Conceptualizing and Identifying Emotions in Anti-Harassment Education for Medical Residents

Byunghoon Ahn, MEd, BSc
PhD Candidate, Department of Surgical and Interventional Sciences
McGill University, Montreal
June 03, 2024

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Doctor of Philosophy

Table of Contents

Acknowledgements

I would like to thank everyone involved in producing this dissertation. Special thanks to my supervisor, Prof. Jason Harley. An idiom from my culture, 師弟同行 (사제동행) comes to mind—a master and pupil journeys together, and with unified vision, collaboration, and research they pursue higher knowledge. I am thankful for their continued dedication, patience, and support, throughout not only my doctoral program, but the majority of my academic journey so far, back from my undergraduate years. I am also thankful to my committee members, Prof. Susanne Lajoie, and Dr. Kevin Lachapelle; along with my research advisory committee chair, Prof. David Labbé for their continued feedback and support to my work. I would like to thank Niki Solis for providing helpful feedback to the simulation. I would also like to thank Dr. Ning-Zi Sun for helping to integrate our training into medical residents' curriculum. Thanks to Leah Moss, PhD, for her feedback towards our educational intervention as well. Special thanks to Caroline White and her team at MUHC-i-SIM center for running our simulation smoothly. Another special thanks to my lab mates who were indispensable in conducting our studies: Myriam Johnson and Negar Matin. I am of course thankful to all the various colleagues, coauthors, staff members at various worksites, and any other supporters I have come across during my years at McGill.

I would also like to thank others who were not involved in producing this dissertation. Thanks, Mr. O'Connor, for introducing design to me. Thanks, Ellen Kim, for your indescribable faith. Thanks, Tom Johnson, and Shannon Scott, for your guidance. Thanks, Sang-hyeok Lee, for being the inspirational Unkillable Demon King. Thanks, Hyun-cheol Jeong, for being relatable. Thanks, Bang-geol Ahn, for the bloodline.

Statement of Support

This dissertation was financially supported by the McGill University Department of Surgery's Graduate Excellence Award from Prof Harley's start-up package and the Social Sciences and Humanities Research Council of Canada (SSHRC). The latter is comprised of a SSHRC grant awarded to my supervisor which in turn supported me (Grant ID: 430-2020-00573) and a doctoral fellowship awarded to me (Application number: 752-2022-1899). My dissertation was also supported by the involvement of staff and use of the facility of the McGill University Health Centre's Interprofessional Simulation Centre.

The views presented in this dissertation are that of the authors alone and do not necessarily reflect those of the McGill University Department of Surgery or SSHRC.

Thesis Format

This dissertation uses the manuscript-based format and includes 3 manuscripts that have either been published or are under review for publication. The first manuscript is a scoping review study has been published in *Advances in Simulation*. The second manuscript is a quasi-experimental study that has been submitted to *Technology, Knowledge and Learning*. The third manuscript is an observational study that has been submitted to *Journal of Computer Assisted Learning*.

This dissertation starts with an introduction as the first chapter and concludes with an overall discussion (Chapter 5) and conclusion (Chapter 6). Each of the manuscript chapters (Chapters 2, 3, and 4) have been formatted to meet their respective journal submission requirements. Note that these chapters have their own reference lists, tables, figures, and supplementary information. The master reference at the end of the dissertation only pertain to Chapter 1, 5, and 6.

Contributions to Original Knowledge

My dissertation contributes original knowledge in the following ways by investigating: 1) how emotions are conceptualized and measured in simulation-based education research studies; 2) the emotional experience of anti-harassment training for medical trainees, specifically, for video-based instructions and simulation-based training; and 3) the potential roles of emotions in anti-harassment education for medical residents. The work presented in my dissertation is entirely original and primarily conducted by me. I have undertaken this research independently, with support and guidance provided by my dissertation committee, co-authors, and my supervisor.

Author Contributions

The following states author contributions for each published manuscript.

For article, Ahn, B., Maurice-Ventouris, M., Bilgic, E., Yang, A., Lau, C. H.-H., Peters, H., Li, K., Chang-Ou, D., & Harley, J. M. (2023). A scoping review of emotions and related constructs in simulation-based education research articles. *Advances in Simulation*, *8*(1), 22:

- Ahn B: Contributed to conceptualization and design of the research project, development of search strategy, develop of task protocols and data extraction sheet, documentations, article screening, data extraction, data analysis and visualization, literature review, manuscript writing, hand searching for articles, revisions, and approval for submission.
- Maurice-Ventouris M: Contributed to conceptualization and design of the research project, develop of search strategy, development of the data extraction sheet, development of task protocols, article screening, data extraction, file management, and approval for submission.
- Bilgic E: Contributed to the development of data extraction sheet, article screening, hand searching, and approval for submission.
- Yang A: Contributed to the development of search strategy, development of the data extraction sheet, article screening, hand searching for articles, and approval for submission.
- Lau H-H: Contributed to the development of the data extraction sheet, article screening, and approval for submission. H.P contributed to the development of the data extraction sheet, article screening, and approval for submission.
- Li K: Contributed to the development of search strategy, development of the data extraction sheet, article screening, hand searching for articles, file management, and approval for submission.
- Chang-Ou D: Contributed to the development of search strategy, development of the data extraction sheet, article screening, hand searching for articles, file management, and approval for submission.
- Harley J.M.: contributed to the conceptualization and design of the research project, development of search strategy, development of task protocols, and data extraction sheet, providing guidance on data analysis, visualization, literature review, manuscript writing, revision, and submission of the article.

For both articles, Ahn, B., Johnson, M., Matin, N., Sun, N-Z., & Harley, J. M. (Under Review). Anti-Harassment Education for Internal Medicine Residents: Learning Activity Sequence, Productive Failure, and Emotions. *Technology, Knowledge and Learning*

And, Ahn, B., Johnson, M., Matin, N., Sun, N-Z., & Harley, J. M. (Under Review). Multimodal Cluster Analysis of Medical Residents' Emotions During High-Fidelity Harassment Bystander Simulation. *Journal of Computer Assisted Learning*

- Ahn B: Contributed to study design, lead instructional development, instructional delivery (simulation training facilitation), data collection, data preparation, data analysis, lead manuscript writing.
- Johnson M: Instructional delivery (simulation training standardized participation), data collection, consultation on qualitative methods, data preparation, support manuscript writing.
- Matin N: Instructional delivery (simulation training standardized participation), data collection, consultation on psychophysiological analysis, support manuscript writing.
- Sun N-Z: Study design, curriculum integration, instructional development, data collection planning
- Harley J: Funding acquisition, supervision, study design, instructional development, data collection planning, support manuscript writing, data analysis

In addition, for the above article, Multimodal Cluster Analysis of Medical Residents' Emotions During High-Fidelity Harassment Bystander Simulation, there is an additional author:

- Lee SY: Data preparation, data analysis, support manuscript writing.

While I have received support and guidance from my dissertation committee and supervisor on study designs, analyses, and writing, I have made the greatest contribution in all intellectual aspects of these manuscripts.

List of Tables and Figures

Chapter 2 – How Emotions are Conceptualized in Measured in Medical Education Research

Chapter 3 – Learning and Simulation Performance in an Anti-Harassment Intervention: Instructional Design and Emotions

Chapter 4 – Multimodal Cluster Analysis of Medical Residents' Emotions During High-Fidelity Harassment Bystander Simulation

List of Abbreviations

CVT	Control Value Theory of Achievement Emotions
EDA	Electrodermal activity
EI	Emotional intelligence
KLI framework	Knowledge Learning Instruction framework
MOFEQ	Multiple object foci emotion questionnaire
PRISMA-ScR	Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews
SBE	Simulation based education
Sim-first	Simulation-first
SP	Standardized participants
TRE	Technology-Rich Environment
Vid-first	Video first

Abstract

Emotions are multicomponent psychological constructs that are expressed in multiple channels, including through affect, physiological, and behavioural processes. While this conceptualization has been embraced in the field of traditional educational psychology, it is unclear whether such adoption is the norm in medical education. Such adoption and pursuit of investigating the emotions of learners is particularly important in understudied subject domains such as anti-harassment education. The importance stems from emotions' role in fostering or hindering achievement, and therefore the need for researchers and educators to provide emotions-aware educational experience for learners. This dissertation identifies the current gaps in simulation-based education regarding conceptualizing and measuring emotions. It then examines the effectiveness of an anti-harassment educational intervention I served as the lead developer on. In relation to the effectiveness, I examine the emotional experiences of medical residents, and employ multiple approaches to analyzing emotions, including a unimodal analysis using self-report data and a multimodal analysis using self-report, electrodermal activity, and simulation debriefing speech content data.

The first manuscript my dissertation features highlights the lack of formal integration of theories in simulation-based education when examining emotions. It identifies simulation educators focusing on anxiety and relying primarily on unimodal analysis, potentially, a symptom of a lack of emotions theories, such as Pekrun's Control Value Theory of Achievement Emotions. The second manuscript highlights the importance of activity sequences in educational interventions, specifically, in terms of the effect the sequence of activities can have on learning outcomes and simulation performance. Further, emotions can be linked to different activities.

Positive emotions were linked to higher knowledge levels while negative emotions were linked to lower knowledge levels. The last manuscript highlights the affordances that a multimodal approach to emotion analysis offers. Examples include the use of a cluster analysis that relies on both self-report and skin conductance data, as well as leveraging simulation debriefing speech content which provides additional discrete emotions omitted by other measures, and insight on potential emotions object-foci. The results of the study also highlights the potential impact of emotions in how residents may intervene in harassment, including how those with lower activation levels may better confront perpetrators. Future directions include examining other related populations including other resident populations, such as general surgery, or other healthcare workers such as nurses. Further, other emotions data channels such as heartrate or voice, along with examining emotion regulation processes medical residents use once specific emotions are elicited, also serve as future directions.

Résumé

Les émotions sont des phénomènes psychologiques à plusieurs composantes qui s'expriment via de multiples perspectives, notamment à travers l'affect et des processus physiologiques et comportementaux. Bien que cette conceptualisation fut adoptée dans le domaine de la psychologie pédagogique traditionnelle, il reste incertain si elle aussi adapté à l'enseignement médical. Une telle application et recherche sur les émotions des apprenants sont particulièrement importantes dans des domaines sous-étudiés tels que la formation contre le harcèlement. Cela est particulièrement important du au rôle des émotions qui peuvent servir d'un levier ou, au contraire, devenir un frein dans la réussite, soulignant la nécessité pour les chercheurs et les éducateurs d'offrir aux apprenants une formation incluant les émotions. Cette thèse identifie les lacunes actuelles dans la conception de simulation éducative et la mesure des émotions durant ces dernières. Elle examine également l'efficacité d'une intervention éducative contre le harcèlement, laquelle je fus le développeur principal. Concernant l'efficacité, j'explore les expériences émotionnelles de résidents en médecine en utilisant diverses approche pour analyser les émotions, y compris une analyse unimodale utilisant des auto-évaluation et une analyse multimodale utilisant des auto-évaluation, l'activité électrodermique et des enregistrements vocaux d'entrevues suivant les simulations.

Le premier manuscrit présenté dans ma thèse met en évidence le manque d'intégration fondamentale de théories en pédagogie par la simulation lors de l'étude des émotions. Il démontre que les éducateurs en simulation se concentrant sur l'anxiété, s'appuient principalement sur une analyse unimodale, potentiellement du à un manque de théories sur les émotions, telles que la théorie Contrôle-Valeur des émotions d'accomplissement de Pekrun. Le

deuxième manuscrit met en évidence l'importance l'ordre d'activités dans les interventions éducatives, notamment en termes de leurs effets sur les résultats d'apprentissage et performances durant une simulation. De plus, les émotions peuvent être liées à différentes activités. Les émotions positives furent liées à des niveaux de connaissances plus élevés, tandis que les émotions négatives, à des niveaux de connaissances inférieurs. Le dernier manuscrit met en évidence les avantages de l'approche multimodale dans l'analyse des émotions. Cela inclut l'utilisation d'une analyse groupée s'appuyant sur les données d'auto-évaluation et de conductance cutanée, l'utilisation d'enregistrements vocaux des entrevues suivant la simulation qui fournit des émotions discrètes supplémentaires omises par d'autres mesures, mais offre également un aperçu des émotions dérivées. Il met aussi en évidence le possible impact des émotions sur la manière dont les apprenants pourraient intervenir en cas de harcèlement, incluant comment ceux possédant des niveaux d'activation sont inférieurs peuvent mieux affronter les agresseurs. Les recherches futures pourraient se concentrer sur d'autres populations médicales, y compris les chirurgiens généraux ou d'autres travailleurs de la santé tels que les infirmières. De plus, d'autres canaux de données sur les émotions, comme la fréquence cardiaque ou la voix, ainsi que l'examen des processus de régulation des émotions utilisés par les résidents en médecine une fois que des émotions spécifiques sont suscitées, servent également d'orientations futures.

Chapter 1: Introduction and Literature Review

The primary problem of interest this dissertation examines is the role of emotions in learning and performance, specifically, in the subject domain of anti-harassment training within the context of medical resident simulation education. I begin with an introduction to how emotions have been studied so far in the general field of education, as well as introduce the context of anti-harassment training in medicine. The chapter will outline why emotions matter in education by explaining what the field already knows about the potential impact of emotions in learning, as well as why studying emotions in anti-harassment matters due to the emotionally charged nature of harassment and its overall negative impacts to trainees.

1.1 Prior Research on Emotions in Education

There are two major characteristics of past emotions research in the context of learning. First, there was a heavy focus on studying anxiety, as evidenced by the ratio of studies focusing on anxiety to other emotions surpassing 100 to 1 during the mid 1970s to 1990.[1] This focus stemmed from identifying anxiety as a key barrier to academic success, as anxiety could lead to diminished working memory capacity and disengagement to learning activities.[2,3] Aligned with the heavy emphasis on anxiety, another characteristic of past emotions research is the view that emotions are undesirable in learning. Emotions were thought to be something that can take away precious cognitive resources and therefore serve as a distraction.[1,4] Further, emotions were often characterized as something that opposed being rational[5], with "thinking" solely tied to cognitive processes and free from any emotions.[6,7] To borrow a quote featured in McNaughton's article[8] that described the roots of how emotions were conceptualized in (western) medical education, emotions were viewed as "a corruption of reason that needs to be transcended".[9(p.184)]

However, notable advancements in emotions research have taken place since. First, the calls to studying more diverse kinds of emotions have been answered.[10] This took form in two ways, where researchers could focus on more diverse sets of discrete emotions other than anxiety, such as anger[11], boredom[12], enjoyment[13], and confusion[14]. Researchers could also study a specific type or subtypes of emotions, such as achievement emotions and related emotions[15,16] (e.g., epistemic[17], social[18,19], topic emotions[20], etc.). Researchers now therefore recognize that not only anxiety, but a whole array of emotions can have significant roles in education.

Further, emotion research has expanded outside of math education, with studies examining emotions in various contexts and subject domains. Research has embraced various contexts of learning, including multimedia[21] or technology-rich environments,[22] distance or online learning settings,[23,24] and even informal learning contexts.[25,26] While studies of emotions in math education remains popular[27], other subject domains frequently studied include biology, chemistry, computer science, engineering, physics, literacy, and health sciences.[28,29] With heavy emphasis on STEM subject domains, the field still calls for examining less studied subject domains such as history.[30,31]

Lastly, emotions research has begun to incorporate various measurement approaches other than self-report questionnaires or surveys.[32] Such approaches can involve analyzing speech, facial expressions, posture, physiology, and brain activities.[33] In addition, studies have begun to integrate multiple of these measurements in chorus, therefore relying on multimodal analyses for studying emotions.[34,35] Studies that take on diverse data channels related to emotions mean they can overcome challenges faced by unimodal analysis, as each approach to detecting emotions offers unique advantages and disadvantages.[36] Considering that most traditional

approaches to emotion research relied on self-report measures, such self-report studies would have faced various limitations. These limitations include participants potentially inaccurately recalling emotions, having different interpretations of emotional terms featured in the self-report measures, facing reluctance to disclosing certain emotions due to social desirability, and the possibility of filling out the self-report measures themselves influencing the emotions the participants experienced.[32] The trend in emotions researchers advocating for multimodal analyses aim to remedy such shortcomings each approach to measuring emotions have.

As my dissertation focuses on utilizing Pekrun's Control Value Theory (CVT) of Achievement Emotions, specifically in the subject domain of anti-harassment training with an emphasis in simulation-based education (SBE), the brief overview of past emotions research I have highlighted contextualizes where in the emotions research landscape my dissertation lies. It lies on the shoulders of giants, where past scholars have: 1) formulated relevant theoretical framework of emotions in educational settings (i.e., CVT); 2) expanded emotions research beyond a narrow focus of anxiety (i.e., achievement and other related emotions); 3) have called for studying emotions in various contexts, including understudied subject domains (i.e., SBE, health professions education); and 4) how to measure emotions, with a particular emphasis on multimodal methods.

The following sections expand on the theoretical framework I primarily relied on, and the subject domain and learning context I have focused on in my dissertation.

1.2 Control-Value Theory of Achievement Emotions

1.2.1 Defining Emotions

Before I discuss achievement emotions, I must operationalize emotions in general. While there are numerous competing theories for conceptualizing emotions[37,38], the following are characteristics of emotions most theorists can agree on.[39] First, emotions are caused by a specific stimulus. This notion is agreed upon by major emotion theories: 1) Basic emotion theories posit that our evolution has hardwired us to experience certain emotions in response to specific stimuli[37,40,41]; 2) Appraisal theories posit that our emotional responses depend on cognitive evaluations of these stimuli[42,43]; and lastly; 3) Psychological construction theories posit that emotions are constructions arising from a combination of psychological, neurological, and social processes triggered by specific events[44,45]. All of these major theories agree that the prerequisite for an emotional experience is the presence of a stimulus that one can focus on.

Second, emotions have multiple components. Emotions are multifaceted, encompassing affective, cognitive, physiological, motivation, and behavioural components.[46] Therefore, when someone (e.g., me) is learning the intricate art of surgical knot tying for the first time, they may feel frustration (affective), think about why the knot is not being made properly (cognitive), start to have sweaty palms (physiological), desire to throw away the task trainer and give up (motivation), and bite their lips (behavioural). While the order, the exact interactions, and implications of these components may differ between theories, emotions are not confined to mere felt experiences.[39]

Third, emotional episodes follow a sequential process. Theories differentiate the specific processes, but the general process involves perceiving a stimulus, appraising it, and experiencing emotions as a response. The entire process is thought to last from a few seconds to minutes[39].

Some have reported emotions being able to last for much longer, especially according to Verduyn's line of work[47,48]. Depending on the severity of the event, the intensity of the emotions, and personal traits of the person experiencing the emotions, emotional episodes can last for hours. Ruminating on a negative, intense, emotion-eliciting event could lead to experiencing the same emotional state for longer, for example.[49] However it is worth noting that ruminating may mean one constantly re-engaging in a stimulus that causes emotions—therefore it stands to reason the finding of certain emotional episodes lasting longer can be attributed to the brief sequential processes being triggered repeatedly to form the long duration.

And finally, emotions are distinct from other affective phenomena. Other related constructs to emotions such as mood, personality, attitudes, reflexes, drives, and stress responses do not adhere to all three of the characteristics outlined above. For example, moods lack an event focus and can persist over much longer periods than emotions.[48] Stress responses are seen as unspecified negative states that are evoked specifically when an individual evaluates the demands of a situation as unmanageable.[50]

In sum, emotions are "multi-componential psychological responses, that are generated towards object foci of situations perceived to be relevant to one's goals",[51(p.2)] a conceptualization congruent with contemporary emotion theories and Pekrun's CVT.[52–54]

1.2.2 Why Control Value Theory?

As stated in the prior section, the definition of emotion I described above is applicable to most theories across the spectrum of basic emotions, appraisal emotions, and psycho-constructivist emotions theories. This sub-section clarifies why this dissertation adopts the CVT

over other frameworks. This explanation requires a brief survey of the three camps of the emotions.

First, basic emotion theories propose that there are a set of biologically hardwired emotions that are universally recognized and have distinct neural and physiological signatures.[37] These theories are advantageous because they are grounded in Darwin's theory of evolution, lending support from evolutionary science and related empirical work. They also provide clear, distinct labels for emotions, suggesting that complex emotions (e.g., jealousy) are products of basic emotions. However, basic emotion theories face criticism due to inconsistent findings regarding specific neural circuits or biological processes supporting discrete emotions.[37] Additionally, they might oversimplify the spectrum of emotional experiences and fail to acknowledge individual variability in emotions, such as those influenced by culture or personal background.[43]

Appraisal theories focus on how emotions are formed through individual evaluations and interpretations of events.[37] These theories take context and individuality into account, emphasizing cognitive processes like evaluation and interpretation, which align well with psychological research on perception and thought. The advantages of appraisal theories include their ability to account for contextual and individual differences in emotional responses. However, they are often criticized for lacking clarity about the specific mechanisms through which appraisals lead to emotional responses.[37,42] Additionally, the varying definitions of appraisals across different theories and the challenges in measuring and linking appraisals to specific emotions due to their subjective nature are significant drawbacks.[42,55]

Finally, psycho-constructivist theories propose that emotions are constructed in the moment from more basic psychological processes.[37] For instance, Barrett's Theory of Constructed Emotion suggests that core affect, conceptualization, and language interact dynamically to create emotional experiences.[45] These theories emphasize the variability and context-dependence of emotional experiences, aligning with empirical research showing that emotions are not fixed but constructed dynamically.[45]

As evolutionary theories, they are supported by studies on the neural and psychological mechanisms leading to emotion construction, though they do not support the universality of emotions. However, the complexity of psycho-constructivist theories makes them difficult to test empirically, particularly due to their focus on dynamic interactions and context-dependent constructions.[37] Additionally, these theories are less intuitive and harder to communicate to researchers outside the psycho-constructivist domain.

The CVT belongs to the camp of appraisal theories. Appraisal theories (as well as psycho-constructivist theories) can integrate both dimensional models and discrete (distinct) emotion models. That is, while the CVT relies on a continuous space defined by multiple underlying dimensions that form a three-dimensional taxonomy (i.e., valence, arousal, object foci), it also utilizes specific terms (e.g., anger, sadness, etc.) to convey more discrete, tighter defined ideas of specific emotional experience.[54] Accordingly, CVT includes propositions for both discrete emotions and emotional quadrants (e.g., positively-valenced, high-activating). Both dimensional models and models of distinct emotions have drawbacks. The former, while able to capture a large range of emotions, lacks specificity and potentially lumps emotions with distinct characteristics into one category; the latter models may suffer issues pertaining to limits of

linguistic labels (e.g., anger vs. rage vs. fury vs. wrath), especially in cross-cultural contexts (e.g., anger vs. гнев vs. রাগ vs. öfke).[56] Additionally, there are statistical power considerations and limitations to examining a large set of discrete emotions. As an appraisal theory, the CVT utilizes both dimensions and specific discrete emotions labels to provide a taxonomy of emotions to offer a comprehensive organization of emotions.

