only used as prompts after the open-ended question, ensuring that we did not lead the participants' answers while still exploring relevant areas when necessary.
5. What inferences can be made about a resident who passes the Basic Exam?
6. What inferences can be made about a resident who fails the Basic Exam?
7. Some programs have noticed that residents change their behavior before the exam. Have you seen a change in resident behavior before the Basic Exam?
    a. Authors' note: Regarding behavioral changes, some participants reported no noticeable differences in resident behavior, which gave us confidence to explore this topic further using the interview question as written. This specific phrasing of the behavioral change question was posed to all participants to maintain consistency. Interestingly, studying was frequently mentioned in response to this question. Some participants observed residents becoming quieter, more stressed, and eager to leave the operating room.

continued on next page

continued from previous page

Supplemental Online Material *continued*

8. How does the Basic Exam impact clinical training? Resident performance?
9. Reflecting on your experience with the Basic Exam, what are your thoughts about the future?
10. Summary of key points provided, and clarification is sought.
11. Thank you for considering our manuscript. Do you have any closing thoughts about your session?

continued on next page

continued from previous page

Supplemental Online Material *continued*

Appendix B. Evolution of Interview Guide

Do you hold any positions that might influence your perspective on the basic exam?

1. Have you taken the Basic Exam?
2. What's your understanding of the purpose of the Basic Exam?
3. How does that compare to what it's actually doing?
4. At the SAAPM, the ABA assessment committee stated that the Program Directors had asked for a Basic Exam. Were you involved in the initial conversations? Vs. What do you think this means?
5. Walk me through your program's Basic Exam policies.
6. Tell me a story about a resident who failed the Basic Exam.
7. Have you ever had a failure who was struggling clinically?
8. Do you have a Basic Exam curriculum?
9. How does the Basic Exam influence recruitment?
10. Do applicants deserve to know the program's policies?
11. What inferences can be made about residents who pass or fail the Basic Exams?
12. What's the value added to the board certification process?
13. You can reform the system in any way you like or not at all. What, if anything, would you change?
14. This concludes my questions. What have I missed?

continued on next page

continued from previous page

Supplemental Online Material *continued*

Appendix C. Participant Sentiment Regarding Basic Examination Deduced From Suggested Revisions

Participant	Sentiment	Rationale	Suggested Revisions
1	Skeptical	Predictive validity	Revert to the traditional system
Quote: I'm just not sure if the basic exam is something that is really necessary for our residents to undertake. I'm sure it's hard for there to be data to demonstrate folks [who] do well on [the] basic exam versus folks [who] don't do well on [the] basic exam, do well clinically. But from my experience... the one resident [who] didn't pass the basic exam twice, clinically he was a strong resident.			
2	Skeptical	ROI; Value-added	Replace Basic with more frequent, low-stakes assessments: MOCA
Quote: So I think it's a problematic test. I think we would do better as a specialty to get rid of it as well as the standardized advanced exam...So I'm not sure there's a lot of added value to having the basic exam. I don't think that if residents are already feeling pressure to study for the ITE exam because it's going to influence their ability to get into fellowships, adding [an] even higher stakes exam that will influence their ability to graduate residency seems not helpful and only potentially harmful and adding stress.			
3	Skeptical	Predictive validity	Replace Basic with more frequent, low-stakes assessments: MOCA
Quote: I'd love for them to […] transition to a system we have as diplomats in terms of kind of more frequent, less high-stakes kind of application of knowledge. So, I don't know, yes, I'm sure there's [a] benefit to knowing the ion channels and how they're activated and things like that, but whether that's going to make the difference in terms of them being a good clinical anesthesiologist, Not quite sure. [More frequent, low-stakes exams] would actually provide more standardized or kind of more consistent longitudinal measure[s] of medical knowledge than just this one high-stakes Exam. […] Clearly, there's a lot of data that says it's probably not a great predictor other than their ability to then take a high-stakes exam in the future.			