Other than the advantages appraisal theory brings (e.g., taking account of various contextual factors, including individual differences), the CVT emphasizes object foci—or exactly *what* is being appraised to elicit certain emotions.[54] In an academic context, this can be particularly relevant as the same emotions may have differing impacts on learning if they are sourced from different object foci (i.e., prospective, concurrent, retrospective). The CVT was also crafted with such academic contexts in mind. Many other emotion theories are meant for general purposes, and lack the specific lens the CVT provides in terms of the relationship appraisal processes and emotions can have with learning outcomes or other learning related psychological processes. Relatedly, there is a wealth of empirical evidence that have been collected with CVT-informed research, with various measurement tools with robust evidence of validity.[10,54]

With these advantages in mind, the CVT was chosen for my dissertation. The CVT, of course, has limitations to consider. It lacks accounting for interpersonal interactions and emotions of multiple learners. It also lacks focus on constructivist principles which emphasizes the interaction of experiences and the environment for emotions to emerge. The latter limitation may be particularly pertinent when one is to study psychosocial constructs such as culture or personal identity in how emotions form. However, given that my dissertation does not focus on

the interplay of complex psychological processes such as cultural orientation in how emotions are formed, the CVT remains the most relevant framework for my research.

1.2.3 Achievement Emotions

According to the CVT, achievement emotions are "emotions tied directly to achievement activities or achievement outcomes".[54(p.317)] In an anti-harassment training context, achievement activities would include reviewing harassment policies, attending lectures, and taking part in role-play exercises. Feeling curiosity when discovering the various reporting process outlined in policy documents would be an example of an activity emotion, while feeling pride when reflecting at how the harassment was prevented in the role-play exercise would be a case of an outcome emotion.

Achievement emotions are elicited through appraisal of control and value associated with the stimulus.[54,57] Specifically, this refers to the degree of control one feels over the situation, and the amount of importance one places on the activity or outcome. CVT proposes that in general, high control is associated with positive emotions, while high value is linked to more intense emotions.[58] The elicited emotions in turn, can be categorized by a three-dimensional taxonomy: valence (negative to positive), activation (deactivating to activating; also referred to as arousal), and object focus (retrospective outcome, activity, to prospective outcome). Therefore, feeling anxious when actively listening to harassment experiences in a workshop would be categorized by CVT as a negative-activating activity emotion. In the context where the object focus is not clear, it is the norm to rely on the first two dimensions (valence and activation), to describe an emotion. For example, anger would be called a negative-activating emotion.

1.2.4 The Role of Emotions in Learning

The CVT outlines ways emotions can impact learning by influencing various processes, including those related to cognition, motivation, and self-regulation.[54] As appraisal is a prerequisite for eliciting emotions, experiencing achievement emotions demand cognitive processes. Learning in itself requires cognitive resources (e.g., working memory, attention), and while positive emotions such as enjoyment may foster dedicating appropriate resources towards learning, other emotions such as anxiety of failing may impede such processes; effectively being a distraction to the learner.

Emotions can impact motivational processes as well.[57] While positive activating emotions should foster motivation, deactivating negative emotions from low-control (e.g., hopelessness) or low-value (e.g., boredom) appraisals can dampen interest and engagement in achievement activities. Positive deactivating emotions (e.g., relief) may initially lower engagement but foster long-term commitment, while activating negative emotions (e.g., anxiety) may undermine intrinsic motivation while potentially boosting extrinsic motivation to avoid failure.

Positive emotions, such as enjoyment, are believed to enhance self-regulated learning, whereas negative emotions, like anxiety, tend to encourage reliance on external regulation.[54] A general sense of joy in learning is likely to foster active goal setting, planning, monitoring, and evaluation by learners themselves. In contrast, emotions of despair or frustration might drive individuals to seek these processes externally. Overall, positive activating emotions are presumed to exert a greater influence on self-regulated learning compared to positive deactivating emotions.

In a nutshell, positive emotions, especially positive activating emotions are deemed as helpful in learning. Positive deactivating and negative activating emotions may have mixed effects. Negative activating emotions such as anger especially may have potential in learning, provided that the emotion is resolved during the learning process.[28,31] Negative deactivating emotions such as hopelessness is thought to be almost always detrimental to learning.

1.3 Simulation Based Education and Emotions

Simulation-Based Education (SBE) has emerged as a pivotal methodology in training healthcare professionals, offering a safe, controlled environment where learners can practice clinical skills, decision-making, and interprofessional collaboration.[59,60] I specifically subscribe to Gaba's definition of simulation, where it is defined as a "technique—not a technology—to replace or amplify real experiences with guided experiences that evoke or replicate substantial aspects of the real world in a fully interactive manner".[61(p.i2)] Simulations therefore aim to replicate experiences a learner would face in real situations using various technological tools, called simulators (e.g., mannikins, task trainers, virtual reality, etc.).

SBE in medical training is designed to replicate not just the physical aspects of real-life clinical scenarios, but also their functional and psychological dimensions.[62,63] Therefore, SBE also includes providing learners with authentic emotions for learners.[64,65] Medical trainees frequently encounter emotionally charged situations[66], such as handling difficult clinical encounters,[67] delivering bad news to patients[68], and even confronting patient death.[69] Learning scenarios involving simulations therefore is filled with emotions for researchers to examine.

Medical educators have varied opinions on the roles of emotions. A recent review from LeBlanc and Posner (2022) have summarized this issue,[70] noting that some believe that

activating emotions regardless of valence can lead to better learning, as high intensity states should translate into better memory retention and retrieval under similar situations. Others have focused on anxiety, in particular, and how knowledge acquired under stressful conditions is desirable and is something learners should become adapted to. The review highlights how these views are often anecdotal and based on personal experiences. There has been a call to anchor how medical educators understand emotions to theories such as the CVT[7], and some contemporary research seems to show such integration.[71,72] In addition to LeBlanc's line of work, others have further highlighted the importance of emotions in medical education and SBE,[73–75] echoing the CVT's description of how emotions can influence learning. Specifically, positive emotions such as enjoyment and relaxation typically fosters learning. While anxiety and other negative emotions are predominant in SBE, they can enhance learning outcomes in some contexts.

1.4 Harassment Education in Medicine

Harassment can be defined as repeated unwanted or unwelcome repeated behaviour that harms a person's dignity or physical integrity, and results in a harmful environment.[76–79] Despite the effort to address harassment, recent articles continue to cover incidents and problems of harassment, with medical education now having accumulated more than 40 years' worth of scholarly articles relating to harassment.[80,81] Silver's commentaries[81,82] were one of the earliest to talk about student mistreatment, where he:

> […] pointed out the striking parallels between the changes that occurred in children who had been abused and those of medical students during the weeks and months after they matriculated. He suggested that medical students, like abused

> children, exhibited alterations in attitude and behaviour that might be the result of avoidable, unnecessary, and harmful abuse that often went unrecognized or, if recognized, was largely ignored.[81(p.527)]

Despite this perception made in the earliest days of formally documenting harassment, I believe Silver's quote above accurately outlines the core issues at hand. Medical students and trainees like children under guardianships can often be helpless in front of their preceptors and seniors; harassment can be abusive, it should be avoidable, but is often ignored, and like any abusive acts, can have devastating consequences.

Recent research has highlighted not only a lack of improvement in addressing harassment, but also a strong association between harassment and negative outcomes such as trainee burnout, suicidal ideation, and diminished academic and career achievement.[83–90] While medical trainees seemingly know that harassment is wrong[91], barriers to effectively addressing it includes the trainees perceiving bullying and humiliation as ingrained aspects of training,[80,92] and the hierarchical nature of medical education.[93–95]

Typical efforts in harassment education are characterized by passive, didactic strategies such as presentations and handouts. Evidence of their effectiveness is limited, particularly due to their tendency to provoke defensiveness and backlash, lack of inclusivity, or failure to sustain attention from learners.[96,97] Contemporary research has examined more innovative approaches such as facilitated discussions[98] and video modules[99]. However, these studies tend to focus on the *perceived* usefulness, feasibility, or effectiveness of these educational interventions,[98–103] as opposed to examining concrete knowledge gains, or behavioural change. Further, while scarce, handful of efforts that integrate SBE into anti-harassment education, most are conference

proceedings that lack much details on methodologies and discussion.[104–106] Relatedly, empirical study articles that feature SBE tend to reflect "black box" research, as it does not investigate the underlying mechanisms and processes involved in learning and dealing with harassment, but simply the learning outcomes.[107] While researchers call for SBE to be integrated into anti-harassment education,[108,109] such endeavors also require incorporating theoretical frameworks to understand the factors for effective education.[110]

Therefore, the current gaps in harassment education research in medicine include: 1) lack of investigating innovative educational interventions; 2) primarily relying on measuring self-perceived usefulness or effectiveness of the training; and 3) lack of investigation in the underlying learning processes. While other factors other than emotions can be pertinent to learning outcomes (e.g., learning activity sequencing, cognitive load, motivation, etc.) emotions are one prominent factor not yet explored in the context of anti-harassment education in medicine. Considering that harassment situations can be deeply emotional,[111] the field of medical education and SBE warrants investigation into emotions of anti-harassment training.

1.5 Dissertation Objectives

The primary objective of this dissertation is to explore and analyze the intersections of emotions, SBE, and harassment education in the context of medical education. It also features an anti-harassment educational intervention I served as lead developer on under the supervision of my supervisor. Serving as the lead developer of the anti-harassment simulation and coordinator for the larger SSHRC-funded project my dissertation is an important part of this project and has several contributions. Primarily, this included the creation of video modules where I supplied much of the text content, made all illustrations and animations from scratch, recorded and edited

the videos, and provided my own voice for narration. The educational intervention also featured a simulation training component. I led a small team of assistants, and we chose an appropriate clinical procedure for our target population, documented and implemented best practices for developing simulation training, planned all materials needed including personnel, scripted the simulation, trained the involved standardized participants (i.e., actors), and developed the corresponding debriefing session and assessment tools. Finally, I also served as the simulation facilitator to all the trainees featured in this dissertation.

Considering the pivotal role emotions can play in the learning process, and with the hopes of bringing uplifting change to the issue of harassment in medical education, my dissertation set out to do the following:

1. Systematically chart how emotions are studied in medical education, specifically in the SBE context. How are emotions and related constructs conceptualized in SBE research?
2. Test the effectiveness of the anti-harassment intervention. What was responsible for learning and simulation performance in the involved medical residents? Were emotions involved?
3. Provide a rich description of what kind of emotional experience educators and researchers can expect from anti-harassment SBE. What are the typical emotions medical residents experience as they try to address harassment in a simulation setting? Do these emotions have any impact in residents' simulation performance?

To achieve these goals, my dissertation contains the following chapters:

1. Chapter 2 features a scoping review study that examined how SBE researchers have conceptualized and measured emotions and related constructs: mood, stress, and emotional intelligence.
2. Chapter 3 features a quasi-experimental study that examined the effectiveness of the anti-harassment educational intervention. It further investigated the impact of activity sequencing, and how emotions were related to residents' learning and simulation performance.
3. Chapter 4 features an observational study that employed cluster analysis based on multimodal emotions data to formulate emotional profiles of learners. It further investigated whether these emotional profiles impacted simulation performance.

In summary this dissertation should serve as an example of theory-driven research approaches of emotions in an understudied subject domain, utilizing contemporary methodologies in investigating the potential roles of emotions in anti-harassment education.

Chapter 2 Bridging Text

In Chapter 1, I provided a definition of emotions, and the importance of understanding its role in health professions education. Emotions are distinct from similar, but related processes such as stress and mood. It is unclear whether SBE researchers make such distinctions, and whether they infer emotions from other measured constructs. It is also unclear how SBE researchers tend to measure emotions of their learners.

In chapter 2, our work centered on synthesizing knowledge through a scoping review. Our focus was on examining the conceptualization and measurement of emotions in SBE. We aimed to identify existing gaps and suggest future research directions for SBE, drawing on the collective insights and findings of previous SBE educators.

Chapter 2 - A Scoping Review of Emotions and Related Constructs in Simulation Based Education Research Articles

Authors: Byunghoon (Tony) Ahn[1], Megane Maurice-Ventouris[1], Elif Bilgic[1,2,3], Alison Yang[1], Clarissa Hin-Hei Lau[1], Hannah Peters[1], Kexin Li[1], Deuscies Chang-Ou[1], Jason M. Harley[1,4,5,6]

[1]*Department of Surgery, McGill University, Montreal, Canada*
[2]*Department of Pediatrics, McMaster University, Hamilton, Canada*
[3]*McMaster Education Research Innovation and Theory (MERIT) Program*
[4]*Research Institute of the McGill University Health Centre*
[5]*Institute for Health Sciences Education, McGill University*
[6]*Steinberg Centre for Simulation and Interactive Learning*

Institution: Department of Surgical and Interventional Sciences (Formerly Department of Surgery)

Advances in Simulation. 2023 Sept; 8(22):1-21 doi:10.1186/s41077-023-00258-z

Correspondence for Reprints to:
Jason M. Harley, PhD
jason.harley@mcgill.ca

Parts of this work have also been presented and/or published as:

Ahn, B.T., Maurice-Ventouris, M., Bilgic, E., Yang, A., Lau, C.H., Peters, H., Li, K., Chang-Ou, D., & Harley, J. (2022 May). *A Scoping Review of Emotions in Medical Simulation Training*. Poster presented at Royal College of Physicians and Surgeons of Canada's Simulation Summit Conference. Virtual Conference.

Ahn, B.T., Maurice-Ventouris, M., Bilgic, E., Yang, A., Lau, C.H., Peters, H., Li, K., Chang-Ou, D., & Harley, J. (2022, May). *Mapping Out Conceptualizations of Emotions in Simulation-Based Education: A Scoping Review*. Presentation presented at McGill University Department of Surgery's Experimental Surgery Research Day. Montreal.

Abstract

Background: While acknowledgement of emotions' importance in simulation-based education is emerging, there are concerns regarding how education researchers understand the concept of emotions for them to deliberately incorporate emotionally-charged scenarios into simulation-based education. This concern is highlighted especially in the context of medical education often lacking strong theoretical integration. To map out how current simulation-based education literature conceptualises emotion, we conducted a scoping review on how emotions and closely related constructs (e.g., stress, and emotional intelligence) are conceptualised in simulation-based education articles that feature medical students, residents, and fellows.

Methods: The scoping review was based on articles published in the last decade identified through database searches (EMBASE and Medline) and hand-searched articles. Data extraction included the constructs featured in the articles, their definitions, instruments used, and the types of emotions captured. Only empirical articles were included (e.g., no review or opinion articles). Data were charted via descriptive analyses.

Results: A total of 141 articles were reviewed. Stress was featured in 88 of the articles, while emotions and emotional intelligence were highlighted in 45 and 34 articles respectively. Conceptualizations of emotions lacked integration of theory. Measurements of emotions mostly relied on self-reports while stress was often measured via physiological and self-report measurements. Negative emotions such as anxiety were sometimes seen as interchangeable with the term stress. No inferences were made about specific emotions of participants from their emotional intelligence.

Conclusions: Our scoping review illustrates that learners in simulation-based education are most often anxious and fearful. However, this is partially due to medical education prioritizing measuring negative emotions. Further theoretical integration when examining emotions and stress may help broaden the scope towards other kinds of emotions and better conceptualizations of their impact. We call for simulation education researchers to reflect on how they understand emotions, and whether their understanding may neglect any specific aspect of affective experiences their simulation participants may have.

Keywords

Emotion, Stress, Emotional Intelligence, Simulation Based Education, Medical Education, Surgical Education, Scoping Review

Introduction

Training for delivering better patient outcomes requires understanding emotions—they are pervasive throughout healthcare environments for both physicians and trainees[1]. Moreover, recent medical education research has highlighted the roles and potential impact of emotions in training future doctors.[2,3] Not only can emotionally-charged scenarios sway clinical decisions[4], but being able to understand, monitor, and manage one's own and others' emotions (i.e., emotional intelligence) is deemed crucial for developing core competencies physicians require[5] as it relates to desirable aptitudes such as resiliency against burnout[6], leadership[7], and communication[8,9]. In sum, medical education can help prepare learners to be emotionally resilient and better emotionally attuned to one another and their patients. Therefore, advancing these goals and informing practice requires a better understanding of medical trainees' emotions during medical education, including simulations.

Simulation-based education (SBE) is no exception to emotions' pervasiveness: a recent conceptual review by LeBlanc and Posner (2022) highlighted how SBE can be "rife with emotional situations"[10,(p.6)]. Indeed, the ubiquitous nature of emotions in medicine, including in SBE has been echoed previously[1,11] and its potential impact on healthcare practice and education has been empirically documented in various domains, including diagnostic reasoning[12], patient-physician communication[13], and patient safety issues related to medical errors[14,15].

LeBlanc and Posner's review[10] also underscored how simulation educators' various beliefs regarding the role of emotions are often based on personal experiences—one concern regarding this is whether educators can consistently be "thoughtful and deliberate"[10,(p.5)] when incorporating emotional experiences into SBE scenarios. This can be especially so when

simulation educators conceptualise and use the term emotions in everyday and professional discourse. For example, McNaughton's discourse analysis (2013) shows how educators can view emotions not only as a "series of biological and neurochemical responses", but also as "skills that can be learned" (i.e., emotional intelligence; EI), and "a set of practices that are constructed by social, cultural, and political arrangements"[11,(p.73)]. To put the concern another way: are medical educators understanding emotions as something interchangeable to EI? Will they distinguish something like stress (also a product of "series of biological and neurochemical responses") from emotions? To truly be purposeful in incorporating emotions in SBE, educators should be aware of how they conceptualise emotions, and why they subscribe to their understanding.

Integrating a theoretical framework into research is a straightforward way of advancing understanding of emotions (e.g., see Coppin & Sander[16]). While medical education research often lacks strong theoretical integration[17,18], there have been calls for integrating frameworks such as the Control Value Theory of Achievement Emotions (CVT)[1,2,19], a prominent theoretical framework in educational psychology[20,21]. CVT's definition of emotion aligns with the consensus mainstream emotion theories and its definition can help educators distinguish emotions from other affective phenomena including moods and stress[19,22]. CVT defines emotions as multi-componential psychological responses produced by coordinated affective, cognitive, motivational, and expressive processes. Emotions are described based on a three-dimensional taxonomy: valence (negative or positive), activation (deactivating or activating, also known as arousal), and object focus (retrospective outcome, concurrent activity, or prospective outcome). Research indicates that positive-activating emotions (e.g., enjoyment) should favor learning outcomes, while other types of emotions tend to hinder learning outcomes, especially negative-

deactivating emotions such as hopelessness[23]. While CVT has been utilised in contemporary research in medical education[12,24], it is unclear how widespread incorporating such frameworks into SBE research and simulation design is.

To help address the concern of whether simulation educators can consistently be deliberate in incorporating emotional experiences into SBE scenarios, a scoping review aiming to distill how simulation educators understand and study emotions was conducted. The review examined not only emotions but also closely related constructs: mood, EI, and stress.

The Peripheral Constructs to Emotions: Mood, Stress, and Emotional Intelligence

Previous research reports that terms such as mood and stress are closely related constructs to emotions, although they are not interchangeable terms and can be differentiated[22,25]. EI, while not used interchangeably with emotions in the psychology or educational psychology literature, is often viewed by medical educators as one way of conceptualizing emotions[11].We first identified mood as a construct of interest for our review, as we thought it may be possible for researchers to mistakenly use the term interchangeably with emotions, despite the general consensus from emotion researchers[22].

We identified stress as another construct of interest. Stress is "inextricably linked to anxiety"[26(p.4)], where such discrete negative emotions are the consequence of a stress-response[25,26]. In other words, where there is stress response, one would expect negative emotions[25]. In addition, like emotions, stress has traditionally been hard to define[27], and there was concern regarding what approach recent SBE research would take. Given the uncertainty of how stress would be conceptualised, and with stress having a close relationship with negative

emotions, we reasoned that studies that examined stress held potential to infer or directly measure emotions.

We lastly considered EI. EI is deemed a highly desirable trait in medical trainees; often measured as emotional quotient (EQ) through instruments such as Bar-On's Emotional Quotient Inventory (EQ-i)[28]. EI is associated with leadership skills, non-technical skills, reduced stress/burnout, higher job satisfaction, and better relationships with patients[29]. While EI is not an affective phenomena and moreover a trait rather than a state, it directly concerns recognizing and regulating emotions[30]. Therefore, with the popularity of examining EI in SBE literature, coupled with its relationship recognizing and managing emotions, we were interested in how the literature approached EI. Our interest included whether any emotional experiences would be inferred from EI measurements.

Objectives and Research Questions

The objective of this scoping review was to map out how recent SBE research approached studying emotions. Our primary research question was the following: How are emotions and closely related constructs (i.e., stress, emotional intelligence, and mood) conceptualised in articles that focus on simulation environments, with medical students and trainees as the population of interest?

We formulated sets of complementary secondary questions (SQ) that align with our objective:

- (SQ1): What is known about how emotions are conceptualised?
    - (SQ1-A) Are emotions defined?
    - (SQ1-B) Do the articles cite a theory?

    - (SQ1-C) How are closely related constructs conceptualised?
- (SQ2): What is known about how emotions are measured?
    - (SQ2-A) How are other closely related constructs measured?
- (SQ3): What are the emotions medical students and trainees experienced?
- (SQ4): What kinds of emotions did medical students and trainees experience in studies that measured their stress levels?
- (SQ5): Can we infer the kinds of emotions medical trainees and students experienced in studies that measured their emotional intelligence?

Methods

Design

We designed and conducted our scoping review based on Arksey and O'Malley's[31] methodological framework. We further consulted guidelines that elaborate on this framework[32], and a librarian with expertise in scoping reviews and health sciences education. Lastly, we referred to the Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews (PRISMA-ScR) checklist[33]. See Appendix A for review protocol details, and Appendix B for the completed PRISMA-ScR checklist. A scoping review was chosen over a systematic review as we set out to identify the types of knowledge and research the current literature offers, and clarify key concepts—a systematic review would be better suited to assess the quality of current practices and can be a natural progression after a scoping review is conducted first[31,34]. Our review of the literature indicates that there has yet to be a scoping review for this topic. Therefore, this knowledge synthesis work appropriately takes place before any systematic reviews.

Stage 1: Identifying the Research Questions

We formulated our primary question based on our primary research objective: to map out how recent SBE research approached studying emotions. We also considered which specific aspects our review should focus on[31], namely which additional constructs related to emotions our scope would cover. There are numerous constructs that are closely related to emotions, and while they can be differentiated from emotions, they have components and features that overlap with emotions[22,35]. Therefore, we believed there may be a range of how these terms are used and applied in research settings, warranting their inclusion.

As per our methodological framework, we refined our research question through an iterative process to balance out the vastness of our scope and the relevancy that our identified articles would yield. This included consulting our librarian and deciding to focus on SBE (versus general medical education) to enhance the review's focus and quality. We iterated through the guidelines from Arksey and O'Malley[31] and ensured our adjusted scope aligned with our research questions and search strategy.

Eligibility Criteria for Constructs Related to Emotions

In terms of eligibility criteria for specific constructs, we consulted a health sciences education librarian to explore related MeSH terms and the existing literature to focus on key constructs. This allowed us to exclude less relevant terms towards our research objective such as "social stress" or terms related to clinical disorders (e.g., mood disorders).