continued on next page

continued from previous page

Supplemental Online Material *continued*

4	Neutral/undecided	Differentiation	Reevaluate Exam's Position in CA1 Year
Quote: I think it's important in the sense that it does force you to learn a subset of information that most people don't want to learn. I think on some level, that is important because again, that is what's going to set you apart from other people. However, I'm not so certain about the timing of when they're taking it. That being said, I'm not sure when you would move it to make it better. In some respects, I've never thought I would have a say, so I haven't really thought about what I would do. Honestly, I don't know. I think the advanced exam and the implied exams are important. I'm still hedging my bets on the basic. I don't know. It hasn't proven its utility to me yet.			
5	Supportive	Studying	No changes
Quote (In favor): No, not really. I don't think I would. I mean, full disclosure, [redacted: acknowledges bias] I mean I think it's, overall, I think it's a good decision. I do think, and I agree with the board, [that] the ABA's assessment is that it has probably caused or induced residents to do more serious reading or study, for lack of a better word, during the middle of their residency, rather than just pushing all of the things to the end, but it has certainly pushed some undesirable behavior like we were talking about right there in the middle of the residency, the expectation that we cover residents and that we relieve them and we don't have them on call.			
6	Skeptical	Predictive validity	Revert to the traditional exam system
Quote: I would get rid of it. It was so much better before when the residents would be able just to come and learn and spend three years with me and learn how to be great anesthesiologists. Then go for their written Exam.			
7	Supportive	Minimum standard	Address sources of "irritation" around the Basic Exam: Operational constraints & resident impact
Quote: And also having a role in the ABA with the BASIC Exam, that sometimes I like, that is a lot of work, a lot of work goes into that exam. Then sometimes it makes me question: why am I doing that?			

continued on next page

continued from previous page

Supplemental Online Material *continued*

What is the value in that? When you start to think about all of these other elements of it. But I guess I still think, in my core, my current perspective is needing to meet a certain minimum standard of medical knowledge, I guess. I just don't feel right about a pass-fail life career.			
8	Neutral/undecided	ROI; Predictive validity	
Quote: Again, this feeling that a fair number of these questions, the relevance is low to our clinical practice in some ways. So it's interesting where they would've come up with what these keywords were to include on the basic science part. Otherwise, I have to admit, I'm one of those pragmatists. And it is what it is, and I wasn't looking to change it in any way. I think it has [a] very low correlation, particularly the Basic, with their ability to care safely for patients. The Advanced Exam [has] slightly more correlation. But I would say that probably the OSCE and the SOE are our best attempt[s] at making a correlation.			
9	Skeptical	Predictive validity	Replace Basic with more frequent, low-stakes assessments: ITE
Quote: But it seems to me more like if the "weed out" would come [and] if they can't perform clinically, can't perform an emergency under stress, [and] have some professionalism issues, not that they can't remember the renal tubules of the kidney.			
10	Supportive	Weed Out	No 6-month extension after 3rd failure OSCE pre-graduation
Quote: The only thing is I would keep the Basic, I just don't understand the extension of training after the second [attempt].			
11	Skeptical	ROI	Revert to the traditional exam system
Quote: So, again, not knowing all the information, I don't know…I don't see a whole lot of benefits. I think our product was pretty good beforehand... it's a little oddly timed. It's at a time when it seems like people were learning a lot already. So, I don't know if it's the best timed. But yeah, it's a lot of logistical cost to orchestrate and program around.			
12	Supportive	Standard	Move the Basic Exam to PGY1 or the first half of the CA1 Year for better

continued on next page

continued from previous page

Supplemental Online Material *continued*

			alignment Offer the exam more frequently and let residents take it when they are ready Max number of attempts set by ABA
Quote: So my proposition would be, as soon as six months into the program, people could opt to take that Basic Exam, get it off the table and the distraction. If they failed it at that point, they could choose to take it as soon as a month later and just keep working on that. So they could have it knocked out by the end of CA1 year, almost certainly. Because I feel like if you fail it six times, you know what? Maybe then, yeah, it would be time to reevaluate your time in the profession. So say you took it at month six and then every month thereafter or something like that. And if you couldn't pass that Exam, well, that's a problem. But if you could get everybody to knock that out in the first six months, then that's great. Then that would serve [the] mission of the ABA to get people to learn that basic content, but not serve as a distraction long term within the training program.			
13	Supportive	Weed Out	Maintain current system; compress current system
Quote: I think the timeframe is really drawn out. If you look at the time you sit for the Basic to the time [of] your Applied Exam […], that is a really, really long timeframe. Years and years and years. I don't know as I'm saying this out loud what the negative ramifications would be, but if you could compress it and put it in closer together, closer to residency training, I think that would be beneficial.			
14	Skeptical	ROI	Replace Basic with more frequent, low-stakes assessments: MOCA
Quote: This is a really interesting question and I've got to say I love the MOCA minute pattern. To me, that seems really highly focused on learning it well, it seems to me [to] be a really great balance of assessment and learning. You've got to plunk down your quarter, you've got to answer the question. You get your score, you get your NDT. But then you also have this opportunity to review, review that question deepen your knowledge.			