We also decided to focus on medical students and trainees, meaning we only included studies that featured medical students, interns, residents, and fellows. If a study featured other population groups (e.g., pre-med students, nurses, attending physicians), it was only included if it

featured our target population. In addition, as we were interested in the context of SBE, we sought studies that had such environments. Hence, studies that asked for general life satisfaction or a survey that asked for emotions concerning day-to-day activities were excluded. When dealing with simulations, we were mindful of whether the simulation content would be related to medical knowledge, procedure, or non-technical skills. For example, simulations of war combat via a video game would not be included in our review as the simulated content is not something we would expect a medical trainee to experience (even if there were educational components featured in the simulation).

Through the above considerations and process of exploring what our scope would be, we were able to finalise the primary research question to focus on emotions, mood, stress, and EI of medical students and trainees in a SBE context. After trials identifying the number of articles we would yield, the team agreed our scope was broad enough to adhere to our research objective.

Stage 2: Identifying Studies

We developed a search strategy with our librarian's consultation. An iterative development process led to identifying a list of relevant keywords and MeSH terms (Table 1). Our search was carried out on *Medline* and *EMBASE* on June 22nd, 2020, exclusively looking at articles published from June 22nd, 2010 to June 22nd 2020. The initial yield of 29,329 articles decreased to 19,508 after deduplication.

We chose electronic databases and decided to complement them with hand-searched articles as outlined in Arksey & O'Malley's guidelines[31]. We chose EMBASE and Medline based on our previous experience conducting scoping reviews related to research in the fields of medical education[36]. We selected Academic Medicine and Medical Education to hand-search

articles because key articles were often featured there (e.g., Artino and Pekrun[2], Pottier and colleagues[37]). Hand searching is a supplementary technique in a scoping review and not intended to be comprehensive.

Stage 3: Study Selection

For our database searches, there were three screening processes. Table 2 shows our inclusion and exclusion criteria for the title-abstract screening. Tables 3 and 4 show the criteria for the first and second full-text screening, respectively. While the first full-text screening included all medical education scenarios, the second full-text screening specifically filtered for simulation-based studies as per the iterated process outlined in Stage 1. Gaba's definition of simulation was used to help screen the articles, where simulation was defined as "a technique—not a technology—to replace or amplify real experiences with guided experiences that evoke or replicate substantial aspects of the real world in a fully interactive manner"[38(p.i2)]. We referred to Gaba's article that describes various dimensions of simulations to operationalise different criteria and types of simulators.

Each of the screening processes involved a team of reviewers (six for title-abstract screening, four for both full-text screenings), where each article was screened by a pair. A pilot preceded every screening process to ensure 75+% interrater reliability before proceeding[39]. Re-calibrations took place after 1/3 and 2/3 of the articles were screened.

In addition to searching electronic databases, screening of hand-searched articles involved four reviewers. These articles (37) were screened for SBE content, similarly to our full-text screenings. After deduplication, two articles were added (please see Figure 1 for the flow diagram).

Stage 4: Data Charting and Reporting

We finalised our data extraction sheet and strategy through consultation of multiple resources, our librarian, and pilot testing. Two reviewers were chosen in the end to extract the articles to ensure high interrater reliability. See Appendix C for details.

The final data extraction sheet included three categories of data: 1) information pertaining to the publication of the article (e.g., author, year, etc.); 2) information about the constructs of interest (e.g., whether the article focused on emotions or closely-related constructs, what measurements were used, etc.); and 3) information on the simulation based on Gaba's dimensions of simulation applications[38].

Stage 5: Collating, Summarizing, and Reporting the Results

We analysed our data descriptively; we summarised general characteristics of the papers through various charts to provide a broad overview of study characteristics. We report our summary to answer our research questions while providing context.

Results

The database search (139) combined with hand searching (2) yielded a total of 141 articles that met the inclusion criteria for review.

Study Characteristics

Our results showed that research in emotions and related constructs have been increasing over the years, with most of the articles published in western countries (e.g., United States, United Kingdom). The vast majority (83.0%; 117 articles) were quantitative studies, while the bulk of the populations featured were medical students (56.7%; 80 articles) and residents (43.4%; 61 articles).

Stress was featured the most in the articles we identified (62.4%; 88 articles), followed by emotions (31.9%; 45 articles) and EI (24.1%; 34 articles). Only 1 article (0.7%) featured mood. Appendix D contains more details about study characteristics, including simulation characteristics.

Conceptualization, Measurement, and Experiences of Emotions

Our results indicated that most of the articles that focused on emotions did not formally define emotions: 35 (77.8%) of the 45 articles on emotions[43–76]. Seven (15.6%) articles[9,77–82] defined a type or a discrete emotion (e.g., anxiety), 2[83,84] (4.4%) provided formal definitions, while 1[85] defined both a formal definition for the term emotion, and separate definitions for discrete emotions (e.g., anxiety). For specific types of emotions, achievement emotions were the only type identified (as opposed to other types such as epistemic emotions—emotions that relate to knowledge and generation of knowledge[86]). Definitions of discrete emotions focused mainly on negatively valanced emotions such as anxiety and embarrassment. Amongst the articles that formally defined emotions, the circumplex model of emotion was referenced once and CVT was referenced four times.

Figure 2 shows that self-report measures (34 articles) were the most common method of collecting data on emotions. Besides custom self-made instruments[46,47,50,53,58,59,66,72,77,87], the State-Trait Anxiety Inventory[43,44,63,67–70,82] (STAI; or a variation of it) was the most employed instrument (featured 8 times in emotions-focused articles). The next most common instruments were the Achievement Emotions Questionnaire[77,80,85] (AEQ), and a scale based on Barrett and Russel's work on emotions[51,56,79] (i.e., the circumplex model[88]), each featured 3 times. See Appendix E for more details on the wide array of other instruments used in studies. Figure 3

shows that anxiety and fear (captured in 25 and 10 articles[47–50,53,55,58,60,62,67], respectively) were the most commonly measured emotions across the 45 articles that focused on emotions. Few positive emotions were captured in the studies, with excitement and enjoyment being the most frequently measured ones, but only being mentioned in four[53,56,59,83] and five articles[56,60,74,77,80], respectively.

Conceptualizations and Measurements of Other Constructs

64 of the 88 articles that focused on stress (72.4%) did not formally define stress with an explicit reference to a theoretical framework. Sixteen (18.2%) articles[37,68,69,69,75,82,89–99] did, however, formally define what stress was, while 6 articles[62,100–104] (6.8%) defined related terms to stress such as distress and stressors. Lastly, 2 articles[105,106] (2.3%) defined both stress and related terms. For the 16 papers that formally defined stress, 8[68,69,75,82,89,95–97] relied on conceptualization of stress stemming from a physiological-oriented (physiogenic) approach (e.g., Selhye's General Adaptation Syndrome[107]), while 6 of the papers[37,69,93,94,98,99] from a psychological-oriented (psychogenic) approach (i.e., Lazarus' Transactional Model[108]). Overall, 3 papers[82,90,105] relied on definitions that drew from multiple approaches. It should be noted that 1 paper that explicitly defined stress did not provide any references (however took a physiological-oriented approach to defining stress)[91].

Figure 4 shows the measurements that stress articles employed. Studies using both self-reports and physiological measures were the most common (33 articles[37,68,69,69,70,75,76,90,93,94,97,99,101,104,109–128]; 37.5%), followed by articles solely relying on either self-reports alone (30 articles [59,63,66,67,71,73,74,87,95,98,102,105,106,129–145]; 34.1%) or physiological measures alone (19 articles[58,65,72,82,89,91,92,96,146–156]; 21.6%). There were two articles[103,157] (2.3%)

that used behavioural analysis in addition to self-reports and physiological measures. Heart rate or heart rate variability related instruments were the most common (32 articles[68,70,72,82,89–93,97,99,101,103,109–113,113,114,116,117,120–123,127,146,148,151,154,155,157]; 36.4%) for physiological measures. For self-report measures, the STAI was the most common (20 articles[37,67,70,75,93,97,101,103,106,106,109,111–113,116,117,119,121,126,127]; 22.7%).

Nine EI articles[7,9,101,158–163] (26.5%) formally defined EI. Thirteen articles[60,63,100,164–173] (38.2%) defined a construct directly related to, or a subordinate construct of EI (e.g., empathy, emotional skills). The most common measurement in these articles was the Jefferson Scale of Physician Empathy[174] (JSE; 6 articles[165,169,170,172,173,175]; 17.6%), followed by The Trait Emotional Intelligence Questionnaire[176] (TEIQue; 4 articles[101,161,168,177]; 11.8%) and The Mayer-Salovey-Caruso Emotional Intelligence Test[178] (MSCEIT; 3 articles[9,162,163]; 8.8%). The EI articles exclusively focused on EI did not infer emotions from their measures.

The one article that focused on mood[179] utilised the Visual Analogue Scales (VAS) from Bond and Lader[180] to assess "high" (positive) and "low" (negative) mood, and hence did not measure any discrete mood or emotions. It did not formally define mood or emotions.

Discussion

Primary Question

Our primary research question was: "How are emotions and closely related constructs (i.e., stress, emotional intelligence, and mood) conceptualised in articles that focus on simulation environments, with medical students and trainees as the population of interest?" Based on our results, the short answer is that the SBE literature concerning emotions and related constructs tend to be light on theory, relying on previous research findings to orient their research. The

literature had an affinity towards focusing on negative emotions (e.g., anxiety), and made connections between negative emotions and stress. Our review also further acknowledged the importance of these psychological constructs in education—hence acknowledging the desirability of EI in trainees (although no emotions are inferred from students' EI levels). The more elaborate version of our answer was obtained through answering the secondary questions we proposed.

Secondary Questions

Conceptualization and Measurements of Emotions

SBE literature often discusses emotions informally, without explicit definitions or references. This was evident as most emotions-focused articles (35/45) did not provide a definition for emotions. We believe this may partially be due to the studies' scopes and intentions, which are shaped by norms and trends in SBE. While a definition by itself is not a theory, it serves as an important part of a psychological theory. As these articles did not formally cite any theories either, we concluded that they lacked a theoretical-based conceptualization of emotions. Based on our results, we believe that recent work in the SBE literature resembles the early emotion research activities in the realm of traditional education research in that it focuses on specific discrete negative-activating emotions and relies on self-report measures[181].

Many studies concentrated on negative emotions. For example, Kim's study[77] looked at how medical students' anxiety and boredom in class were associated to levels of anxiety in Objective Structured Clinical Examinations (OSCEs). Fraser and colleagues'[51] focused on negative emotions elicited from simulated patient death. It was therefore not surprising that measurements towards emotions largely included self-reports (34 articles of 45 that focused on

emotions) that focused on negative emotions such as anxiety (8 articles using STAI[182] to measure this) and other negative emotions such as fear (measured via the Fear of Negative Evaluation Scale[183]) and shame (measured via Experimental Shame Scale[184]). We noted that studies that did not rely on self-reports also followed the trend of identifying negative emotions in SBE[48,55,61,62].

The emphasis on negative emotions is understandable. This is especially so when we consider how medical trainees face training scenarios that are *meant* to elicit negative emotions, due to the nature of the topic the scenarios deal with, and the pressure for achieving high performance. For example, articles such as Groot's article[61] featured medical students taking on residency-level simulations that featured emergency room acute care cases (e.g., anaphylaxis, acute myocardial infarction, etc.). These simulations placed the students in an emotionally-charged situation where negative emotions such as anxiety (i.e., stress) would be elicited, in part, due to the advanced and challenging nature of the medical issues they needed to manage. Students strived to perform well but ultimately reported stress and disappointment in themselves failing to meet their own expectations. Our other reviewed articles follow this example in terms of dealing with emotionally-charged topics and scenarios likely to elicit negative emotions. Bloomfield and others' article[54] featured students communicating with dying patients and their family members. Bauer and others' article[82] had residents for their first time be given a high-fidelity mannikin in a scenario where the simulated patient was intubated during an intra-hospital transportation, but had oxygen desaturation, and was under mechanical ventilation. Summing up these examples, we report that the typical SBE scenario in our review featured stress-inducing, negative emotion eliciting experiences—explaining why the literature perhaps tends to focus on negative emotions.

Our results showed that four articles referenced the CVT to define emotions, with three using related measures (i.e., Achievement Emotion Questionnaire; AEQ[185]). While focusing on specific emotions such as anxiety has merit, extending the acknowledgement of emotions' role beyond a specific discrete emotion such as anxiety, or a type of emotion would benefit the SBE literature by providing a more comprehensive picture of the potential role other emotions play. Broad examples of this include studying emotions during simulation versus after simulation (debriefing); effects of positive-activating emotions (enjoyment) versus positive-deactivating emotions (relief) and so on. We make a positive note that CVT—one of the most suitable frameworks for being applied to SBE research—was cited by several of the few articles that *did* rely on theory. This suggests that SBE researchers are on the right track in conceptualizing emotions in academic contexts.

While we highlighted the need to look beyond negative emotions in SBE research, studies such as Butteris and colleague's[53] illuminates the rationale for focusing on negative emotions. Unlike the rule of thumb that negative emotions are generally undesirable in education, the study's facilitators' consensus seemed to be that completing a simulation scenario involving a neonatal death or caring for a simulated HIV-positive toddler requires trainees to experience negative emotions so that they are motivated to reflect on their competency and preparedness[53]. We note that the study emphasised post-simulation debriefings to help trainees adaptively digest these negative emotions. Though the non-profession specific educational (e.g., high school, higher education) emotion literature acknowledges that emotions such as anger may be beneficial in specific contexts[23,186], the contexts featured are difficult to compare to what medical students and trainees experience.

Conceptualizations and Measurements of Related Constructs to Emotions

Like emotions, most articles focusing on stress (64/88) did not offer formal definitions. Instead, they referenced previous work and expert opinions to establish research directions. Further, papers featuring stress tended to simply conceptualise stress from a biological, physiological approach. The assumption may be that readers of medical education journals do not expect formal explanations of stress, as they are familiar with the biological components of stress responses (e.g., activation of the hypothalamic-pituitary-adrenal and the sympathetic-adrenal-medullary axes). Hence, many authors simply mention this physiological side of stress in lieu of citing a theoretical framework. We note that authors who measured constructs such as anxiety as an indicator of stress were measuring a construct that is different from stress, albeit related[187].

Measurements of stress-related articles were similar to emotion-related articles in that they often employed instruments such as the STAI[182]. In other words, articles that focused on stress, while not studying a wide variety of emotions, still often examined negative-activating emotions such as anxiety. This is supported by our findings that of the 18 articles that examined both emotions and stress, 16 featured STAI for measuring anxiety. What was different from emotion related articles, however, was the frequent reliance on multiple channels of data: 36.4% of stress articles included both self-report and physiological measures to infer stress as opposed to just 2.1% of the emotion articles. We believe this is a symptom of the SBE literature not embracing formal definitions of emotions. Formal definitions of emotions will tend to agree that emotions are multi-componential psychological responses which include a combination of affective, cognitive, physiological, motivational, and expressive processes[22]. Therefore,

measuring emotions should go beyond self-report measures and should also measure the physiological (e.g., skin conductance, heart rates) and expressive (e.g., facial expressions, speech) aspects of emotions.

The conceptualization of EI was more formal relative to emotions and stress (26.5% of the articles with formal definitions versus 4.4% and 18.2% of the articles with formal definitions for emotions and stress respectively). Articles often cited ideas related to Goleman[188] (2020) and Mayer[189] (1990; theories claiming EI as a type of intelligence), in addition to citing Petrides[190] (2007) in reference to the theory of trait emotional intelligence. If we consider that 13 articles that did not define EI still defined the construct they were measuring (e.g., empathy), 64.7% of EI articles featured formal definitions. This finding seems to signal a trend in SBE where constructs such as emotions are just emerging and therefore lack theoretical integration, while emotional intelligence may be a more established topic with a more matured approach.

We lastly note that we only identified one article that focused on mood. We report that the SBE literature does not seem to interchangeably use the term mood and emotions.

Emotions of Medical Students and Trainees

We report that anxiety and fear were the most captured emotions in our reviewed articles. According to CVT's classification, both are negative-activating emotions, indicating their similarities. The captured emotions reflected how the chosen instruments aligned with the studies' objectives of investigating negative-activating emotions. In other words, if the study sought to investigate anxiety in medical students, the emotional profile reported will mainly be anxiety. Overall, as discussed earlier, we report that while students and trainees do experience positive emotions (e.g., excitement, enjoyment) during certain simulation scenarios, the literature

in our review more often captured negative emotions (e.g., anxiety, fear, frustration, guilt, etc.) due to the intensity of the scenarios (e.g., simulated patient death[51]), and the high expectations set for learners. According to the CVT, high expectations means high appraisal of value (i.e., learners perceive their performance in a simulation to be important), which lead to emotions with high levels of intensity[191]. This is especially the case when one's control over a situation is uncertain or low (i.e., the difficulty of the simulation is high, or there are uncontrollable factors in a simulation)[191].

Inferring Emotions from Stress and Emotional Intelligence

Studies that focused on both stress and emotions (predominantly anxiety) tended to infer anxiety from stress levels of their participants. It is also interesting to note that two of the studies[61,64] that focused on emotion reported "stress" as a type of emotion their participants experienced. Overall, SBE articles tended to acknowledge that anxiety is an expressive component of stress. However, we think caution is warranted in using the term stress, emotions, and anxiety interchangeably. Anxiety is just one of many different stress responses[192], and hence sole reliance on measurement of anxiety may be limited as opposed to relying on multiple measures that also take into account physiological measures or behavioural coding.

Further, when examining what other discrete emotions educators could intend on introducing and measuring in SBE, considering that the CVT illustrates how interchanging stress is with the term emotion, this interchangeability may potentially lead to a narrow capture of emotions. While anxiety is a negative-activating emotion like anger and shame, they are not identical and have different implications for learning[181,193]. Therefore, measuring stress may capture whether one is feeling anxious or not, but not adequately capture anger or shame. For

other emotions that are still negative but are deactivating (e.g., boredom, hopelessness, sadness, disappointment) this issue becomes much more prominent. Finally, measuring stress would not capture positive emotions, missing emotions such as enjoyment and curiosity.

While McNaughton's discourse analysis revealed that educators can view emotions as a skillset, we believe, from an educational psychology perspective, that there is a distinction between having the ability to understand one's own and others' emotions (i.e., EI)[29,189] and the actual experience of feeling specific emotions. Aligned with this, our results showed that the researchers that focused on EI did not infer emotions from EI measurements. The closest inference would be Dohms and others[194] (2020) reporting that students with higher empathy will have better emotion regulation, leading to a calmer emotional profile in stressful situations relative to their peers. We therefore presume that SBE researchers do not confuse experiencing specific discrete emotions with levels of emotional intelligence.

Specific calls to action based on our research partially echoes Leblanc and Posner's review[10]: simulation designers and researchers should ask what emotions they are deliberately or potentially introducing to the participants and consider the impact they may have. However, in thinking and conceptualizing this, we would like to extend their call by specifying *how* to be deliberate and consider integrating a theoretical framework that can formally define what emotions are.

Contextual Factors and Future Directions

When interpreting our study, certain contextual factors matter. Our results show that SBE literature is prominent in the western world, with USA, UK, Germany, Canada, and France compromising 66.6% of the articles. The role of culture may influence studies[195,196] and

increasing cultural diversity in SBE emotions research is therefore an important future direction. Specifically examining SBEs offered in institutions in various nations is one example of this direction. Further, our results indicate that certain types of simulation structures and contexts were underrepresented. For example, only 6.4% of our studies featured interprofessional teams, while certain simulators such as augmented reality or virtual reality simulators were featured much less (3.5%). Focusing the investigation of emotions in such specific SBE contexts could be valuable future directions.

We note limitations of our scoping review process, including our limited selection of electronic databases, not drawing on more than two journals for hand-searching, especially from journals that are SBE-focused. Other limitations include our review focusing on just empirical articles, lacking additional screenings of identified articles' reference lists, and the lack of consulting content experts. In addition, as our study inclusion criteria was quite broad (e.g., including individual-based, team-based simulations, and a wide range of simulators), our findings may have different applications when focused on specific types of SBE. While our results are applicable to the general landscape of the SBE literature, extrapolating our findings to specific subfields of SBE may warrant care.

Our study had some notable strengths as well. Other than the expected standards of following a scoping review framework and being consulted on our search decisions by an expert librarian, we have also conducted screenings, data-extraction and analysis via evidence-based practices. From identifying the affect-related constructs to be studied, to how we conceptualised simulations and their different features, decisions were based on prominent ideas to ensure consistent and accurate intake and analysis of data. Our appendix and body of the manuscript

also offer full transparency in all the steps we have taken based on Arksey and O'Malley's framework.

Future studies could focus on exploring more databases with additional consultation from content experts. Further, our study only extends to the middle of 2020. At the time we wrote this article, the COVID-19 pandemic had left its impact on healthcare education worldwide; it would be interesting to see whether there are any shifts in directions and activities within the SBE literature. Other future directions include this review being a basis for a systematic review on how theoretical frameworks guide emotion-focused SBE research and the interpretation of the results.

Conclusion

We presented a scoping review that aimed to describe the current state of SBE literature pertaining to the conceptualization of emotions and related constructs: stress, emotional intelligence, and mood. Our results revealed that authors of SBE tend to omit including a theoretical framework for conceptualizing emotion-related constructs in their study. We also highlight another tendency amongst SBE studies: capturing negative emotions such as anxiety, where studies that examine stress also often evaluate the levels of anxiety of learners. This tendency of capturing negative emotions reveals that, at first glance, medical trainees and students are stricken with anxiety, fear, and guilt. However, we note that this finding partially stems from studies setting out to measure negative emotions that are also identified as *important* in medical education. Unlike in more traditional fields of education (i.e., Kindergarden-12, Higher Ed), negative emotions seem to have a more profound and authentic role in facilitating learning in SBE. For example, Butteris and colleagues identified how negative emotions such as

frustration and helplessness facilitated motivation for preparedness[53]. Future research that embraces theoretical frameworks such as CVT should equip researchers with the tools they need to critically interpret the impact of such emotions in SBE. In addition, synthesis work focusing on specific types of SBE environments in relation to emotion-related constructs (e.g., interprofessional SBE using virtual reality simulators), and the role of emotion regulation[21] supporting instructional design can serve as next steps. This line of work can illuminate the roles of emotions in SBE and how to best support students' experiences of these influential psychological states that are associated with learning, performance, and psychological well-being.