continued on next page

continued from previous page

Supplemental Online Material *continued*

15	Skeptical	ROI; Predictive validity	n/a
Quote: No. To be very honest with you, I do not believe the ABA provides us with enough information to be able to assess from their perspective...They're the guardians of our ability to actually take care of the public.			
16	Skeptical	ROI; Predictive validity	Replace Basic with more frequent, low-stakes assessments: MOCA
Quote: I think clinical time made me a better doctor [and] has made me a better anesthesiologist. I think time with high-stakes patients has made me a better doctor. I think getting challenging consults has made me a better doctor. I don't know that the Basic Exam made… a better anything. Again, I think, yeah, again, honestly, again, I never thought I would say this because I was petrified before Oral Boards, but leaving there is like, "This is probably the most useful test I've ever taken. "			
17	Skeptical	ROI; Predictive validity	Revert to the traditional exam system
Quote: No: Nothing. That they're skillful in taking standardized exams [and] that they have a modicum of knowledge. I've never seen anyone that I'm aware of. I don't think anyone that has failed. . . I've never seen anyone not get a fellowship even if they failed one time or a job. So, I've never seen that. But I have seen confidence deteriorate and it's very distracting for the student.			
18	Skeptical	ROI; Predictive validity	Replace Basic with more frequent, low-stakes assessments: ITE
Quote: No: I would get rid of it... the lack of that data itself is enough for me to delete that exam. I would keep the ITE as the sort of […] litmus test for gaps and knowledge and all that stuff. And then it's up to the programs to create their own. Our program has a policy that you have to so score about a certain point in ITE, otherwise they were going to get an unsatisfactory well. I'm sure most programs have it... So, the Basic really doesn't add anything to us.			
19	Skeptical	ROI	Mini applied exams at the local level

continued on next page

continued from previous page

Supplemental Online Material *continued*

Quote: I think I will take away. What I will do, probably, instead of doing this written board exam, I have an idea and that is that end of the six-month evaluation of the residents a few skills that we are looking for. The ABA says, okay, [in] the first six month[s], every CA1 should be able to accomplish these attributes and skills. Every resident need[s] to be assessed on those skills by having a, let's say, not mark but OR, like one cheat sheet.			
20	Neutral/undecided	Standard	Evidence supporting correlation with clinical performance
Quote: Well, you've asked me some tough questions, and I wish I had better responses. I'm someone who's probably neutral [on] some of these points. I'm not impassioned to say, yes, the basics should go away, or yes, we have to keep it. I think refining the purpose or utility, it's good for programs to understand and PDs to know. It's crushing to see a resident not pass. I take it personally because it's majorly disruptive [to] their planning. It leads to, in a general sense, concerns that I have about their health, mental health. I don't know if one can make any profound inferences on that. I sort of wish, as a PD, that I could see the Exam questions for the Basic. What are they testing exactly? I have a sense of it in terms of basic facts and concepts, which I'm sure are critical to basic practice and understanding of some of the elements, I hope. I think. Right? But either we're not doing a good job as a program assessing the clinical strength of residents, or this exam doesn't correlate well with it, meaning one can be [clinical], at least thought of as a very strong resident: they perform well on cases, they manage patients well, they make good decisions, they show good judgment, but yet, why can't they pass an exam?			