References

1. Artino AR, Holmboe ES, Durning SJ. Can achievement emotions be used to better understand motivation, learning, and performance in medical education? *Medical Teacher*. 2012;34(3):240-244. doi:10.3109/0142159X.2012.643265

2. Artino AR, Pekrun R. Using Control-Value Theory to Understand Achievement Emotions in Medical Education. *Academic Medicine*. 2014;89(12):1696. doi:doi: 10.1097/ACM.0000000000000536

3. LeBlanc VR, McConnell MM, Monteiro SD. Predictable chaos: a review of the effects of emotions on attention, memory and decision making. *Advances in Health Sciences Education*. 2015;20(1):265-282. doi:10.1007/s10459-014-9516-6

4. Kozlowski D, Hutchinson M, Hurley J, Rowley J, Sutherland J. The role of emotion in clinical decision making: an integrative literature review. *BMC Med Educ*. 2017;17(1):255. doi:10.1186/s12909-017-1089-7

5. Arora S, Ashrafian H, Davis R, Athanasiou T, Darzi A, Sevdalis N. Emotional intelligence in medicine: a systematic review through the context of the ACGME competencies: Emotional intelligence in medicine. *Medical Education*. 2010;44(8):749-764. doi:10.1111/j.1365-2923.2010.03709.x

6. Shahid R, Stirling J, Adams W. Promoting wellness and stress management in residents through emotional intelligence training. *AMEP*. 2018;Volume 9:681-686. doi:10.2147/AMEP.S175299

7. Cerrone SA, Adelman P, Akbar S, Yacht AC, Fornari A. Using Objective Structured Teaching Encounters (OSTEs) to prepare chief residents to be emotionally intelligent leaders. *Medical Education Online*. 2017;22(1):1320186. doi:10.1080/10872981.2017.1320186

8. Weng HC, Hung CM, Liu YT, et al. Associations between emotional intelligence and doctor burnout, job satisfaction and patient satisfaction: Emotional intelligence and doctor and patient satisfaction. *Medical Education*. 2011;45(8):835-842. doi:10.1111/j.1365-2923.2011.03985.x

9. Cherry MG, Fletcher I, O'Sullivan H. Validating relationships among attachment, emotional intelligence and clinical communication. *Med Educ*. 2014;48(10):988-997. doi:10.1111/medu.12526

10. LeBlanc VR, Posner GD. Emotions in simulation-based education: friends or foes of learning? *Adv Simul*. 2022;7(1):3. doi:10.1186/s41077-021-00198-6

11. McNaughton N. Discourse(s) of emotion within medical education: the ever-present absence: Discourses of emotion in medical education. *Medical Education*. 2013;47(1):71-79. doi:10.1111/j.1365-2923.2012.04329.x

12. Harley JM, Jarrell A, Lajoie SP. Emotion regulation tendencies, achievement emotions, and physiological arousal in a medical diagnostic reasoning simulation. *Instructional Science*. 2019;47(2):151-180. doi:10.1007/s11251-018-09480-z

13. Kazemitabar M, Lajoie SP, Doleck T. Analysis of emotion regulation using posture, voice, and attention: A qualitative case study. *Computers and Education Open*. 2021;2:100030. doi:10.1016/j.caeo.2021.100030

14. Heyhoe J, Birks Y, Harrison R, O'Hara JK, Cracknell A, Lawton R. The role of emotion in patient safety: Are we brave enough to scratch beneath the surface? *J R Soc Med*. 2016;109(2):52-58. doi:10.1177/0141076815620614

15. Ali S, Thomson D, Graham TAD, Rickard S, Stang A. High stakes and high emotions: providing safe care in Canadian emergency departments. *OAEM*. 2017;Volume 9:23-26. doi:10.2147/OAEM.S122646

16. Coppin G, Sander D. Theoretical Approaches to Emotion and Its Measurement. In: *Emotion Measurement*. Elsevier; 2016:3-30. doi:10.1016/B978-0-08-100508-8.00001-1

17. Cleary TJ, Durning SJ, Gruppen LD, Hemmer PA, Artino AR. Self-regulated learning in medical education. In: Walsh K, ed. *Oxford Textbook of Medical Education*. Oxford University Press; 2013:465-477. doi:10.1093/med/9780199652679.003.0040

18. Ajjawi R, McIlhenny C. Researching in Surgical Education: An Orientation. In: Nestel D, Dalrymple K, Paige JT, Aggarwal R, eds. *Advancing Surgical Education*. Vol 17. Innovation and Change in Professional Education. Springer Singapore; 2019:341-352. doi:10.1007/978-981-13-3128-2_30

19. Pekrun R. The Control-Value Theory of Achievement Emotions: Assumptions, Corollaries, and Implications for Educational Research and Practice. *Educational Psychology Review*. 2006;18(4):315-341. doi:10.1007/s10648-006-9029-9

20. Eliot JAR, Hirumi A. Emotion theory in education research practice: an interdisciplinary critical literature review. *Education Tech Research Dev*. 2019;67(5):1065-1084. doi:10.1007/s11423-018-09642-3

21. Harley JM, Pekrun R, Taxer JL, Gross JJ. Emotion Regulation in Achievement Situations: An Integrated Model. *Educational Psychologist*. 2019;54(2):106-126. doi:10.1080/00461520.2019.1587297

22. Jarrell A, Lajoie SP. The regulation of achievements emotions: Implications for research and practice. *Canadian Psychology/Psychologie canadienne*. 2017;58(3):276-287. doi:10.1037/cap0000119

23. Loderer K, Pekrun R, Lester JC. Beyond cold technology: A systematic review and meta-analysis on emotions in technology-based learning environments. *Learning and Instruction*. Published online November 2018:S095947521730436X. doi:10.1016/j.learninstruc.2018.08.002

24. Duffy MC, Lajoie SP, Pekrun R, Lachapelle K. Emotions in medical education: Examining the validity of the Medical Emotion Scale (MES) across authentic medical learning environments. *Learning and Instruction*. 2020;70:101150. doi:10.1016/j.learninstruc.2018.07.001

25. Gross JJ, Thompson RA. Emotion regulation: Conceptual foundations. In: *Handbook of Emotion Regulation*. Guilford Publications; 2007:22.

26. Fink G. Stress, Definitions, Mechanisms, and Effects Outlined. In: *Stress: Concepts, Cognition, Emotion, and Behavior*. Elsevier; 2016:3-11. doi:10.1016/B978-0-12-800951-2.00001-7

27. Koolhaas JM, Bartolomucci A, Buwalda B, et al. Stress revisited: A critical evaluation of the stress concept. *Neuroscience & Biobehavioral Reviews*. 2011;35(5):1291-1301. doi:10.1016/j.neubiorev.2011.02.003

28. Dawda D, Hart SD. Assessing emotional intelligence: reliability and validity of the Bar-On Emotional Quotient Inventory (EQ-i) in university students. *Personality and Individual Differences*. 2000;28(4):797-812. doi:10.1016/S0191-8869(99)00139-7

29. Sharp G, Bourke L, Rickard MJFX. Review of emotional intelligence in health care: an introduction to emotional intelligence for surgeons. *ANZ Journal of Surgery*. 2020;90(4):433-440. doi:10.1111/ans.15671

30. Mayer JD, Caruso DR, Salovey P. The Ability Model of Emotional Intelligence: Principles and Updates. *Emotion Review*. 2016;8(4):290-300. doi:10.1177/1754073916639667

31. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. *International Journal of Social Research Methodology*. 2005;8(1):19-32. doi:10.1080/1364557032000119616

32. Thomas A, Lubarsky S, Durning SJ, Young ME. Knowledge Syntheses in Medical Education: Demystifying Scoping Reviews. *Academic Medicine*. 2017;92(2):161-166. doi:10.1097/ACM.0000000000001452

33. Tricco AC, Lillie E, Zarin W, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. *Ann Intern Med*. 2018;169(7):467-473. doi:10.7326/M18-0850

34. Munn Z, Peters MDJ, Stern C, Tufanaru C, McArthur A, Aromataris E. Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. *BMC Med Res Methodol*. 2018;18(1):143. doi:10.1186/s12874-018-0611-x

35. Shuman V, Scherer KR. Concepts and Structures of Emotions. In: *International Handbook of Emotions in Education*. Routledge; 2014. doi:10.4324/9780203148211.ch2

36. Maurice-Ventouris M, Muller Moran HR, Alharbi M, Ahn BT, Harley JM, Lachapelle KJ. The study of visuospatial abilities in trainees: A scoping review and proposed model. *Surgery Open Science*. 2021;5:25-33. doi:10.1016/j.sopen.2021.05.001

37. Pottier P, Hardouin JB, Dejoie T, et al. Stress responses in medical students in ambulatory and in-hospital patient consultations: Stress responses in medical students. *Medical Education*. 2011;45(7):678-687. doi:10.1111/j.1365-2923.2011.03935.x

38. Gaba DM. The future vision of simulation in health care. *Quality and Safety in Health Care*. 2004;13(suppl_1):i2-i10. doi:10.1136/qshc.2004.009878

39. Peters M, Godfrey C, McInerney P, Munn Z, Trico A, Khalil H. Chapter 11: Scoping Reviews. In: Aromataris E, Munn Z, eds. *JBI Manual for Evidence Synthesis*. JBI; 2020. doi:10.46658/JBIMES-20-12

40. Tricco AC, Lillie E, Zarin W, et al. A scoping review on the conduct and reporting of scoping reviews. *BMC Medical Research Methodology*. 2016;16(1). doi:10.1186/s12874-016-0116-4

41. Pelaccia T, Plotnick LH, Audétat MC, et al. A Scoping Review of Physicians' Clinical Reasoning in Emergency Departments. *Annals of Emergency Medicine*. 2020;75(2):206-217. doi:10.1016/j.annemergmed.2019.06.023

42. Thomas A, Lubarsky S, Varpio L, Durning SJ, Young ME. Scoping reviews in health professions education: challenges, considerations and lessons learned about epistemology and methodology. *Adv in Health Sci Educ*. 2020;25(4):989-1002. doi:10.1007/s10459-019-09932-2

43. Yutani M, Takahashi M, Miyaoka H. Changes in communication skills of clinical residents through psychiatric training: Communication skills training. *Psychiatry and Clinical Neurosciences*. 2011;65(6):561-566. doi:10.1111/j.1440-1819.2011.02265.x

44. Romo-Barrientos C, Criado-Álvarez JJ, González-González J, et al. Anxiety levels among health sciences students during their first visit to the dissection room. *BMC Med Educ*. 2020;20(1):109. doi:10.1186/s12909-020-02027-2

45. Zhou Y, Collinson A, Laidlaw A, Humphris G. How Do Medical Students Respond to Emotional Cues and Concerns Expressed by Simulated Patients during OSCE Consultations? – A Multilevel Study. Houser D, ed. *PLoS ONE*. 2013;8(10):e79166. doi:10.1371/journal.pone.0079166

46. Beckers SK, Biermann H, Sopka S, et al. Influence of pre-course assessment using an emotionally activating stimulus with feedback: A pilot study in teaching Basic Life Support. *Resuscitation*. 2012;83(2):219-226. doi:10.1016/j.resuscitation.2011.08.024

47. Russa AD, Mligiliche NuruL. Inspiring Tanzanian medical students into the profession: appraisal of cadaveric dissection stress and coping strategies. *Italian Journal of Anatomy and Embryology*. 2015;Vol 119:268-276 Pages, 276 kB. doi:10.13128/IJAE-15560

48. Romotzky V, Galushko M, Düsterdiek A, et al. "It's Not that Easy"—Medical Students' Fears and Barriers in End-of-Life Communication. *J Canc Educ*. 2015;30(2):333-339. doi:10.1007/s13187-014-0712-0

49. Chang HJ, Kim HJ, Rhyu IJ, Lee YM, Uhm CS. Emotional experiences of medical students during cadaver dissection and the role of memorial ceremonies: a qualitative study. *BMC Med Educ*. 2018;18(1):255. doi:10.1186/s12909-018-1358-0

50. Bernhardt V, Rothkötter HJ, Kasten E. Psychische Belastungen durch die Dissektion am Leichnam im anatomischen Präparierkurs bei Erstsemestern des Studienfachs Medizin. *GMS Zeitschrift für Medizinische Ausbildung; 29(1):Doc12; ISSN 1860-3572*. Published online 2012. doi:10.3205/ZMA000782

51. Fraser K, Huffman J, Ma I, et al. The Emotional and Cognitive Impact of Unexpected Simulated Patient Death. *Chest*. 2014;145(5):958-963. doi:10.1378/chest.13-0987

52. Lefroy J, Brosnan C, Creavin S. Some like it hot: medical student views on choosing the emotional level of a simulation: Student views on emotion in simulation. *Medical Education*. 2011;45(4):354-361. doi:10.1111/j.1365-2923.2010.03881.x

53. Butteris SM, Gladding SP, Eppich W, Hagen SA, Pitt MB. Simulation Use for Global Away Rotations (SUGAR): Preparing Residents for Emotional Challenges Abroad—A Multicenter Study. *Academic Pediatrics*. 2014;14(5):533-541. doi:10.1016/j.acap.2014.05.004

54. Bloomfield JG, O'Neill B, Gillett K. Enhancing student communication during end-of-life care: A pilot study. *Pall Supp Care*. 2015;13(6):1651-1661. doi:10.1017/S147895151500022X

55. Kerins J, Smith SE, Phillips EC, Clarke B, Hamilton AL, Tallentire VR. Exploring transformative learning when developing medical students' non-technical skills. *Med Educ*. 2020;54(3):264-274. doi:10.1111/medu.14062

56. Fraser K, Ma I, Teteris E, Baxter H, Wright B, McLaughlin K. Emotion, cognitive load and learning outcomes during simulation training: Emotion and cognitive load during simulation. *Medical Education*. 2012;46(11):1055-1062. doi:10.1111/j.1365-2923.2012.04355.x

57. Noll C, von Jan U, Raap U, Albrecht UV. Mobile Augmented Reality as a Feature for Self-Oriented, Blended Learning in Medicine: Randomized Controlled Trial. *JMIR Mhealth Uhealth*. 2017;5(9):e139. doi:10.2196/mhealth.7943

58. Robb A, Kopper R, Ambani R, et al. Leveraging Virtual Humans to Effectively Prepare Learners for Stressful Interpersonal Experiences. *IEEE Trans Visual Comput Graphics*. 2013;19(4):662-670. doi:10.1109/TVCG.2013.35

59. Rajeh NA, Badroun LE, Alqarni AK, et al. Cadaver dissection: A positive experience among Saudi female medical students. *Journal of Taibah University Medical Sciences*. 2017;12(3):268-272. doi:10.1016/j.jtumed.2016.07.005

60. Dupuy L, Micoulaud-Franchi JA, Cassoudesalle H, et al. Evaluation of a virtual agent to train medical students conducting psychiatric interviews for diagnosing major depressive disorders. *Journal of Affective Disorders*. 2020;263:1-8. doi:10.1016/j.jad.2019.11.117

61. Groot F, Jonker G, Rinia M, ten Cate O, Hoff RG. Simulation at the Frontier of the Zone of Proximal Development: A Test in Acute Care for Inexperienced Learners. *Academic Medicine*. 2020;95(7):1098-1105. doi:10.1097/ACM.0000000000003265

62. Lewis G, McCullough M, Maxwell AP, Gormley GJ. Ethical reasoning through simulation: a phenomenological analysis of student experience. *Adv Simul*. 2016;1(1):26. doi:10.1186/s41077-016-0027-9

63. Meunier J, Libert Y, Merckaert I, et al. How much is residents' distress detection performance during a clinical round related to their characteristics? *Patient Education and Counseling*. 2011;85(2):180-187. doi:10.1016/j.pec.2010.09.022

64. Torres A, Horodeńska M, Witkowski G, Torres K. Hybrid simulation of pediatric gynecologic examination: a mix-methods study of learners' attitudes and factors affecting learning. *BMC Med Educ*. 2020;20(1):164. doi:10.1186/s12909-020-02076-7

65. O'Rourke SR, Branford KR, Brooks TL, Ives LT, Nagendran A, Compton SN. The Emotional and Behavioral Impact of Delivering Bad News to Virtual versus Real Standardized Patients: A Pilot Study. *Teaching and Learning in Medicine*. 2020;32(2):139-149. doi:10.1080/10401334.2019.1652180

66. Bob MH, Popescu CA, Armean MS, Suciu SM, Buzoianu AD. Ethical views, attitudes and reactions of Romanian medical students to the dissecting room. *Rev Med Chir Soc Med Nat Iasi*. 2014;118(4):1078-1085.

67. Leboulanger N. First cadaver dissection: Stress, preparation, and emotional experience. *European Annals of Otorhinolaryngology, Head and Neck Diseases*. 2011;128(4):175-183. doi:10.1016/j.anorl.2011.01.007

68. Hulsman RL, Pranger S, Koot S, Fabriek M, Karemaker JM, Smets EMA. How stressful is doctor–patient communication? Physiological and psychological stress of medical students in simulated history taking and bad-news consultations. *International Journal of Psychophysiology*. 2010;77(1):26-34. doi:10.1016/j.ijpsycho.2010.04.001

69. Piquette D, Tarshis J, Sinuff T, Fowler RA, Pinto R, Leblanc VR. Impact of Acute Stress on Resident Performance During Simulated Resuscitation Episodes: A Prospective Randomized Cross-Over Study. *Teaching and Learning in Medicine*. 2014;26(1):9-16. doi:10.1080/10401334.2014.859932

70. Harvey A, Bandiera G, Nathens AB, LeBlanc VR. Impact of stress on resident performance in simulated trauma scenarios: *The Journal of Trauma and Acute Care Surgery*. 2012;72(2):497-503. doi:10.1097/TA.0b013e31821f84be

71. McNeer RR, Bennett CL, Dudaryk R. Intraoperative Noise Increases Perceived Task Load and Fatigue in Anesthesiology Residents: A Simulation-Based Study. *Anesthesia & Analgesia*. 2016;122(2):512-525. doi:10.1213/ANE.0000000000001067

72. Clarke S, Horeczko T, Cotton D, Bair A. Heart rate, anxiety and performance of residents during a simulated critical clinical encounter: a pilot study. *BMC Med Educ*. 2014;14(1):153. doi:10.1186/1472-6920-14-153

73. Hurst SA, Baroffio A, Ummel M, Layat Burn C. Helping medical students to acquire a deeper understanding of truth-telling. *Medical Education Online*. 2015;20(1):28133. doi:10.3402/meo.v20.28133

74. Hagemann V, Herbstreit F, Kehren C, et al. Does teaching non-technical skills to medical students improve those skills and simulated patient outcome? *Int J Med Educ*. 2017;8:101-113. doi:10.5116/ijme.58c1.9f0d

75. Crewther BT, Shetty K, Jarchi D, et al. Skill acquisition and stress adaptations following laparoscopic surgery training and detraining in novice surgeons. *Surg Endosc*. 2016;30(7):2961-2968. doi:10.1007/s00464-015-4584-0

76. Cook CJ, Crewther BT. The impact of a competitive learning environment on hormonal and emotional stress responses and skill acquisition and expression in a medical student domain. *Physiology & Behavior*. 2019;199:252-257. doi:10.1016/j.physbeh.2018.11.027

77. Kim KJ. Factors associated with medical student test anxiety in objective structured clinical examinations: a preliminary study. *Int J Med Educ*. 2016;7:424-427. doi:10.5116/ijme.5845.caec

78. Hautz WE, Schröder T, Dannenberg KA, et al. Shame in Medical Education: A Randomized Study of the Acquisition of Intimate Examination Skills and Its Effect on Subsequent Performance. *Teaching and Learning in Medicine*. 2017;29(2):196-206. doi:10.1080/10401334.2016.1254636

79. Fraser K, McLaughlin K. Temporal pattern of emotions and cognitive load during simulation training and debriefing. *Medical Teacher*. 2019;41(2):184-189. doi:10.1080/0142159X.2018.1459531

80. Behrens CC, Dolmans DH, Gormley GJ, Driessen EW. Exploring undergraduate students achievement emotions during ward round simulation: a mixed-method study. *BMC Med Educ*. 2019;19(1):316. doi:10.1186/s12909-019-1753-1

81. Carenzo L, Braithwaite EC, Carfagna F, et al. Cognitive appraisals and team performance under stress: A simulation study. *Med Educ*. 2020;54(3):254-263. doi:10.1111/medu.14050

82. Bauer C, Rimmelé T, Duclos A, et al. Anxiety and stress among anaesthesiology and critical care residents during high-fidelity simulation sessions. *Anaesthesia Critical Care & Pain Medicine*. 2016;35(6):407-416. doi:10.1016/j.accpm.2016.01.004

83. Rogers T, Andler C, O'Brien B, van Schaik S. Self-Reported Emotions in Simulation-Based Learning: Active Participants vs. Observers. *Sim Healthcare*. 2019;14(3):140-145. doi:10.1097/SIH.0000000000000354

84. Toivonen AK, Lindblom-Ylänne S, Louhiala P, Pyörälä E. Medical students' reflections on emotions concerning breaking bad news. *Patient Education and Counseling*. 2017;100(10):1903-1909. doi:10.1016/j.pec.2017.05.036

85. Naismith LM, Lajoie SP. Motivation and emotion predict medical students' attention to computer-based feedback. *Adv in Health Sci Educ*. 2018;23(3):465-485. doi:10.1007/s10459-017-9806-x

86. Vogl E, Pekrun R, Loderer K. Epistemic Emotions and Metacognitive Feelings. In: Moraitou D, Metallidou P, eds. *Trends and Prospects in Metacognition Research across the Life Span*. Springer International Publishing; 2021:41-58. doi:10.1007/978-3-030-51673-4_3

87. Daly SC, Wilson NA, Rinewalt DE, Bines SD, Luu MB, Myers JA. A Subjective Assessment of Medical Student Perceptions on Animal Models in Medical Education. *Journal of Surgical Education*. 2014;71(1):61-64. doi:10.1016/j.jsurg.2013.06.017

88. Feldman Barrett L, Russell JA. Independence and bipolarity in the structure of current affect. *Journal of Personality and Social Psychology*. 1998;74(4):967-984. doi:10.1037/0022-3514.74.4.967

89. Hurley AM, Kennedy PJ, O'Connor L, et al. SOS save our surgeons: Stress levels reduced by robotic surgery. *Gynecol Surg*. 2015;12(3):197-206. doi:10.1007/s10397-015-0891-7

90. Tramèr L, Becker C, Schumacher C, et al. Association of self-esteem, personality, stress and gender with performance of a resuscitation team: A simulation-based study. Roma PG, ed. *PLoS ONE*. 2020;15(5):e0233155. doi:10.1371/journal.pone.0233155

91. Tramèr L, Becker C, Hochstrasser S, Marsch S, Hunziker S. Association of electrocardiogram alterations of rescuers and performance during a simulated cardiac arrest: A prospective simulation study. Lazzeri C, ed. *PLoS ONE*. 2018;13(6):e0198661. doi:10.1371/journal.pone.0198661

92. Bensouda B, Mandel R, Mejri A, Lachapelle J, St-Hilaire M, Ali N. Effect of an audience on trainee stress and performance during simulated neonatal intubation: a randomized crossover trial. *BMC Med Educ*. 2018;18(1):230. doi:10.1186/s12909-018-1338-4

93. Bajunaid K, Mullah MAS, Winkler-Schwartz A, et al. Impact of acute stress on psychomotor bimanual performance during a simulated tumor resection task. *JNS*. 2017;126(1):71-80. doi:10.3171/2015.5.JNS15558

94. Pottier P, Dejoie T, Hardouin JB, et al. Effect of stress on clinical reasoning during simulated ambulatory consultations. *Medical Teacher*. 2013;35(6):472-480. doi:10.3109/0142159X.2013.774336