continued on next page

continued from previous page

Supplemental Online Material *continued*

Appendix D. Final Codebook With Definitions

Code		Comment	Code Groups
1	ABA financial implications	Money exchange taking place to support resident enrollment in the board certification process and the ABA construction/delivery of the examinations	Interpretive Use Consequences Resident Implications
2	ABA policy	ABA's policy surrounding the basic exam: unlimited number of attempts, must be completed prior to graduation, 6-month extension after 3rd failed attempt, cannot matriculate to CA3 year.	Interpretive Use Consequences Equity Variations in Policy
3	Access test-taking resources	A program's (indirectly, a resident's) access to tools or personnel to support test-taking skills or preparation for any attempt at the basic exam	Consequences Equity Cycle of Failure Programmatic Implication Resident Implications
4	ACGME	Accreditation Council of Graduate Medical Education Program Accountability Program Citation Influence of ACGME policy on program (i.e., smaller programs impacted more) Influence of ACGME's policy on ABA or local policies surrounding the basic exam;	Purpose/Value-Added Consequences Variations in Policy
5	AKT	Merged from AKT and Anesthesia Knowledge Test	Purpose/Value-Added

continued from previous page

Supplemental Online Material *continued*

6	Alignment of exam content	Presence or absence of alignment between the Basic Exam as an assessment and the Teaching and Learning activities of the training program including clinical activities or didactic learning activities; Better engagement when a lecture is exam-focused; Distracting to clinical focus	Constructive Alignment
7	Assessment Bias	Merged from Assessment bias and Addressing bias	Purpose/Value-Added Consequences Equity Cycle of Failure Resident Implications
8	Backlash	A strong and adverse reaction expressed by many people surrounding the Basic Exam- Negative perception by a community, specific stakeholder group, or research; Program director/leadership or resident dissatisfaction	Consequences
9	Blindsided	Basic Exam results that were unexpected considering other exam/clinical performance results	Purpose/Value-Added
10	Clinical Performance implications	Refers to the direct or indirect, positive or negative implications of the Basic Exam on patient care; Distracting to clinical focus (i.e., doing questions in the OR); Knowledge acquisition facilitated by the Basic Exam which improves patient care	Consequences Cycle of Failure Programmatic Implication Resident Implications
11	Confidentiality	Pertaining to protecting the identity of residents who failed the basic exam	Consequences

continued on next page

continued from previous page

Supplemental Online Material *continued*

			Cycle of Failure Variations in Policy Resident Implications
13	Correlation to clinical performance	Correlation between residents' pass/fail results on Basic Exam and positive/negative clinical performance; Knowledge vs skills; Good test performance, poor clinical performance; scores not correlated to practice/performance; Disconnect from clinical practice	Purpose/Value-Added Constructive Alignment Consequences
14	Curriculum	Description of the program's didactic curriculum; includes a dedicated curriculum created to support basic exam performance	Consequences Equity Programmatic Implication
15	Early intervention	Remediation of a resident prior to 1st Basic Exam attempt in response to previous test performance including step 1, ITE, AKT. Co-regulated early intervention; Individualized education/learning plan as an early intervention or Individual support; Studying accountability (specifically pre-BE-failure); Meeting/discussion with leadership	Purpose/Value-Added Consequences Equity Variations in Policy Programmatic Implication
16	Efficiency of standardized exam	Refers to the feasibility of a testing format, acknowledging that selected-response questions are most easily constructed and administered	Purpose/Value-Added
17	Elbow grease	The anticipated early intervention work and resources associated with selecting an applicant with marginal standardized test-taking ability	Consequences Equity Programmatic Implication

continued on next page

continued from previous page

Supplemental Online Material *continued*

18	Exam construction	How and by whom the basic exam is constructed; Evolving exam content; Writing the exam; Separation of Basic and Advanced Content; Structure of exam (referring to which content is included and the weight assigned to individual content; similar to content determination); Content determination; ASA involvement	Constructive Alignment Equity
19	Exam timing	Reference to the June versus November opportunity to take the Basic Exam; Limitations imposed by the biannual opportunity to take the exam; Delay exam (foregoing the June exam in favor of 1st attempt in November)	Purpose/Value-Added Consequences Cycle of Failure Variations in Policy
20	Exception to policy	When a program deviates from its stated policy on behalf of an individual resident	Equity Cycle of Failure Variations in Policy
21	Faculty's ability to teach exam content	Refers to faculty's familiarity with Basic Exam and/or the ease with which a program's faculty can deliver lectures covering basic exam content; Faculty need to (re)learn content to teach it	Constructive Alignment Equity
22	Fellowship	Any reference to fellowship regarding the Basic Exam. Use of ITE scores by fellowships is NOT included in this code (see ITE).	Consequences Equity Variations in Policy
23	High stakes	Stakes of Basic Exam including implications of not passing from the programmatic perspective for residents	Consequences Equity