95. Oriot D, Trigolet M, Kessler DO, Auerbach MA, Ghazali DA. Stress: A Factor Explaining the Gap Between Simulated and Clinical Procedure Success. *Pediatr Emer Care*. 2021;37(12):e1192-e1196. doi:10.1097/PEC.0000000000001962

96. Phitayakorn R, Minehart RD, Pian-Smith MCM, Hemingway MW, Petrusa ER. Practicality of using galvanic skin response to measure intraoperative physiologic autonomic activation in operating room team members. *Surgery*. 2015;158(5):1415-1420. doi:10.1016/j.surg.2015.04.009

97. Baker BG, Bhalla A, Doleman B, et al. Simulation fails to replicate stress in trainees performing a technical procedure in the clinical environment. *Medical Teacher*. 2017;39(1):53-57. doi:10.1080/0142159X.2016.1230188

98. Chia TI, Oyeniran OI, Ajagbe AO, Onigbinde OA, Oraebosi MI. The symptoms and stress experienced by medical students in anatomy dissection halls. *Journal of Taibah University Medical Sciences*. 2020;15(1):8-13. doi:10.1016/j.jtumed.2020.01.001

99. Hunziker S, Semmer NK, Tschan F, Schuetz P, Mueller B, Marsch S. Dynamics and association of different acute stress markers with performance during a simulated resuscitation. *Resuscitation*. 2012;83(5):572-578. doi:10.1016/j.resuscitation.2011.11.013

100. Harendza S, Soll H, Prediger S, Kadmon M, Berberat PO, Oubaid V. Assessing core competences of medical students with a test for flight school applicants. *BMC Med Educ*. 2019;19(1):9. doi:10.1186/s12909-018-1438-1

101. Arora S, Russ S, Petrides KV, et al. Emotional Intelligence and Stress in Medical Students Performing Surgical Tasks: *Academic Medicine*. 2011;86(10):1311-1317. doi:10.1097/ACM.0b013e31822bd7aa

102. Böckers A, Baader C, Fassnacht UK, Öchsner W, Böckers TM. Reduction of mental distress in the dissection course by introducing the body donor experience through anatomical demonstrations of organ systems. *Anat Sci Ed*. 2012;5(6):321-329. doi:10.1002/ase.1292

103. Andreatta PB, Hillard M, Krain LP. The impact of stress factors in simulation-based laparoscopic training. *Surgery*. 2010;147(5):631-639. doi:10.1016/j.surg.2009.10.071

104. Ontrup G, Vogel M, Wolf OT, Zahn PK, Kluge A, Hagemann V. Does simulation-based training in medical education need additional stressors? An experimental study. *Ergonomics*. 2020;63(1):80-90. doi:10.1080/00140139.2019.1677948

105. Boyle E, Kennedy AM, Doherty E, O'Keeffe D, Traynor O. Coping with stress in surgery: the difficulty of measuring non-technical skills. *Ir J Med Sci*. 2011;180(1):215-220. doi:10.1007/s11845-010-0578-z

106. Pottier P, Hardouin JB, Dejoie T, et al. Effect of Extrinsic and Intrinsic Stressors on Clinical Skills Performance in Third-Year Medical Students. *J GEN INTERN MED*. 2015;30(9):1259-1269. doi:10.1007/s11606-015-3314-6

107. Selye H. Stress and the general adaptation syndrome. *Br Med J*. 1950;1(4667):1383-1392. doi:10.1136/bmj.1.4667.1383

108. Biggs A, Brough P, Drummond S. Lazarus and Folkman's Psychological Stress and Coping Theory. In: Cooper CL, Quick JC, eds. *The Handbook of Stress and Health*. John Wiley & Sons, Ltd; 2017:349-364. doi:10.1002/9781118993811.ch21

109. Maher Z, Milner R, Cripe J, Gaughan J, Fish J, Goldberg AJ. Stress training for the surgical resident. *The American Journal of Surgery*. 2013;205(2):169-174. doi:10.1016/j.amjsurg.2012.10.007

110. Bong CL, Lee S, Ng ASB, Allen JC, Lim EHL, Vidyarthi A. The effects of active (hot-seat) versus observer roles during simulation-based training on stress levels and non-technical performance: a randomized trial. *Adv Simul*. 2017;2(1):7. doi:10.1186/s41077-017-0040-7

111. Poolton JM, Wilson MR, Malhotra N, Ngo K, Masters RSW. A comparison of evaluation, time pressure, and multitasking as stressors of psychomotor operative performance. *Surgery*. 2011;149(6):776-782. doi:10.1016/j.surg.2010.12.005

112. Malhotra N, Poolton JM, Wilson MR, Ngo K, Masters RSW. Conscious monitoring and control (reinvestment) in surgical performance under pressure. *Surg Endosc*. 2012;26(9):2423-2429. doi:10.1007/s00464-012-2193-8

113. Dias RD, Scalabrini-Neto A. Acute stress in residents playing different roles during emergency simulations: a preliminary study. *Int J Med Educ*. 2017;8:239-243. doi:10.5116/ijme.5929.60f1

114. Bakhsh A, Martin GFJ, Bicknell CD, Pettengell C, Riga C. An Evaluation of the Impact of High-Fidelity Endovascular Simulation on Surgeon Stress and Technical Performance. *Journal of Surgical Education*. 2019;76(3):864-871. doi:10.1016/j.jsurg.2018.10.015

115. Finan E, Bismilla Z, Whyte HE, LeBlanc V, McNamara PJ. High-fidelity simulator technology may not be superior to traditional low-fidelity equipment for neonatal resuscitation training. *J Perinatol*. 2012;32(4):287-292. doi:10.1038/jp.2011.96

116. Daglius Dias R, Scalabrini Neto A. Stress levels during emergency care: A comparison between reality and simulated scenarios. *Journal of Critical Care*. 2016;33:8-13. doi:10.1016/j.jcrc.2016.02.010

117. Flinn JT, Miller A, Pyatka N, Brewer J, Schneider T, Cao CGL. The effect of stress on learning in surgical skill acquisition. *Medical Teacher*. 2016;38(9):897-903. doi:10.3109/0142159X.2015.1114597

118. DeMaria S, Silverman ER, Lapidus KAB, et al. The impact of simulated patient death on medical students' stress response and learning of ACLS. *Medical Teacher*. 2016;38(7):730-737. doi:10.3109/0142159X.2016.1150986

119. Harvey A, Nathens AB, Bandiera G, LeBlanc VR. Threat and challenge: cognitive appraisal and stress responses in simulated trauma resuscitations: Appraisal and stress in simulated trauma resuscitations. *Medical Education*. 2010;44(6):587-594. doi:10.1111/j.1365-2923.2010.03634.x

120. Łoś K, Chmielewski J, Łuczyński W. Relationship between Executive Functions, Mindfulness, Stress, and Performance in Pediatric Emergency Simulations. *IJERPH*. 2020;17(6):2040. doi:10.3390/ijerph17062040

121. Waterland P, Khan FS, Ismaili E, Cheruvu C. Environmental Noise as an Operative Stressor During Simulated Laparoscopic Surgery. *Surgical Laparoscopy, Endoscopy & Percutaneous Techniques*. 2016;26(2):133-136. doi:10.1097/SLE.0000000000000250

122. Modi HN, Singh H, Fiorentino F, et al. Association of Residents' Neural Signatures With Stress Resilience During Surgery. *JAMA Surg*. 2019;154(10):e192552. doi:10.1001/jamasurg.2019.2552

123. Modi HN, Singh H, Orihuela-Espina F, et al. Temporal Stress in the Operating Room: Brain Engagement Promotes "Coping" and Disengagement Prompts "Choking." *Annals of Surgery*. 2018;267(4):683-691. doi:10.1097/SLA.0000000000002289

124. Lizotte MH, Janvier A, Latraverse V, et al. The Impact of Neonatal Simulations on Trainees' Stress and Performance: A Parallel-Group Randomized Trial*. *Pediatric Critical Care Medicine*. 2017;18(5):434-441. doi:10.1097/PCC.0000000000001119

125. Deinzer R, Kiupel S, Weik U. Endocrine and psychological stress response in simulated doctor-patient interactions in medical education. *Psychoneuroendocrinology*. 2019;105:172-177. doi:10.1016/j.psyneuen.2018.09.028

126. Keitel A, Ringleb M, Schwartges I, et al. Endocrine and psychological stress responses in a simulated emergency situation. *Psychoneuroendocrinology*. 2011;36(1):98-108. doi:10.1016/j.psyneuen.2010.06.011

127. Goldberg MB, Mazzei M, Maher Z, et al. Optimizing performance through stress training — An educational strategy for surgical residents. *The American Journal of Surgery*. 2018;216(3):618-623. doi:10.1016/j.amjsurg.2017.11.040

128. Geeraerts T, Roulleau P, Cheisson G, et al. Physiological and self-assessed psychological stress induced by a high fidelity simulation course among third year anesthesia and critical care residents: An observational study. *Anaesthesia Critical Care & Pain Medicine*. 2017;36(6):403-406. doi:10.1016/j.accpm.2017.06.002

129. Hochberg MS, Berman RS, Kalet AL, Zabar SR, Gillespie C, Pachter HL. The stress of residency: recognizing the signs of depression and suicide in you and your fellow residents. *The American Journal of Surgery*. 2013;205(2):141-146. doi:10.1016/j.amjsurg.2012.08.003

130. Ishiyama Y, Inaki N, Bando H, Yamada T. Assessment of Intracorporeal Suturing in Single-Port Surgery Using an Experimental Suturing Model. *Indian J Surg*. 2017;79(2):137-142. doi:10.1007/s12262-016-1445-6

131. Maschuw K, Schlosser K, Kupietz E, Slater EP, Weyers P, Hassan I. Do Soft Skills Predict Surgical Performance?: A Single-Center Randomized Controlled Trial Evaluating Predictors of Skill Acquisition in Virtual Reality Laparoscopy. *World J Surg*. 2011;35(3):480-486. doi:10.1007/s00268-010-0933-2

132. Larkin AC, Cahan MA, Whalen G, et al. Human Emotion and Response in Surgery (HEARS): A Simulation-Based Curriculum for Communication Skills, Systems-Based

Practice, and Professionalism in Surgical Residency Training. *Journal of the American College of Surgeons*. 2010;211(2):285-292. doi:10.1016/j.jamcollsurg.2010.04.004

133. Hunziker S, Pagani S, Fasler K, Tschan F, Semmer NK, Marsch S. Impact of a stress coping strategy on perceived stress levels and performance during a simulated cardiopulmonary resuscitation: a randomized controlled trial. *BMC Emerg Med*. 2013;13(1):8. doi:10.1186/1471-227X-13-8

134. Pai DR, Ram S, Madan SS, Soe HHK, Barua A. Causes of stress and their change with repeated sessions as perceived by undergraduate medical students during high-fidelity trauma simulation. *THE NATIONAL MEDICAL JOURNAL OF INDIA*. 2014;27(4):7.

135. Nicolaides M, Theodorou E, Emin EI, et al. Team performance training for medical students: Low vs high fidelity simulation. *Annals of Medicine and Surgery*. 2020;55:308-315. doi:10.1016/j.amsu.2020.05.042

136. Bragard I, Etienne AM, Merckaert I, Libert Y, Razavi D. Efficacy of a Communication and Stress Management Training on Medical Residents' Self-efficacy, Stress to Communicate and Burnout: A Randomized Controlled Study. *J Health Psychol*. 2010;15(7):1075-1081. doi:10.1177/1359105310361992

137. Murugavadigal N. The impact of stressors on the learning outcome of high-fidelity patient simulation in undergraduate medical students. 2020;75(3):7.

138. Schwill S, Fahrbach-Veeser J, Moeltner A, et al. Peers as OSCE assessors for junior medical students – a review of routine use: a mixed methods study. *BMC Med Educ*. 2020;20(1):17. doi:10.1186/s12909-019-1898-y

139. Bragard I, Seghaye MC, Farhat N, et al. Implementation of a 2-Day Simulation-Based Course to Prepare Medical Graduates on Their First Year of Residency: *Pediatric Emergency Care*. 2018;34(12):857-861. doi:10.1097/PEC.0000000000000930

140. Klein MI, Mouraviev V, Craig C, et al. Mental stress experienced by first-year residents and expert surgeons with robotic and laparoscopic surgery interfaces. *J Robotic Surg*. 2014;8(2):149-155. doi:10.1007/s11701-013-0446-8

141. Klein MI, Warm JS, Riley MA, et al. Mental Workload and Stress Perceived by Novice Operators in the Laparoscopic and Robotic Minimally Invasive Surgical Interfaces. *Journal of Endourology*. 2012;26(8):1089-1094. doi:10.1089/end.2011.0641

142. Anyanwu GE, Nto JN, Agu AU, Ekezie J, Esom EA. Musical preferences and learning outcome of medical students in cadaver dissection laboratory: A Nigerian survey. *Annals of Anatomy - Anatomischer Anzeiger*. 2016;208:228-233. doi:10.1016/j.aanat.2016.07.010

143. Alt-Epping B, Lohse C, Viebahn C, Steinbüchel N von, Benze G, Nauck F. On death and dying – an exploratory and evaluative study of a reflective, interdisciplinary course element

in undergraduate anatomy teaching. *BMC Med Educ*. 2014;14(1):15. doi:10.1186/1472-6920-14-15

144. Anton NE, Bean EA, Myers E, Stefanidis D. Optimizing learner engagement during mental skills training: A pilot study of small group vs. individualized training. *The American Journal of Surgery*. 2020;219(2):335-339. doi:10.1016/j.amjsurg.2019.12.022

145. Khairat S, Coleman C, Ottmar P, Bice T, Koppel R, Carson SS. Physicians' gender and their use of electronic health records: findings from a mixed-methods usability study. *Journal of the American Medical Informatics Association*. 2019;26(12):1505-1514. doi:10.1093/jamia/ocz126

146. Feng X, Morandi A, Boehne M, et al. 3-Dimensional (3D) laparoscopy improves operating time in small spaces without impact on hemodynamics and psychomental stress parameters of the surgeon. *Surg Endosc*. 2015;29(5):1231-1239. doi:10.1007/s00464-015-4083-3

147. Pavlidis I, Zavlin D, Khatri AR, Wesley A, Panagopoulos G, Echo A. Absence of Stressful Conditions Accelerates Dexterous Skill Acquisition in Surgery. *Sci Rep*. 2019;9(1):1747. doi:10.1038/s41598-019-38727-z

148. Cleary DR, Siler DA, Whitney N, Selden NR. A microcontroller-based simulation of dural venous sinus injury for neurosurgical training. *Journal of Neurosurgery*. 2018;128(5):1553-1559. doi:10.3171/2016.12.JNS162165

149. Maddox MM, Lopez A, Mandava SH, et al. Electroencephalographic Monitoring of Brain Wave Activity During Laparoscopic Surgical Simulation to Measure Surgeon Concentration and Stress: Can the Student Become the Master? *Journal of Endourology*. 2015;29(12):1329-1333. doi:10.1089/end.2015.0239

150. Le Lous M, De Chanaud N, Bourret A, et al. Improving the quality of transvaginal ultrasound scan by simulation training for general practice residents. *Adv Simul*. 2017;2(1):24. doi:10.1186/s41077-017-0056-z

151. Fukushima Y, Yamada A, Imoto N, Iba T. Does TeamSTEPPS affect psychological status? *IJHCQA*. 2019;32(1):11-20. doi:10.1108/IJHCQA-08-2017-0156

152. Marvin RK, Saepoo MB, Ye S, et al. Salivary protein changes in response to acute stress in medical residents performing advanced clinical simulations: a pilot proteomics study. *Biomarkers*. 2017;22(3-4):372-382. doi:10.1080/1354750X.2017.1279215

153. LaPorta AJ, McKee J, Hoang T, et al. Stress Inoculation: Preparing Outside the Box in Surgical Resuscitation and Education. *Curr Trauma Rep*. 2017;3(2):135-143. doi:10.1007/s40719-017-0090-2

154. Timberlake MD, Stefanidis D, Gardner AK. Examining the impact of surgical coaching on trainee physiologic response and basic skill acquisition. *Surg Endosc*. 2018;32(10):4183-4190. doi:10.1007/s00464-018-6163-7

155. Dedmon MM, O'Connell BP, Yawn RJ, et al. Measuring Mental Stress During Otologic Surgery Using Heart Rate Variability Analysis. *Otology & Neurotology*. 2019;40(4):529-534. doi:10.1097/MAO.0000000000002187

156. Kromann CB, Jensen ML, Ringsted C. Test-enhanced learning may be a gender-related phenomenon explained by changes in cortisol level: Test-enhanced learning, gender and cortisol. *Medical Education*. 2011;45(2):192-199. doi:10.1111/j.1365-2923.2010.03790.x

157. Cheung EO, Barsuk JH, Mitra D, et al. Preliminary Efficacy of a Brief Mindfulness Intervention for Procedural Stress in Medical Intern Simulated Performance: A Randomized Controlled Pilot Trial. *The Journal of Alternative and Complementary Medicine*. 2020;26(4):282-290. doi:10.1089/acm.2019.0209

158. Johnson JM, Stern TA. Teaching Residents About Emotional Intelligence and Its Impact on Leadership. *Acad Psychiatry*. 2014;38(4):510-513. doi:10.1007/s40596-014-0048-4

159. Reed S, Kassis K, Nagel R, Verbeck N, Mahan JD, Shell R. Does emotional intelligence predict breaking bad news skills in pediatric interns? A pilot study. *Medical Education Online*. 2015;20(1):24245. doi:10.3402/meo.v20.24245

160. Dugan JW, Weatherly RA, Girod DA, Barber CE, Tsue TT. A Longitudinal Study of Emotional Intelligence Training for Otolaryngology Residents and Faculty. *JAMA Otolaryngol Head Neck Surg*. 2014;140(8):720. doi:10.1001/jamaoto.2014.1169

161. Nayar SK, Musto L, Fernandes R, Bharathan R. Emotional Intelligence Predicts Accurate Self-Assessment of Surgical Quality: A Pilot Study. *Journal of Surgical Research*. 2020;245:383-389. doi:10.1016/j.jss.2019.07.051

162. Cherry MG, Fletcher I, O'Sullivan H. Exploring the relationships among attachment, emotional intelligence and communication: Emotional intelligence and clinical communication. *Medical Education*. 2013;47(3):317-325. doi:10.1111/medu.12115

163. Lim G, Gardner AK. Emotional Intelligence and Delivering Bad News: The Jury is Still Out. *Journal of Surgical Education*. 2019;76(3):779-784. doi:10.1016/j.jsurg.2018.09.017

164. Kanter JW, Rosen DC, Manbeck KE, et al. Addressing microaggressions in racially charged patient-provider interactions: a pilot randomized trial. *BMC Med Educ*. 2020;20(1):88. doi:10.1186/s12909-020-02004-9

165. Graupe T, Fischer MR, Strijbos JW, Kiessling C. Development and piloting of a Situational Judgement Test for emotion-handling skills using the Verona Coding Definitions of

Emotional Sequences (VR-CoDES). *Patient Education and Counseling*. 2020;103(9):1839-1845. doi:10.1016/j.pec.2020.04.001

166. Farrell MH. Child Health Providers' Precautionary Discussion of Emotions During Communication About Results of Newborn Genetic Screening. *Arch Pediatr Adolesc Med*. 2012;166(1):62. doi:10.1001/archpediatrics.2011.696

167. Ruiz-Moral R, Pérula de Torres L, Monge D, García Leonardo C, Caballero F. Teaching medical students to express empathy by exploring patient emotions and experiences in standardized medical encounters. *Patient Education and Counseling*. 2017;100(9):1694-1700. doi:10.1016/j.pec.2017.04.018

168. Bourgeon L, Bensalah M, Vacher A, Ardouin JC, Debien B. Role of emotional competence in residents' simulated emergency care performance: a mixed-methods study. *BMJ Qual Saf*. 2016;25(5):364-371. doi:10.1136/bmjqs-2015-004032

169. Schweller M, Costa FO, Antônio MÂRGM, Amaral EM, de Carvalho-Filho MA. The Impact of Simulated Medical Consultations on the Empathy Levels of Students at One Medical School: *Academic Medicine*. 2014;89(4):632-637. doi:10.1097/ACM.0000000000000175

170. Sevrain-Goideau M, Gohier B, Bellanger W, Annweiler C, Campone M, Coutant R. Forum theater staging of difficult encounters with patients to increase empathy in students: evaluation of efficacy at The University of Angers Medical School. *BMC Med Educ*. 2020;20(1):58. doi:10.1186/s12909-020-1965-4

171. Nelson S, Germann C, MacVane C, Bloch R, Fallon T, Strout T. Intern as Patient: A Patient Experience Simulation to Cultivate Empathy in Emergency Medicine Residents. *WestJEM*. Published online January 18, 2018:41-48. doi:10.5811/westjem.2017.11.35198

172. Ogle J, Bushnell JA, Caputi P. Empathy is related to clinical competence in medical care. *Med Educ*. 2013;47(8):824-831. doi:10.1111/medu.12232

173. Berg K, Majdan JF, Berg D, Veloski J, Hojat M. Medical Students' Self-Reported Empathy and Simulated Patients' Assessments of Student Empathy: An Analysis by Gender and Ethnicity: *Academic Medicine*. 2011;86(8):984-988. doi:10.1097/ACM.0b013e3182224f1f

174. Hojat M, Mangione S, Nasca TJ, et al. The Jefferson Scale of Physician Empathy: Development and Preliminary Psychometric Data. *Educational and Psychological Measurement*. 2001;61(2):349-365. doi:10.1177/00131640121971158

175. LaNoue MD, Roter DL. Exploring patient-centeredness: The relationship between self-reported empathy and patient-centered communication in medical trainees. *Patient Education and Counseling*. 2018;101(6):1143-1146. doi:10.1016/j.pec.2018.01.016

176. Andrei F, Smith MM, Surcinelli P, Baldaro B, Saklofske DH. The Trait Emotional Intelligence Questionnaire: Internal Structure, Convergent, Criterion, and Incremental Validity in an Italian Sample. *Measurement and Evaluation in Counseling and Development*. 2016;49(1):34-45. doi:10.1177/0748175615596786

177. Trickey AW, Newcomb AB, Porrey M, et al. Assessment of Surgery Residents' Interpersonal Communication Skills: Validation Evidence for the Communication Assessment Tool in a Simulation Environment. *Journal of Surgical Education*. 2016;73(6):e19-e27. doi:10.1016/j.jsurg.2016.04.016

178. Mayer JD, Salovey P, Caruso DR. Mayer-Salovey-Caruso Emotional Intelligence Test. Published online May 7, 2012. doi:10.1037/t05047-000

179. Huizinga CRH, de Kam ML, Stockmann HBAC, van Gerven JMA, Cohen AF, van der Bogt KEA. Evaluating Fitness to Perform in Surgical Residents after Night Shifts and Alcohol Intoxication: The development of a "Fit-to-Perform" test. *Journal of Surgical Education*. 2018;75(4):968-977. doi:10.1016/j.jsurg.2018.01.010

180. Bond A, Lader M. The use of analogue scales in rating subjective feelings. *British Journal of Medical Psychology*. 1974;47(3):211-218. doi:10.1111/j.2044-8341.1974.tb02285.x

181. Pekrun R, Perry RP. Control-Value Theory of Achievement Emotions. In: *International Handbook of Emotions in Education*. Routledge; 2014. doi:10.4324/9780203148211.ch7

182. Spielberger CD, Gonzalez-Reigosa F, Martinez-Urrutia A, Natalicio LF, Natalicio DS. The state-trait anxiety inventory. *Revista Interamericana de Psicologia/Interamerican journal of psychology*. 1971;5(3 & 4).

183. Watson D, Friend R. Measurement of social-evaluative anxiety. *J Consult Clin Psychol*. 1969;33(4):448-457. doi:10.1037/h0027806

184. Turner JE. Researching state shame with the Experiential Shame Scale. *J Psychol*. 2014;148(5):577-601. doi:10.1080/00223980.2013.818927

185. Pekrun R, Goetz T, Frenzel AC, Barchfeld P, Perry RP. Measuring emotions in students' learning and performance: The Achievement Emotions Questionnaire (AEQ). *Contemporary Educational Psychology*. 2011;36(1):36-48. doi:10.1016/j.cedpsych.2010.10.002

186. Rowe A, Fitness J. Understanding the Role of Negative Emotions in Adult Learning and Achievement: A Social Functional Perspective. *Behavioral Sciences*. 2018;8(2):27. doi:10.3390/bs8020027

187. Gross JJ. The Extended Process Model of Emotion Regulation: Elaborations, Applications, and Future Directions. *Psychological Inquiry*. 2015;26(1):130-137. doi:10.1080/1047840X.2015.989751

188. Goleman D. *Emotional Intelligence*. 25th anniversary edition. Bantam Books; 2020.

189. Salovey P, Mayer JD. Emotional Intelligence. *Imagination, Cognition and Personality*. 1990;9(3):185-211. doi:10.2190/DUGG-P24E-52WK-6CDG

190. Petrides KV, Pita R, Kokkinaki F. The location of trait emotional intelligence in personality factor space. *British Journal of Psychology*. 2007;98(2):273-289. doi:10.1348/000712606X120618

191. Pekrun R. Achievement emotions: A control-value theory perspective. In: *Emotions in Late Modernity*. Routledge; 2019:142-157.

192. Crosswell AD, Lockwood KG. Best practices for stress measurement: How to measure psychological stress in health research. *Health Psychology Open*. 2020;7(2):205510292093307. doi:10.1177/2055102920933072

193. Ahn B "Tony", Harley J. Facial expressions when learning with a Queer History App: Application of the Control Value Theory of Achievement Emotions. 2020;51(5):14.

194. Dohms MC. Video-based feedback using real consultations for a formative assessment in communication skills. Published online 2020:9.

195. Ogarkova A. Cross-lingual translatability of emotion terms. In: *Emotion Measurement*. Elsevier; 2021:909-935. doi:10.1016/B978-0-12-821124-3.00028-4

196. Chentsova Dutton YE, Lyons SH. Different ways of measuring emotions cross-culturally. In: *Emotion Measurement*. Elsevier; 2021:937-974. doi:10.1016/B978-0-12-821124-3.00029-6

References for Included Articles for Review

1. Romotzky V, Galushko M, Düsterdiek A, et al. "It's Not that Easy"—Medical Students' Fears and Barriers in End-of-Life Communication. *J Canc Educ*. 2015;30(2):333-339. doi:10.1007/s13187-014-0712-0

2. Feng X, Morandi A, Boehne M, et al. 3-Dimensional (3D) laparoscopy improves operating time in small spaces without impact on hemodynamics and psychomental stress parameters of the surgeon. *Surg Endosc*. 2015;29(5):1231-1239. doi:10.1007/s00464-015-4083-3

3. Poolton JM, Wilson MR, Malhotra N, Ngo K, Masters RSW. A comparison of evaluation, time pressure, and multitasking as stressors of psychomotor operative performance. *Surgery*. 2011;149(6):776-782. doi:10.1016/j.surg.2010.12.005

4. Dugan JW, Weatherly RA, Girod DA, Barber CE, Tsue TT. A Longitudinal Study of Emotional Intelligence Training for Otolaryngology Residents and Faculty. *JAMA Otolaryngol Head Neck Surg*. 2014;140(8):720. doi:10.1001/jamaoto.2014.1169

5. Cleary DR, Siler DA, Whitney N, Selden NR. A microcontroller-based simulation of dural venous sinus injury for neurosurgical training. *Journal of Neurosurgery*. 2018;128(5):1553-1559. doi:10.3171/2016.12.JNS162165

6. Grayson-Sneed KA, Smith RC. A research coding method to evaluate a smoking cessation model for training residents—A preliminary report. *Patient Education and Counseling*. 2018;101(3):541-545. doi:10.1016/j.pec.2017.09.010

7. Grayson-Sneed KA, Smith RC. A research coding method to evaluate medical clinicians conduct of behavioral health care in patients with unexplained symptoms. *Patient Education and Counseling*. 2018;101(4):743-749. doi:10.1016/j.pec.2017.10.006

8. Daly SC, Wilson NA, Rinewalt DE, Bines SD, Luu MB, Myers JA. A Subjective Assessment of Medical Student Perceptions on Animal Models in Medical Education. *Journal of Surgical Education*. 2014;71(1):61-64. doi:10.1016/j.jsurg.2013.06.017

9. Pavlidis I, Zavlin D, Khatri AR, Wesley A, Panagopoulos G, Echo A. Absence of Stressful Conditions Accelerates Dexterous Skill Acquisition in Surgery. *Sci Rep*. 2019;9(1):1747. doi:10.1038/s41598-019-38727-z

10. Dias RD, Scalabrini-Neto A. Acute stress in residents playing different roles during emergency simulations: a preliminary study. *Int J Med Educ*. 2017;8:239-243. doi:10.5116/ijme.5929.60f1

11. Kanter JW, Rosen DC, Manbeck KE, et al. Addressing microaggressions in racially charged patient-provider interactions: a pilot randomized trial. *BMC Med Educ*. 2020;20(1):88. doi:10.1186/s12909-020-02004-9

12. Bakhsh A, Martin GFJ, Bicknell CD, Pettengell C, Riga C. An Evaluation of the Impact of High-Fidelity Endovascular Simulation on Surgeon Stress and Technical Performance. *Journal of Surgical Education*. 2019;76(3):864-871. doi:10.1016/j.jsurg.2018.10.015

13. Bauer C, Rimmelé T, Duclos A, et al. Anxiety and stress among anaesthesiology and critical care residents during high-fidelity simulation sessions. *Anaesthesia Critical Care & Pain Medicine*. 2016;35(6):407-416. doi:10.1016/j.accpm.2016.01.004

14. Romo-Barrientos C, Criado-Álvarez JJ, González-González J, et al. Anxiety levels among health sciences students during their first visit to the dissection room. *BMC Med Educ*. 2020;20(1):109. doi:10.1186/s12909-020-02027-2

15. Harendza S, Soll H, Prediger S, Kadmon M, Berberat PO, Oubaid V. Assessing core competences of medical students with a test for flight school applicants. *BMC Med Educ*. 2019;19(1):9. doi:10.1186/s12909-018-1438-1

16. Posner G, Naik V, Bidlake E, et al. Assessing Residents' Disclosure of Adverse Events: Traditional Objective Structured Clinical Examinations Versus Mixed Reality. *Journal of Obstetrics and Gynaecology Canada*. 2012;34(4):367-373. doi:10.1016/S1701-2163(16)35219-7

17. Ishiyama Y, Inaki N, Bando H, Yamada T. Assessment of Intracorporeal Suturing in Single-Port Surgery Using an Experimental Suturing Model. *Indian J Surg*. 2017;79(2):137-142. doi:10.1007/s12262-016-1445-6

18. Trickey AW, Newcomb AB, Porrey M, et al. Assessment of Surgery Residents' Interpersonal Communication Skills: Validation Evidence for the Communication Assessment Tool in a Simulation Environment. *Journal of Surgical Education*. 2016;73(6):e19-e27. doi:10.1016/j.jsurg.2016.04.016

19. Tramèr L, Becker C, Hochstrasser S, Marsch S, Hunziker S. Association of electrocardiogram alterations of rescuers and performance during a simulated cardiac arrest: A prospective simulation study. Lazzeri C, ed. *PLoS ONE*. 2018;13(6):e0198661. doi:10.1371/journal.pone.0198661

20. Modi HN, Singh H, Fiorentino F, et al. Association of Residents' Neural Signatures With Stress Resilience During Surgery. *JAMA Surg*. 2019;154(10):e192552. doi:10.1001/jamasurg.2019.2552

21. Tramèr L, Becker C, Schumacher C, et al. Association of self-esteem, personality, stress and gender with performance of a resuscitation team: A simulation-based study. Roma PG, ed. *PLoS ONE*. 2020;15(5):e0233155. doi:10.1371/journal.pone.0233155

22. Rajeh NA, Badroun LE, Alqarni AK, et al. Cadaver dissection: A positive experience among Saudi female medical students. *Journal of Taibah University Medical Sciences*. 2017;12(3):268-272. doi:10.1016/j.jtumed.2016.07.005

23. Pai DR, Ram S, Madan SS, Soe HHK, Barua A. Causes of stress and their change with repeated sessions as perceived by undergraduate medical students during high-fidelity trauma simulation. *THE NATIONAL MEDICAL JOURNAL OF INDIA*. 2014;27(4):7.

24. Yutani M, Takahashi M, Miyaoka H. Changes in communication skills of clinical residents through psychiatric training: Communication skills training. *Psychiatry and Clinical Neurosciences*. 2011;65(6):561-566. doi:10.1111/j.1440-1819.2011.02265.x

25. Farrell MH. Child Health Providers' Precautionary Discussion of Emotions During Communication About Results of Newborn Genetic Screening. *Arch Pediatr Adolesc Med*. 2012;166(1):62. doi:10.1001/archpediatrics.2011.696

26. Carenzo L, Braithwaite EC, Carfagna F, et al. Cognitive appraisals and team performance under stress: A simulation study. *Med Educ*. 2020;54(3):254-263. doi:10.1111/medu.14050

27. Malhotra N, Poolton JM, Wilson MR, Ngo K, Masters RSW. Conscious monitoring and control (reinvestment) in surgical performance under pressure. *Surg Endosc*. 2012;26(9):2423-2429. doi:10.1007/s00464-012-2193-8

28. Boyle E, Kennedy AM, Doherty E, O'Keeffe D, Traynor O. Coping with stress in surgery: the difficulty of measuring non-technical skills. *Ir J Med Sci*. 2011;180(1):215-220. doi:10.1007/s11845-010-0578-z

29. Graupe T, Fischer MR, Strijbos JW, Kiessling C. Development and piloting of a Situational Judgement Test for emotion-handling skills using the Verona Coding Definitions of Emotional Sequences (VR-CoDES). *Patient Education and Counseling*. 2020;103(9):1839-1845. doi:10.1016/j.pec.2020.04.001

30. Maschuw K, Schlosser K, Kupietz E, Slater EP, Weyers P, Hassan I. Do Soft Skills Predict Surgical Performance?: A Single-Center Randomized Controlled Trial Evaluating Predictors of Skill Acquisition in Virtual Reality Laparoscopy. *World J Surg*. 2011;35(3):480-486. doi:10.1007/s00268-010-0933-2

31. Reed S, Kassis K, Nagel R, Verbeck N, Mahan JD, Shell R. Does emotional intelligence predict breaking bad news skills in pediatric interns? A pilot study. *Medical Education Online*. 2015;20(1):24245. doi:10.3402/meo.v20.24245

32. Ontrup G, Vogel M, Wolf OT, Zahn PK, Kluge A, Hagemann V. Does simulation-based training in medical education need additional stressors? An experimental study. *Ergonomics*. 2020;63(1):80-90. doi:10.1080/00140139.2019.1677948

33. Hagemann V, Herbstreit F, Kehren C, et al. Does teaching non-technical skills to medical students improve those skills and simulated patient outcome? *Int J Med Educ*. 2017;8:101-113. doi:10.5116/ijme.58c1.9f0d

34. Fukushima Y, Yamada A, Imoto N, Iba T. Does TeamSTEPPS affect psychological status? *IJHCQA*. 2019;32(1):11-20. doi:10.1108/IJHCQA-08-2017-0156

35. Hunziker S, Semmer NK, Tschan F, Schuetz P, Mueller B, Marsch S. Dynamics and association of different acute stress markers with performance during a simulated resuscitation. *Resuscitation*. 2012;83(5):572-578. doi:10.1016/j.resuscitation.2011.11.013

36. Pottier P, Hardouin JB, Dejoie T, et al. Effect of Extrinsic and Intrinsic Stressors on Clinical Skills Performance in Third-Year Medical Students. *J GEN INTERN MED*. 2015;30(9):1259-1269. doi:10.1007/s11606-015-3314-6

37. Pottier P, Dejoie T, Hardouin JB, et al. Effect of stress on clinical reasoning during simulated ambulatory consultations. *Medical Teacher*. 2013;35(6):472-480. doi:10.3109/0142159X.2013.774336

38. Bragard I, Etienne AM, Merckaert I, Libert Y, Razavi D. Efficacy of a Communication and Stress Management Training on Medical Residents' Self-efficacy, Stress to Communicate and Burnout: A Randomized Controlled Study. *J Health Psychol*. 2010;15(7):1075-1081. doi:10.1177/1359105310361992

39. Maddox MM, Lopez A, Mandava SH, et al. Electroencephalographic Monitoring of Brain Wave Activity During Laparoscopic Surgical Simulation to Measure Surgeon Concentration and Stress: Can the Student Become the Master? *Journal of Endourology*. 2015;29(12):1329-1333. doi:10.1089/end.2015.0239

40. Fraser K, Ma I, Teteris E, Baxter H, Wright B, McLaughlin K. Emotion, cognitive load and learning outcomes during simulation training: Emotion and cognitive load during simulation. *Medical Education*. 2012;46(11):1055-1062. doi:10.1111/j.1365-2923.2012.04355.x

41. Chang HJ, Kim HJ, Rhyu IJ, Lee YM, Uhm CS. Emotional experiences of medical students during cadaver dissection and the role of memorial ceremonies: a qualitative study. *BMC Med Educ*. 2018;18(1):255. doi:10.1186/s12909-018-1358-0

42. Lim G, Gardner AK. Emotional Intelligence and Delivering Bad News: The Jury is Still Out. *Journal of Surgical Education*. 2019;76(3):779-784. doi:10.1016/j.jsurg.2018.09.017

43. Arora S, Russ S, Petrides KV, et al. Emotional Intelligence and Stress in Medical Students Performing Surgical Tasks: *Academic Medicine*. 2011;86(10):1311-1317. doi:10.1097/ACM.0b013e31822bd7aa

44. Nayar SK, Musto L, Fernandes R, Bharathan R. Emotional Intelligence Predicts Accurate Self-Assessment of Surgical Quality: A Pilot Study. *Journal of Surgical Research*. 2020;245:383-389. doi:10.1016/j.jss.2019.07.051

45. Ogle J, Bushnell JA, Caputi P. Empathy is related to clinical competence in medical care. *Med Educ*. 2013;47(8):824-831. doi:10.1111/medu.12232

46. Deinzer R, Kiupel S, Weik U. Endocrine and psychological stress response in simulated doctor-patient interactions in medical education. *Psychoneuroendocrinology*. 2019;105:172-177. doi:10.1016/j.psyneuen.2018.09.028

47. Keitel A, Ringleb M, Schwartges I, et al. Endocrine and psychological stress responses in a simulated emergency situation. *Psychoneuroendocrinology*. 2011;36(1):98-108. doi:10.1016/j.psyneuen.2010.06.011

48. Bloomfield JG, O'Neill B, Gillett K. Enhancing student communication during end-of-life care: A pilot study. *Pall Supp Care*. 2015;13(6):1651-1661. doi:10.1017/S147895151500022X

49. Waterland P, Khan FS, Ismaili E, Cheruvu C. Environmental Noise as an Operative Stressor During Simulated Laparoscopic Surgery. *Surgical Laparoscopy, Endoscopy & Percutaneous Techniques*. 2016;26(2):133-136. doi:10.1097/SLE.0000000000000250

50. Lewis G, McCullough M, Maxwell AP, Gormley GJ. Ethical reasoning through simulation: a phenomenological analysis of student experience. *Adv Simul*. 2016;1(1):26. doi:10.1186/s41077-016-0027-9

51. Bob MH, Popescu CA, Armean MS, Suciu SM, Buzoianu AD. Ethical views, attitudes and reactions of Romanian medical students to the dissecting room. *Rev Med Chir Soc Med Nat Iasi*. 2014;118(4):1078-1085.

52. Huizinga CRH, de Kam ML, Stockmann HBAC, van Gerven JMA, Cohen AF, van der Bogt KEA. Evaluating Fitness to Perform in Surgical Residents after Night Shifts and Alcohol Intoxication: The development of a "Fit-to-Perform" test. *Journal of Surgical Education*. 2018;75(4):968-977. doi:10.1016/j.jsurg.2018.01.010

53. Dupuy L, Micoulaud-Franchi JA, Cassoudesalle H, et al. Evaluation of a virtual agent to train medical students conducting psychiatric interviews for diagnosing major depressive disorders. *Journal of Affective Disorders*. 2020;263:1-8. doi:10.1016/j.jad.2019.11.117

54. Timberlake MD, Stefanidis D, Gardner AK. Examining the impact of surgical coaching on trainee physiologic response and basic skill acquisition. *Surg Endosc*. 2018;32(10):4183-4190. doi:10.1007/s00464-018-6163-7

55. LaNoue MD, Roter DL. Exploring patient-centeredness: The relationship between self-reported empathy and patient-centered communication in medical trainees. *Patient Education and Counseling*. 2018;101(6):1143-1146. doi:10.1016/j.pec.2018.01.016

56. Cherry MG, Fletcher I, O'Sullivan H. Exploring the relationships among attachment, emotional intelligence and communication: Emotional intelligence and clinical communication. *Medical Education*. 2013;47(3):317-325. doi:10.1111/medu.12115

57. Kerins J, Smith SE, Phillips EC, Clarke B, Hamilton AL, Tallentire VR. Exploring transformative learning when developing medical students' non-technical skills. *Med Educ*. 2020;54(3):264-274. doi:10.1111/medu.14062

58. Behrens CC, Dolmans DH, Gormley GJ, Driessen EW. Exploring undergraduate students achievement emotions during ward round simulation: a mixed-method study. *BMC Med Educ*. 2019;19(1):316. doi:10.1186/s12909-019-1753-1

59. Kim KJ. Factors associated with medical student test anxiety in objective structured clinical examinations: a preliminary study. *Int J Med Educ*. 2016;7:424-427. doi:10.5116/ijme.5845.caec

60. Leboulanger N. First cadaver dissection: Stress, preparation, and emotional experience. *European Annals of Otorhinolaryngology, Head and Neck Diseases*. 2011;128(4):175-183. doi:10.1016/j.anorl.2011.01.007

61. Sevrain-Goideau M, Gohier B, Bellanger W, Annweiler C, Campone M, Coutant R. Forum theater staging of difficult encounters with patients to increase empathy in students: evaluation of efficacy at The University of Angers Medical School. *BMC Med Educ*. 2020;20(1):58. doi:10.1186/s12909-020-1965-4

62. Clarke S, Horeczko T, Cotton D, Bair A. Heart rate, anxiety and performance of residents during a simulated critical clinical encounter: a pilot study. *BMC Med Educ*. 2014;14(1):153. doi:10.1186/1472-6920-14-153

63. Hurst SA, Baroffio A, Ummel M, Layat Burn C. Helping medical students to acquire a deeper understanding of truth-telling. *Medical Education Online*. 2015;20(1):28133. doi:10.3402/meo.v20.28133

64. Finan E, Bismilla Z, Whyte HE, LeBlanc V, McNamara PJ. High-fidelity simulator technology may not be superior to traditional low-fidelity equipment for neonatal resuscitation training. *J Perinatol*. 2012;32(4):287-292. doi:10.1038/jp.2011.96

65. Zhou Y, Collinson A, Laidlaw A, Humphris G. How Do Medical Students Respond to Emotional Cues and Concerns Expressed by Simulated Patients during OSCE Consultations? – A Multilevel Study. Houser D, ed. *PLoS ONE*. 2013;8(10):e79166. doi:10.1371/journal.pone.0079166

66. Meunier J, Libert Y, Merckaert I, et al. How much is residents' distress detection performance during a clinical round related to their characteristics? *Patient Education and Counseling*. 2011;85(2):180-187. doi:10.1016/j.pec.2010.09.022

67. Hulsman RL, Pranger S, Koot S, Fabriek M, Karemaker JM, Smets EMA. How stressful is doctor–patient communication? Physiological and psychological stress of medical students in simulated history taking and bad-news consultations. *International Journal of Psychophysiology*. 2010;77(1):26-34. doi:10.1016/j.ijpsycho.2010.04.001

68. Larkin AC, Cahan MA, Whalen G, et al. Human Emotion and Response in Surgery (HEARS): A Simulation-Based Curriculum for Communication Skills, Systems-Based Practice, and Professionalism in Surgical Residency Training. *Journal of the American College of Surgeons*. 2010;211(2):285-292. doi:10.1016/j.jamcollsurg.2010.04.004

69. Torres A, Horodeńska M, Witkowski G, Torres K. Hybrid simulation of pediatric gynecologic examination: a mix-methods study of learners' attitudes and factors affecting learning. *BMC Med Educ*. 2020;20(1):164. doi:10.1186/s12909-020-02076-7

70. Hunziker S, Pagani S, Fasler K, Tschan F, Semmer NK, Marsch S. Impact of a stress coping strategy on perceived stress levels and performance during a simulated cardiopulmonary resuscitation: a randomized controlled trial. *BMC Emerg Med*. 2013;13(1):8. doi:10.1186/1471-227X-13-8

71. Bajunaid K, Mullah MAS, Winkler-Schwartz A, et al. Impact of acute stress on psychomotor bimanual performance during a simulated tumor resection task. *JNS*. 2017;126(1):71-80. doi:10.3171/2015.5.JNS15558

72. Piquette D, Tarshis J, Sinuff T, Fowler RA, Pinto R, Leblanc VR. Impact of Acute Stress on Resident Performance During Simulated Resuscitation Episodes: A Prospective Randomized Cross-Over Study. *Teaching and Learning in Medicine*. 2014;26(1):9-16. doi:10.1080/10401334.2014.859932

73. Harvey A, Bandiera G, Nathens AB, LeBlanc VR. Impact of stress on resident performance in simulated trauma scenarios. *Journal of Trauma and Acute Care Surgery*. 2012;72(2):497-503. doi:10.1097/TA.0b013e31821f84be

74. Wasp GT, Cullinan AM, Chamberlin MD, Hayes C, Barnato AE, Vergo MT. Implementation and Impact of a Serious Illness Communication Training for Hematology-Oncology Fellows. *J Canc Educ*. 2021;36(6):1325-1332. doi:10.1007/s13187-020-01772-8