continued on next page

continued from previous page

Supplemental Online Material *continued*

			Cycle of Failure Variations in Policy
	Subcodes	Subcodes: Balance with clinical prep; board eligibility; graduation/dismissal; extension of training; fellowship; milestone adjustment; unsatisfactory; promotion to CA3 year; stereotype threat; test anxiety; training without credit; wellness; work/life balance	Consequences Equity Cycle of Failure Variations in Policy
24	Implications of failure	Merged from the impact of failure on residents, implications of failure, and implications of not achieving board certification	Consequences Equity Cycle of Failure Variations in Policy
25	Inferences	Inferences that can be made about a resident on behalf of Basic Exam performance.	Purpose/Value-Added Constructive Alignment Equity
26	ITE	Anything pertaining to the ITE including fellowship or ITE use for risk stratification; ITE performance's correlation Basic/other exams; ITE stakes (use by fellowships or for probationary purposes in local policy); ITE role in residency	Purpose/Value-Added Consequences Variations in Policy
27	Knowledge retention	Refers to whether the information tested on the Basic exam is retained by residents or serves a role as the foundational basis for future (advanced knowledge)	Purpose/Value-Added
28	Mixed response	Mixed sentiment expressed surrounding the Basic Exam	Purpose/Value-Added

continued on next page

continued from previous page

Supplemental Online Material *continued*

39	Neuropsychiatric evaluation	Neuropsychological testing of a resident struggling with exams, with or without medical diagnosis and with or without test accommodations, medication, or cognitive behavioral therapy as a result. Medical diagnoses including ADD, ADHD, Dyslexia	Consequences Variations in Policy
30	Other specialties	Comparison of the ABA board certification series to other specialties including those with and without high-stakes exams DURING training	Consequences
31	Participant feelings	Code to capture some latent themes regarding PD sentiment about residents who have experienced a failure. The participant's internal dialogue of sorts. Examples from transcript 12: "I really felt like his executioner at the time"; "Ugh. I just felt awful about that [dismissal]"; "So I'm happy with the end result of that story, but if you had asked me at the time, and even for years afterward, I felt awful about that."	Consequences
32	Participant's personal experience with Basic Exam	Refers to whether the interview participant or program leadership has personally taken the exam as a resident or faculty (as part of a standard setting study); Implications of a participant or program leadership's familiarity (or lack thereof) with the exam; Program leadership has/has not taken exam	Purpose/Value-Added Interpretive Use
33	Pre-basic subspecialty rotations	Interplay of basic exam and subspecialty rotation timing; Waiting until after Basic Exam to	Constructive Alignment Consequences

continued on next page

continued from previous page

Supplemental Online Material *continued*

		introduce subspecialties; Subspecialty rotations before Basic Exam	Equity Programmatic Implication Resident Implications
34	Program experience	Regarding failures or top 10% on the Basic Exam	Cycle of Failure
35	Program policy	The current or historical policy pertaining to the Basic Exam for the participant's institution (locally)	Consequences Equity Cycle of Failure Variations in Policy
36	Program size impact	Impact of any exam in the ABA board certification series that has a different or disproportionate impact on small-sized anesthesiology training programs	Consequences Equity Programmatic Implication
37	Purpose of Basic Exam	Purpose of the basic exam as understood by the participant; Value of the basic exam; Clarity or lack of clarity surrounding the purpose of the Basic Exam	Purpose/Value-Added Interpretive Use
38	Purpose of board certification	Provides public confidence; Proves competence; Public perception of qualifications; Minimum standard; Osteopathic certification; Differentiation between CRNAs/Dos/AAs; Rigor	Purpose/Value-Added Constructive Alignment Interpretive Use
39	Remediation	Term describing the learning intervention that takes place after the first or subsequent failure of the Basic Exam; Co-regulated remediation: remediation via faculty or peer mentor; Individualized education/learning plan; Studying	Equity Cycle of Failure Variations in Policy