75. Bragard I, Seghaye MC, Farhat N, et al. Implementation of a 2-Day Simulation-Based Course to Prepare Medical Graduates on Their First Year of Residency: *Pediatric Emergency Care*. 2018;34(12):857-861. doi:10.1097/PEC.0000000000000930

76. Le Lous M, De Chanaud N, Bourret A, et al. Improving the quality of transvaginal ultrasound scan by simulation training for general practice residents. *Adv Simul*. 2017;2(1):24. doi:10.1186/s41077-017-0056-z

77. Beckers SK, Biermann H, Sopka S, et al. Influence of pre-course assessment using an emotionally activating stimulus with feedback: A pilot study in teaching Basic Life Support. *Resuscitation*. 2012;83(2):219-226. doi:10.1016/j.resuscitation.2011.08.024

78. Russa AD, Mligiliche NuruL. Inspiring Tanzanian medical students into the profession: appraisal of cadaveric dissection stress and coping strategies. *Italian Journal of Anatomy and Embryology*. 2015;Vol 119:268-276 Pages, 276 kB. doi:10.13128/IJAE-15560

79. Nelson S, Germann C, MacVane C, Bloch R, Fallon T, Strout T. Intern as Patient: A Patient Experience Simulation to Cultivate Empathy in Emergency Medicine Residents. *WestJEM*. Published online January 18, 2018:41-48. doi:10.5811/westjem.2017.11.35198

80. McNeer RR, Bennett CL, Dudaryk R. Intraoperative Noise Increases Perceived Task Load and Fatigue in Anesthesiology Residents: A Simulation-Based Study. *Anesthesia & Analgesia*. 2016;122(2):512-525. doi:10.1213/ANE.0000000000001067

81. Robb A, Kopper R, Ambani R, et al. Leveraging Virtual Humans to Effectively Prepare Learners for Stressful Interpersonal Experiences. *IEEE Trans Visual Comput Graphics*. 2013;19(4):662-670. doi:10.1109/TVCG.2013.35

82. Singh S, Modi C, Patel C, Pathak A. Low-fidelity simulation to enhance understanding of infection control among undergraduate medical students. *Natl Med J India*. 2017;30(4):215. doi:10.4103/0970-258X.218677

83. Dedmon MM, O'Connell BP, Yawn RJ, et al. Measuring Mental Stress During Otologic Surgery Using Heart Rate Variability Analysis. *Otology & Neurotology*. 2019;40(4):529-534. doi:10.1097/MAO.0000000000002187

84. Tiuraniemi J, Läärä R, Kyrö T, Lindeman S. Medical and psychology students' self-assessed communication skills: A pilot study. *Patient Education and Counseling*. 2011;83(2):152-157. doi:10.1016/j.pec.2010.05.013

85. Toivonen AK, Lindblom-Ylänne S, Louhiala P, Pyörälä E. Medical students' reflections on emotions concerning breaking bad news. *Patient Education and Counseling*. 2017;100(10):1903-1909. doi:10.1016/j.pec.2017.05.036

86. Berg K, Majdan JF, Berg D, Veloski J, Hojat M. Medical Students' Self-Reported Empathy and Simulated Patients' Assessments of Student Empathy: An Analysis by Gender and Ethnicity: *Academic Medicine*. 2011;86(8):984-988. doi:10.1097/ACM.0b013e3182224f1f

87. Klein MI, Mouraviev V, Craig C, et al. Mental stress experienced by first-year residents and expert surgeons with robotic and laparoscopic surgery interfaces. *J Robotic Surg*. 2014;8(2):149-155. doi:10.1007/s11701-013-0446-8

88. Klein MI, Warm JS, Riley MA, et al. Mental Workload and Stress Perceived by Novice Operators in the Laparoscopic and Robotic Minimally Invasive Surgical Interfaces. *Journal of Endourology*. 2012;26(8):1089-1094. doi:10.1089/end.2011.0641

89. Noll C, von Jan U, Raap U, Albrecht UV. Mobile Augmented Reality as a Feature for Self-Oriented, Blended Learning in Medicine: Randomized Controlled Trial. *JMIR Mhealth Uhealth*. 2017;5(9):e139. doi:10.2196/mhealth.7943

90. Naismith LM, Lajoie SP. Motivation and emotion predict medical students' attention to computer-based feedback. *Adv in Health Sci Educ*. 2018;23(3):465-485. doi:10.1007/s10459-017-9806-x

91. Anyanwu GE, Nto JN, Agu AU, Ekezie J, Esom EA. Musical preferences and learning outcome of medical students in cadaver dissection laboratory: A Nigerian survey. *Annals of Anatomy - Anatomischer Anzeiger*. 2016;208:228-233. doi:10.1016/j.aanat.2016.07.010

92. Alt-Epping B, Lohse C, Viebahn C, Steinbüchel N von, Benze G, Nauck F. On death and dying – an exploratory and evaluative study of a reflective, interdisciplinary course element in undergraduate anatomy teaching. *BMC Med Educ*. 2014;14(1):15. doi:10.1186/1472-6920-14-15

93. Anton NE, Bean EA, Myers E, Stefanidis D. Optimizing learner engagement during mental skills training: A pilot study of small group vs. individualized training. *The American Journal of Surgery*. 2020;219(2):335-339. doi:10.1016/j.amjsurg.2019.12.022

94. Goldberg MB, Mazzei M, Maher Z, et al. Optimizing performance through stress training — An educational strategy for surgical residents. *The American Journal of Surgery*. 2018;216(3):618-623. doi:10.1016/j.amjsurg.2017.11.040

95. Smink GM, Jeffe DB, Hayashi RJ, Al-Hammadi N, Fehr JJ. Paediatric oncology simulation training for resident education. *BMJ STEL*. 2019;5(3):155-160. doi:10.1136/bmjstel-2018-000347

96. Schwill S, Fahrbach-Veeser J, Moeltner A, et al. Peers as OSCE assessors for junior medical students – a review of routine use: a mixed methods study. *BMC Med Educ*. 2020;20(1):17. doi:10.1186/s12909-019-1898-y

97. Khairat S, Coleman C, Ottmar P, Bice T, Koppel R, Carson SS. Physicians' gender and their use of electronic health records: findings from a mixed-methods usability study. *Journal of the American Medical Informatics Association*. 2019;26(12):1505-1514. doi:10.1093/jamia/ocz126

98. Geeraerts T, Roulleau P, Cheisson G, et al. Physiological and self-assessed psychological stress induced by a high fidelity simulation course among third year anesthesia and critical care residents: An observational study. *Anaesthesia Critical Care & Pain Medicine*. 2017;36(6):403-406. doi:10.1016/j.accpm.2017.06.002

99. Phitayakorn R, Minehart RD, Pian-Smith MCM, Hemingway MW, Petrusa ER. Practicality of using galvanic skin response to measure intraoperative physiologic autonomic activation in operating room team members. *Surgery*. 2015;158(5):1415-1420. doi:10.1016/j.surg.2015.04.009

100. Cheung EO, Barsuk JH, Mitra D, et al. Preliminary Efficacy of a Brief Mindfulness Intervention for Procedural Stress in Medical Intern Simulated Performance: A Randomized Controlled Pilot Trial. *The Journal of Alternative and Complementary Medicine*. 2020;26(4):282-290. doi:10.1089/acm.2019.0209

101. Bernhardt V, Rothkötter HJ, Kasten E. Psychische Belastungen durch die Dissektion am Leichnam im anatomischen Präparierkurs bei Erstsemestern des Studienfachs Medizin. *GMS Zeitschrift für Medizinische Ausbildung; 29(1):Doc12; ISSN 1860-3572*. Published online 2012. doi:10.3205/ZMA000782

102. Böckers A, Baader C, Fassnacht UK, Öchsner W, Böckers TM. Reduction of mental distress in the dissection course by introducing the body donor experience through anatomical demonstrations of organ systems. *Anat Sci Ed*. 2012;5(6):321-329. doi:10.1002/ase.1292

103. Łoś K, Chmielewski J, Łuczyński W. Relationship between Executive Functions, Mindfulness, Stress, and Performance in Pediatric Emergency Simulations. *IJERPH*. 2020;17(6):2040. doi:10.3390/ijerph17062040

104. Bourgeon L, Bensalah M, Vacher A, Ardouin JC, Debien B. Role of emotional competence in residents' simulated emergency care performance: a mixed-methods study. *BMJ Qual Saf*. 2016;25(5):364-371. doi:10.1136/bmjqs-2015-004032

105. Marvin RK, Saepoo MB, Ye S, et al. Salivary protein changes in response to acute stress in medical residents performing advanced clinical simulations: a pilot proteomics study. *Biomarkers*. 2017;22(3-4):372-382. doi:10.1080/1354750X.2017.1279215

106. Rogers T, Andler C, O'Brien B, van Schaik S. Self-Reported Emotions in Simulation-Based Learning: Active Participants vs. Observers. *Sim Healthcare*. 2019;14(3):140-145. doi:10.1097/SIH.0000000000000354

107. Hautz WE, Schröder T, Dannenberg KA, et al. Shame in Medical Education: A Randomized Study of the Acquisition of Intimate Examination Skills and Its Effect on Subsequent Performance. *Teaching and Learning in Medicine*. 2017;29(2):196-206. doi:10.1080/10401334.2016.1254636

108. Groot F, Jonker G, Rinia M, ten Cate O, Hoff RG. Simulation at the Frontier of the Zone of Proximal Development: A Test in Acute Care for Inexperienced Learners. *Academic Medicine*. 2020;95(7):1098-1105. doi:10.1097/ACM.0000000000003265

109. Baker BG, Bhalla A, Doleman B, et al. Simulation fails to replicate stress in trainees performing a technical procedure in the clinical environment. *Medical Teacher*. 2017;39(1):53-57. doi:10.1080/0142159X.2016.1230188

110. Butteris SM, Gladding SP, Eppich W, Hagen SA, Pitt MB. Simulation Use for Global Away Rotations (SUGAR): Preparing Residents for Emotional Challenges Abroad—A Multicenter Study. *Academic Pediatrics*. 2014;14(5):533-541. doi:10.1016/j.acap.2014.05.004

111. Crewther BT, Shetty K, Jarchi D, et al. Skill acquisition and stress adaptations following laparoscopic surgery training and detraining in novice surgeons. *Surg Endosc*. 2016;30(7):2961-2968. doi:10.1007/s00464-015-4584-0

112. Lefroy J, Brosnan C, Creavin S. Some like it hot: medical student views on choosing the emotional level of a simulation: Student views on emotion in simulation. *Medical Education*. 2011;45(4):354-361. doi:10.1111/j.1365-2923.2010.03881.x

113. Hurley AM, Kennedy PJ, O'Connor L, et al. SOS save our surgeons: Stress levels reduced by robotic surgery. *Gynecol Surg*. 2015;12(3):197-206. doi:10.1007/s10397-015-0891-7

114. LaPorta AJ, McKee J, Hoang T, et al. Stress Inoculation: Preparing Outside the Box in Surgical Resuscitation and Education. *Curr Trauma Rep*. 2017;3(2):135-143. doi:10.1007/s40719-017-0090-2

115. Daglius Dias R, Scalabrini Neto A. Stress levels during emergency care: A comparison between reality and simulated scenarios. *Journal of Critical Care*. 2016;33:8-13. doi:10.1016/j.jcrc.2016.02.010

116. Pottier P, Hardouin JB, Dejoie T, et al. Stress responses in medical students in ambulatory and in-hospital patient consultations: Stress responses in medical students. *Medical Education*. 2011;45(7):678-687. doi:10.1111/j.1365-2923.2011.03935.x

117. Maher Z, Milner R, Cripe J, Gaughan J, Fish J, Goldberg AJ. Stress training for the surgical resident. *The American Journal of Surgery*. 2013;205(2):169-174. doi:10.1016/j.amjsurg.2012.10.007

118. Oriot D, Trigolet M, Kessler DO, Auerbach MA, Ghazali DA. Stress: A Factor Explaining the Gap Between Simulated and Clinical Procedure Success. *Pediatr Emer Care*. 2021;37(12):e1192-e1196. doi:10.1097/PEC.0000000000001962

119. Brouwers MH, Bor H, Laan R, van Weel C, van Weel-Baumgarten E. Students' experiences with a longitudinal skills training program on breaking bad news: A follow-up study. *Patient Education and Counseling*. 2018;101(9):1639-1644. doi:10.1016/j.pec.2018.05.008

120. Ruiz-Moral R, Pérula de Torres L, Monge D, García Leonardo C, Caballero F. Teaching medical students to express empathy by exploring patient emotions and experiences in standardized medical encounters. *Patient Education and Counseling*. 2017;100(9):1694-1700. doi:10.1016/j.pec.2017.04.018

121. Johnson JM, Stern TA. Teaching Residents About Emotional Intelligence and Its Impact on Leadership. *Acad Psychiatry*. 2014;38(4):510-513. doi:10.1007/s40596-014-0048-4

122. Nicolaides M, Theodorou E, Emin EI, et al. Team performance training for medical students: Low vs high fidelity simulation. *Annals of Medicine and Surgery*. 2020;55:308-315. doi:10.1016/j.amsu.2020.05.042

123. Fraser K, McLaughlin K. Temporal pattern of emotions and cognitive load during simulation training and debriefing. *Medical Teacher*. 2019;41(2):184-189. doi:10.1080/0142159X.2018.1459531

124. Modi HN, Singh H, Orihuela-Espina F, et al. Temporal Stress in the Operating Room: Brain Engagement Promotes "Coping" and Disengagement Prompts "Choking." *Annals of Surgery*. 2018;267(4):683-691. doi:10.1097/SLA.0000000000002289

125. Kromann CB, Jensen ML, Ringsted C. Test-enhanced learning may be a gender-related phenomenon explained by changes in cortisol level: Test-enhanced learning, gender and cortisol. *Medical Education*. 2011;45(2):192-199. doi:10.1111/j.1365-2923.2010.03790.x

126. Flinn JT, Miller A, Pyatka N, Brewer J, Schneider T, Cao CGL. The effect of stress on learning in surgical skill acquisition. *Medical Teacher*. 2016;38(9):897-903. doi:10.3109/0142159X.2015.1114597

127. Bong CL, Lee S, Ng ASB, Allen JC, Lim EHL, Vidyarthi A. The effects of active (hot-seat) versus observer roles during simulation-based training on stress levels and non-technical performance: a randomized trial. *Adv Simul*. 2017;2(1):7. doi:10.1186/s41077-017-0040-7

128. O'Rourke SR, Branford KR, Brooks TL, Ives LT, Nagendran A, Compton SN. The Emotional and Behavioral Impact of Delivering Bad News to Virtual versus Real Standardized Patients: A Pilot Study. *Teaching and Learning in Medicine*. 2020;32(2):139-149. doi:10.1080/10401334.2019.1652180

129. Fraser K, Huffman J, Ma I, et al. The Emotional and Cognitive Impact of Unexpected Simulated Patient Death. *Chest*. 2014;145(5):958-963. doi:10.1378/chest.13-0987

130. Cook CJ, Crewther BT. The impact of a competitive learning environment on hormonal and emotional stress responses and skill acquisition and expression in a medical student domain. *Physiology & Behavior*. 2019;199:252-257. doi:10.1016/j.physbeh.2018.11.027

131. Lizotte MH, Janvier A, Latraverse V, et al. The Impact of Neonatal Simulations on Trainees' Stress and Performance: A Parallel-Group Randomized Trial*. *Pediatric Critical Care Medicine*. 2017;18(5):434-441. doi:10.1097/PCC.0000000000001119

132. Schweller M, Costa FO, Antônio MÂRGM, Amaral EM, de Carvalho-Filho MA. The Impact of Simulated Medical Consultations on the Empathy Levels of Students at One Medical School: *Academic Medicine*. 2014;89(4):632-637. doi:10.1097/ACM.0000000000000175

133. DeMaria S, Silverman ER, Lapidus KAB, et al. The impact of simulated patient death on medical students' stress response and learning of ACLS. *Medical Teacher*. 2016;38(7):730-737. doi:10.3109/0142159X.2016.1150986

134. Andreatta PB, Hillard M, Krain LP. The impact of stress factors in simulation-based laparoscopic training. *Surgery*. 2010;147(5):631-639. doi:10.1016/j.surg.2009.10.071

135. Nachiappan M, Bikramjit P, Aung WT, Htoo HKS, Sudipta P. The impact of stressors on the learning outcome of high-fidelity patient simulation in undergraduate medical students. *Med J Malaysia*. 2020;75(3):209-215.

136. Hochberg MS, Berman RS, Kalet AL, Zabar SR, Gillespie C, Pachter HL. The stress of residency: recognizing the signs of depression and suicide in you and your fellow residents. *The American Journal of Surgery*. 2013;205(2):141-146. doi:10.1016/j.amjsurg.2012.08.003

137. Chia TI, Oyeniran OI, Ajagbe AO, Onigbinde OA, Oraebosi MI. The symptoms and stress experienced by medical students in anatomy dissection halls. *Journal of Taibah University Medical Sciences*. 2020;15(1):8-13. doi:10.1016/j.jtumed.2020.01.001

138. Harvey A, Nathens AB, Bandiera G, LeBlanc VR. Threat and challenge: cognitive appraisal and stress responses in simulated trauma resuscitations: Appraisal and stress in simulated trauma resuscitations. *Medical Education*. 2010;44(6):587-594. doi:10.1111/j.1365-2923.2010.03634.x

139. Cerrone SA, Adelman P, Akbar S, Yacht AC, Fornari A. Using Objective Structured Teaching Encounters (OSTEs) to prepare chief residents to be emotionally intelligent leaders. *Medical Education Online*. 2017;22(1):1320186. doi:10.1080/10872981.2017.1320186

140. Cherry MG, Fletcher I, O'Sullivan H. Validating relationships among attachment, emotional intelligence and clinical communication. *Med Educ*. 2014;48(10):988-997. doi:10.1111/medu.12526

141. Dohms MC, Collares CF, Tibério IC. Video-based feedback using real consultations for a formative assessment in communication skills. *BMC Med Educ*. 2020;20(1):57. doi:10.1186/s12909-020-1955-6

Table 1. Search Terms for Scoping Review

Medline	Embase
1. Resident*.tw,kf.	1. Resident*.tw,kw.
2. Residenc*.tw,kf.	2. Residenc*.tw,kw.
3. (Intern of interns*).tw,kf.	3. (Intern or Interns*).tw,kw.
4. Trainee*.tw,kf.	4. Trainee*.tw,kw.
5. Med* Student*.tw,kf.	5. Med* Student*.tw,kw.
6. Fellow*.tw,kf.	6. Fellow*.tw,kw.
7. 1 or 2 or 3 or 4 or 5 or 6	7. 1 or 2 or 3 or 4 or 5 or 6
8. Emotions/	8. Emotions/
9. Stress, Psychological/	9. Emotional Intelligence/
10. Emotional Intelligence/	10. Emotional Stress/
11. Emotion*.tw,kf.	11. Mental Stress/
12. Stress*.tw,kf	12. Emotion*.tw,kw.
13. mood*.tw,kf.	13. Stress*.tw,kw.
14. 8 or 9 or 10 or 11 or 12 or 13	14. Mood*.tw,kw.
15. 7 and 14	15. 8 or 9 or 10 or 11 or 12 or 13 or 14
16. 15 and 2010:2021.(sa_year).	16. 7 and 15
	17. 16 and 2010:2020.(sa_year).
This search yielded 11,653 hits. The search was executed on June 22nd, 2020.	This search yielded 17,676 hits. The search was executed on June 22nd, 2020.

Table 2. Inclusion and Exclusion Criteria for Title-Abstract Screening

Inclusion criteria	Exclusion criteria
Empirical study, reviews, conference proceedings, commentaries, editorials, conference papers	Magazine articles, one-off diagrams, supplementary articles, theses, dissertations, abstracts
Medical students, interns, residents, and fellows (if any of these are included)	Nurses, dentistry students, paramedics, physicians, doctors, patients (if only these are exclusively included)
Emotions, mood, stress, and emotional intelligence	Depression, burnout, attitudes.

Table 3. Inclusion and Exclusion Criteria for the First Full-Text Screening

Inclusion criteria	Exclusion criteria
English articles	Non-English articles
Empirical studies	Review, editorial, opinion pieces, conference abstracts, conference listings, authors' response, letter to the editor
Medical students, interns, residents, and fellows (if any of these are included)	Nurses, dentistry students, paramedics, physicians, doctors, patients (if only these are exclusively included
Emotions, emotion regulation, discrete emotions (e.g., anxiety, anger, sadness)	Attitudes, personality
Mood	Depression, post-traumatic stress disorder, mood disorders
Stress, stressors, distress, stress coping (regulation), physiological indicators of stress	Burnout, emotional exhaustion, depersonalization, empathy (if only these are exclusively included)
Emotional intelligence	

Table 4. Inclusion and Exclusion Criteria for the Second Full-Text Screening

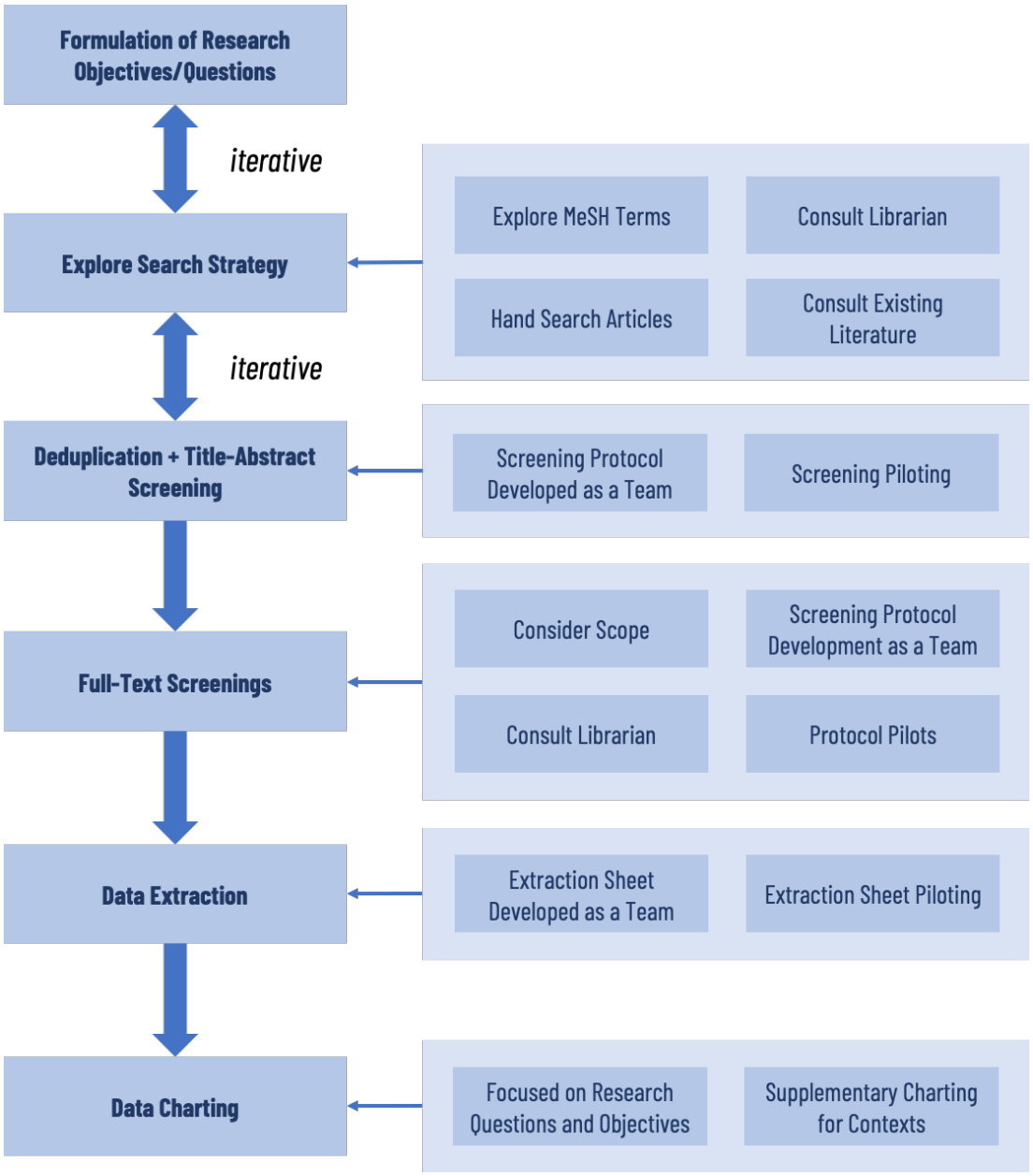
Inclusion criteria	Exclusion criteria
There is at least one simulator mentioned: cadavers, mannequins, screen-based simulators, VR/AR simulator, role playing, standardized patients, task-trainer.	There are no simulators mentioned
Objective Structured Clinical Examination (OSCE), Fundamentals of Laparoscopic Surgery (FLS) training, "cases"	Non-medical simulations, such as computer simulations (e.g., simulations of data), flight simulations, computer games (that are not focused on medical-related tasks)
Participants: Medical students, interns, residents, and fellows (if any of these are included)	Nurses, dentistry students, paramedics, physicians, doctors, patients (if only these are exclusively included

Appendix A

Review Protocol

Below is the general protocol we followed for our scoping review, based on Arksey and O'Malley's framework[1] for conducting scoping reviews. We did not register our protocol.