continued on next page

continued from previous page

Supplemental Online Material *continued*

		accountability (specifically post-BE-failure); Meeting/discussion with leadership	
40	Research	Basic exam's impact on resident engagement in research. Example: "Okay. Yes, you must know a topic for the QI project, but the charter isn't due until after the Basic. You don't have to do it before the Basic."	Purpose/Value-Added Consequences Programmatic Implication Resident Implications
41	Resident Recruitment	Absence or presence of an influence exerted by the basic exam policy (local or national) on the resident selection process; broadly used to code resident recruitment. May also refer to an applicant's perception of a program; Cannot take a chance/gamble on selecting a resident who will not be able to pass the basic exam (unwilling or unable to provide elbow grease); Selection criteria	Consequences Equity Variations in Policy
42	Resident resignation	Resident is given the option to resign after any number of Basic Exam failures; "Soft" termination from the training program; Exit residents from the program in a way other than frankly firing them; Counseled out of the program	Consequences Equity Cycle of Failure Variations in Policy
43	Resident story	Response to tell me a story of a resident affected by policies. Can refer to any resident experience as told by the PD—including success, failure, and dismissal—as they pertain to an individual resident	Purpose/Value-Added Consequences Equity Cycle of Failure Variations in Policy

continued on next page

continued from previous page

Supplemental Online Material *continued*

44	Schedule changes post-failure	Scheduling modifications supporting basic exam preparation or in accommodation of the exam itself, taking place AFTER a failing exam result	Consequences Equity Cycle of Failure Variations in Policy Programmatic Implication
45	Schedule changes pre-Basic	Scheduling modifications supporting basic exam preparation or in accommodation of the exam itself, taking place prior to the first/June attempt of the exam; Scheduling night float; Relief from clinical duties to prepare	Consequences Equity Cycle of Failure Programmatic Implication
46	Scoring	How the Basic Exam is scored and reported; Percentile does not matter; Pass/Fail; Standard setting; transparency in pass/fail cut-off; Commendation for top 10%	Purpose/Value-Added Equity Cycle of Failure
47	Self-fulfilling prophecy	The phenomenon whereby a person's or a group's expectation for the behavior of another person or group serves to bring about the prophesied or expected behavior; Other-imposed prophecy may have some overlap with stereotype threat whereby resident performance will mirror the stereotype assigned to them by others	Purpose/Value-Added Cycle of Failure
48	Stress	Stress or adverse emotions experience pre-basic exam attempt #1; Detriment/negative impact on resident wellness or balance; Depression; Anxiety	Purpose/Value-Added Consequences Equity Cycle of Failure Resident Implications

continued on next page

continued from previous page

Supplemental Online Material *continued*

49	Studying	In reference to the stated purpose of the exam; Longitudinal studying/preparation; Question-bank preparation; Question-banks are not enough; Cramming; Proactive studying	Purpose/Value-Added Consequences Resident Implications
50	Suggested Revisions	Merged from CBME and Suggested Revisions	Purpose/Value-Added Constructive Alignment Equity
51	Supporting evidence	presence or absence of supporting/refuting data to show effectiveness or support use of the Basic Exam	Purpose/Value-Added Interpretive Use
52	Test anxiety	a type of performance anxiety — a feeling someone might have in a situation where performance really counts on the Basic Exam (regardless of attempt number or previous failure)	Purpose/Value-Added Consequences Equity Cycle of Failure
53	Test-taking ability	A resident's historical performance on multiple-choice, high-stakes exams. Merged from Exam history and Test-taking ability	Purpose/Value-Added Consequences Equity Cycle of Failure
54	Transparency about policies	Refers to the program's practices around sharing/or not sharing their local basic exam policies with applicants PRIOR to match day; also refers to the program's communication or conveyance of historical basic exam performance to applicants	Consequences Equity Variations in Policy

continued on next page

continued from previous page

Supplemental Online Material *continued*

55	UME	Refers to Step 1, Step 2, or Step 3; shelf exams; or the context of preparation for these exams	Purpose/Value-Added Consequences Equity
56	URM students	Any aspect of the Basic Exam, UME exam performance (historical test-taking ability), influence on resident selection with specific mention of the impact on URM students; Disparities in opportunities for URM students/residents; Marginalized students/populations/residents	Purpose/Value-Added Consequences Equity