Figure A1. General Protocol.



Appendix B

Table B1. PRISMA Extension for Scoping Reviews (PRISMA-ScR) Checklist

Section	Item	PRISMA-ScR Checklist Item	Check Results
Title	1	Identify the report as a scoping review.	Yes (Manuscript)
Abstract			
Structured Summary	2	Provide a structured summary that includes (as applicable) background, objectives, eligibility criteria, sources of evidence, charting methods, results, and conclusions that relate to the review questions and objectives.	Yes (Manuscript)
Introduction			
Rationale	3	Describe the rationale for the review in the context of what is already known. Explain why the review questions/objectives lend themselves to a scoping review approach.	Yes (Manuscript)
Objectives	4	Provide an explicit statement of the questions and objectives being addressed with reference to their key elements (e.g., population or participants, concepts, and context) or other relevant key elements used to conceptualize the review questions and/or objectives.	Yes (Manuscript)
Methods			
Protocol and registration	5	Indicate whether a review protocol exists; state if and where it can be accessed (e.g., a Web address); and if available, provide registration information, including the registration number.	Yes (Appendix A)
Eligibility criteria	6	Specify characteristics of the sources of evidence used as eligibility criteria (e.g., years considered, language, and publication status), and provide a rationale.	Yes (Manuscript)
Information sources	7	Describe all information sources in the search (e.g., databases with dates of coverage and contact with authors to identify additional sources), as well as the date the most recent search was executed.	Yes (Manuscript)

Search	8	Present the full electronic search strategy for at least 1 database, including any limits used, such that it could be repeated.	Yes (Manuscript)
Selection of sources of evidence	9	State the process for selecting sources of evidence (i.e., screening and eligibility) included in the scoping review.	Yes (Manuscript)
Data charting process	10	Describe the methods of charting data from the included sources of evidence (e.g., calibrated forms or forms that have been tested by the team before their use, and whether data charting was done independently or in duplicate) and any processes for obtaining and confirming data from investigators.	Yes (Appendix C)
Data items	11	List and define all variables for which data were sought and any assumptions and simplifications made.	Yes (Appendix D)
Critical appraisal of individual sources of evidence	12	If done, provide a rationale for conducting a critical appraisal of included sources of evidence; describe the methods used and how this information was used in any data synthesis (if appropriate).	Not applicable
Summary measures	13	Not applicable for scoping reviews.	Not applicable
Synthesis of results	14	Describe the methods of handling and summarizing the data that were charted.	Yes (Manuscript)
Risk of bias across studies	15	Not applicable for scoping reviews.	Not applicable
Additional analyses	16	Not applicable for scoping reviews.	Not applicable
Results			
Selection of sources of evidence	17	Give numbers of sources of evidence screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally using a flow diagram.	Yes (Manuscript)
Characteristics of sources of evidence	18	For each source of evidence, present characteristics for which data were charted and provide the citations.	Yes (Manuscript). However, due to the scope of the review, not all evidence were examined

			thoroughly in detail in the main manuscript
Critical appraisal within sources of evidence	19	If done, present data on critical appraisal of included sources of evidence (see item 12).	Not applicable
Results of individual sources of evidence	20	For each included source of evidence, present the relevant data that were charted that relate to the review questions and objectives.	Yes (Manuscript) However, same caution from item 18 applies here.
Synthesis of results	21	Summarize and/or present the charting results as they relate to the review questions and objectives.	Yes (Manuscript)
Risk of bias across studies	22	Not applicable for scoping reviews.	Not applicable
Additional analyses	23	Not applicable for scoping reviews.	Not applicable
Discussion			
Summary of evidence	24	Summarize the main results (including an overview of concepts, themes, and types of evidence available), link to the review questions and objectives, and consider the relevance to key groups.	Yes (Manuscript)
Limitations	25	Discuss the limitations of the scoping review process.	Yes (Manuscript)
Conclusions	26	Provide a general interpretation of the results with respect to the review questions and objectives, as well as potential implications and/or next steps.	Yes (Manuscript)
Funding	27	Describe sources of funding for the included sources of evidence, as well as sources of funding for the scoping review. Describe the role of the funders of the scoping review.	Not applicable (no funding)

Appendix C

Data Charting Process

We have developed a data extraction sheet so that we could systematically extract data for charting. We ran pilots of our data extraction sheet, where the first three pilots included six reviewers. However, when it came to extracting data that were more inference-based (e.g., deciding what construct the article focused on) we realized that our interrater reliability was below the cut-off point of 75%. We therefore decided to continue collecting feedback from our reviewers, and refine our extraction sheet, while we also narrowed down the number of our reviewers. Our fifth and final pilot included four reviewers. BA (the first author of this paper) was the lead reviewer for this project, and so we decided to compare others' data extraction results to BA's and calculate the number of Excel cells that significantly differed from his extraction. Reviewers CL, EB, and MM yielded agreement ratings of 88.81%, 91.53%, and 95.59% with BA. We in the end decided that MM and BA would extract the data from the articles.

Appendix D

Study Characteristics

Below is a figure (Figure D1) that shows the articles published by year and illustrates research in emotions and related constructs have been increasing over recent years.

Figure D1. Articles Published by Year

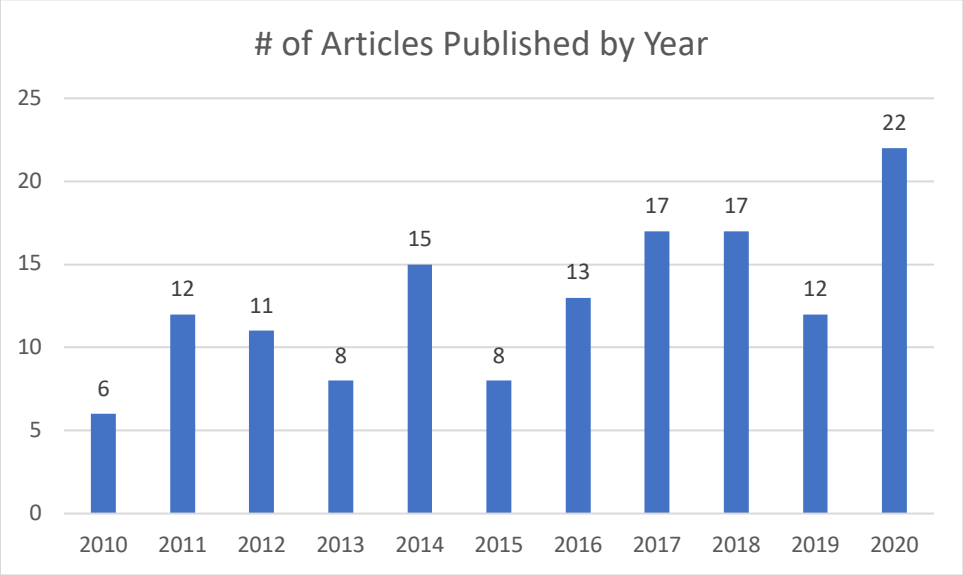
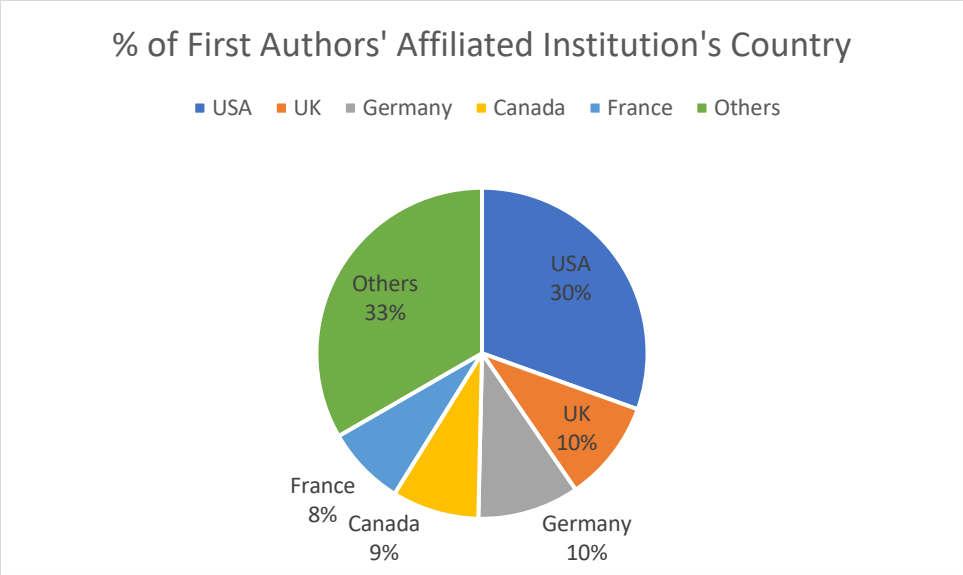


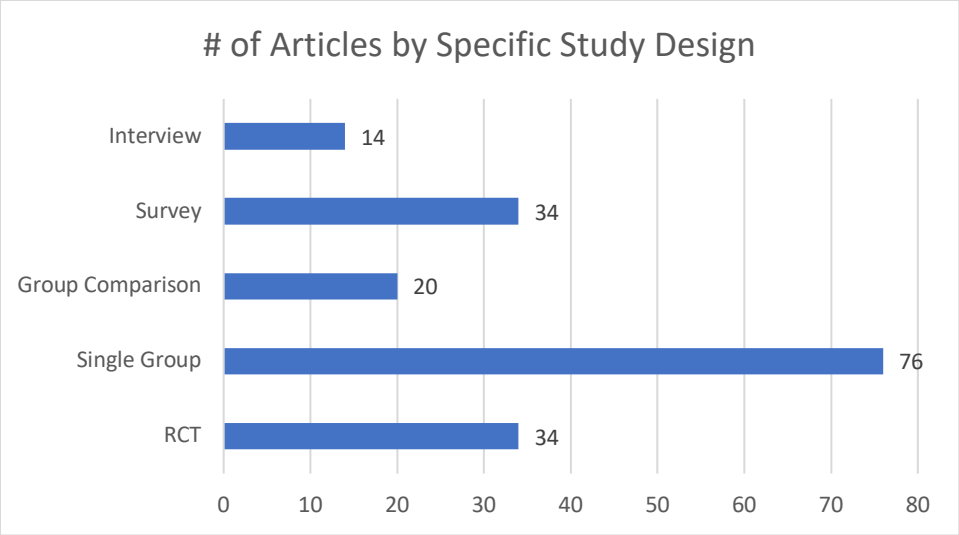
Figure D2 shows the percentage of first/corresponding authors' affiliated institution's country. The United States (30.5%; 43 articles), United Kingdom (9.9%; 14 articles), and Germany (9.9%; 14 articles) were the most common countries affiliated with the first author.

Figure D2. Percentage of First (or Corresponding) Authors' Affiliated Institution's Country



Our results also showed that the vast majority (83.0%; 117 articles) were quantitative studies, while only 5.0% (7 articles) and 12.1% (17 articles) were qualitative and mixed/multi-method studies, respectively. Figure D3 illustrates articles featuring specific study designs. It should be noted that an article may feature multiple specific study designs simultaneously (e.g., have a focus group interview, along with a group comparison design). While 34 studies (24.1%) featured a randomized controlled trial (RCT) design, the single group design (53.9%; 76 articles) was the most employed design.

Figure D3. Number of articles by specific study design



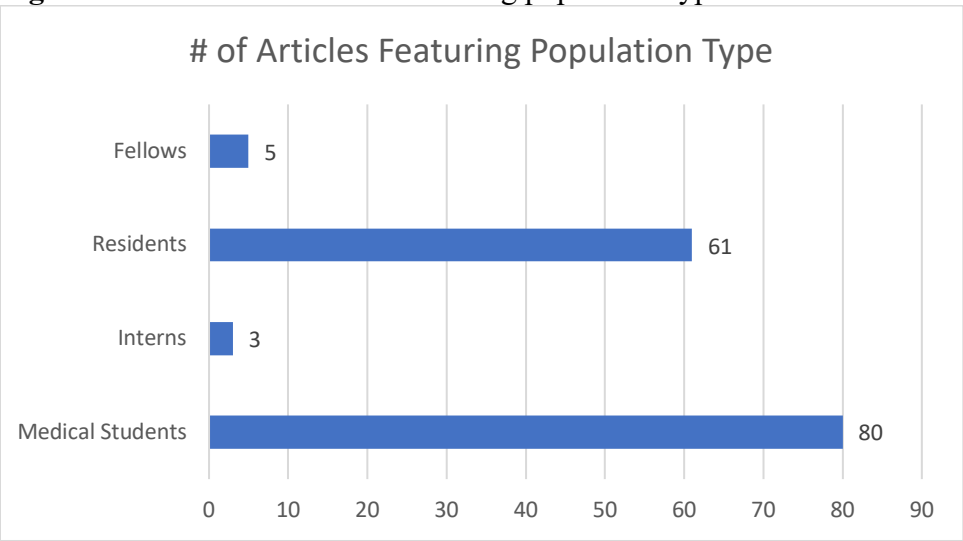
For interpreting data from Figure D2 and D3, it should be noted that there were two types of categorizations we have made with the articles in our review when it came to study design: broad and specific. The former focused on categorizing articles based on whether they applied quantitative, qualitative, or mixed/multi-method analyses. We acknowledge that there is a distinction between mixed and multi-method[2], but have opted to group them up together as doing so still gives us a clear picture concerning whether the literature was dominated by quantitative methods.

The latter categorization aimed to delve into the specific types of study design. While we did start out with various terms guided by work from Turner and others[3], these terms were eventually replaced by the terms produced through the iterative refinement process from our pilot phases; these were based on some of the co-authors' experience conducting scoping reviews[4]: 1) Single Group: If all the participants experienced the same training/education; 2) Group Comparison: If there are two or more groups and are being compared to each other; 3) RCT (Randomized Controlled Trials): If there is randomization and a control group for comparison (regardless of blinding). This includes "single group" studies if participants were

randomly assigned to specific sub groups and counterbalanced for comparison (i.e.,, randomized crossover study), despite the fact that, technically, the students were experiencing the same training in the end; 4) Survey Study: if the study solely relies on a survey sent out to students to poll data, without featuring any intervention or educational activity in their study itself (e.g., ask about a simulation experience but not have the students go through the simulation); and 5) Interview: Studies that involve interviewing. This included studies that featured focus groups. With these terms our reviewers all agreed upon, we were able to conduct the data extraction consistently.

Figure D4 shows the type of relevant populations featured in the articles (again, studies can feature multiple types simultaneously). Medical students (56.7%; 80 articles) and residents (43.3%; 61 articles) made up the bulk of the populations featured. Interns and fellows only represented 3 (2.1%) and 5 articles (3.5%), respectively. We further note that 18 articles (12.8%) featured populations that were not relevant to our scope (e.g., physicians, nursing students).

Figure D4. Number of articles featuring population type



Figures D5 and D6 exhibit the number of articles featuring the constructs our scope sought to capture. Stress was the construct that was featured the most (62.4%; 88 articles), followed by emotions (31.9%; 45 articles), and emotional intelligence (24.1%; 34 articles). There was only 1 article (0.7%) that featured mood. As figure 7 shows, we note that there were only 2 articles (1.4%) that examined emotions, stress, *and* emotional intelligence; 3 articles (2.1%) that looked at both stress and emotional intelligence, and 2 articles (1.4%) that looked at both emotions and emotional intelligence. There were 18 articles that looked at both emotions and stress (12.8%).

Figure D5. Number of times Mood, EI, Stress, or Emotions Featured in the Articles

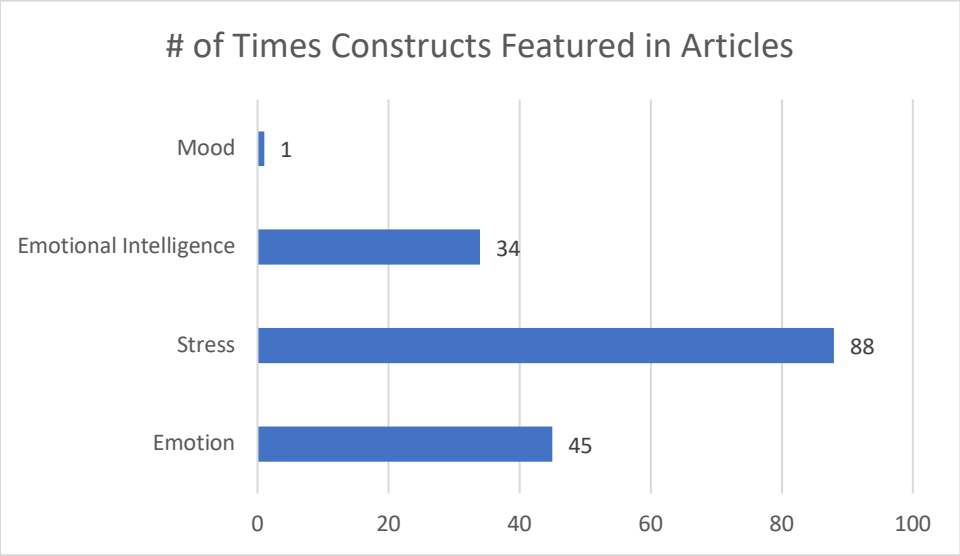
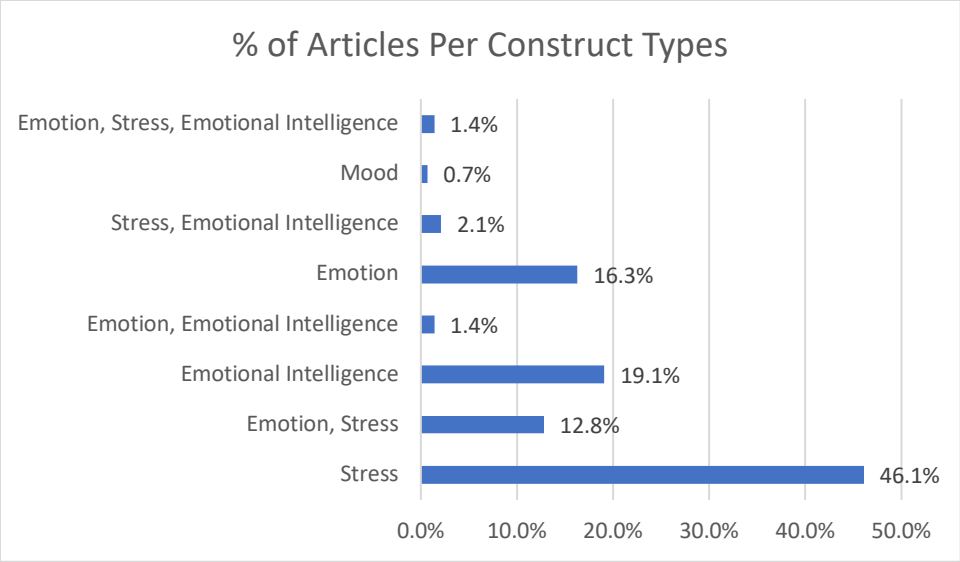


Figure D6. Percentage of Articles by Combination of Constructs Featured



Regarding simulation characteristics, we identified 90 (63.8%) articles that featured simulations done individually, with 49 (34.8%) featuring single-profession teams, and 9 (6.4%) featuring interprofessional teams. We also identified 35 (24.8%) articles that focused on both technical and non-technical skills, 63 articles (45.4%) that focused on just technical skills, and 36 articles (25.5%) that focused on just non-technical skills. For the type of simulators, we found that use of standardized participants (49 articles; 34.8%) and mannequins (48 articles; 34.0%) were the most common. See Table D1 below.

Table D1. Types of Simulators Featured

Simulator Type	**Count**	**%**
Role Playing	28	19.9%
Standardized Participants (SPs)	49	34.8%
Task Trainer	26	18.4%
Screen-Based	9	6.4%
AR/VR Based (with head-sets)	5	3.5%

Mannequin	48	34.0%
Human Cadavers	12	8.5%
Animal Cadavers	1	0.7%
Live Animals	1	0.7%

We note that we differentiated the use of the Role Playing and Standardized Participants (SPs). We counted the article towards former if the simulation featured participants (not SPs) role playing (e.g., a medical student role playing as a family member of the patient). Therefore, even though SPs are technically role-playing, if the simulation did not have participants role-playing, SP articles did not count towards role playing.

Further, while full-body mannequins were counted towards the mannequin category in our review, if it was just a part of a mannequin (e.g., a limb of a mannequin), we considered it to be a task trainer, not a mannequin. Our term Screen-Based simulator only accounted for simulators that had digital interfaces on a traditional flat screen (e.g., desktop monitors, laptops, TV screens, etc.). Augmented reality and virtual reality (AR/VR) simulations were only counted if they were utilized through a headset.

Appendix E

List of Scales Featured for Measuring Emotions

Instrument Name	# of Times Featured
Custom Instrument	11
State-Trait Anxiety Inventory (STAI)	8
Interview: Thematic Analysis	3
Barrett and Russell's Semantic Structure of Emotion Scale	3
Achievement Emotions Questionnaire (AEQ)	3
17-item Training Evaluation Inventory (TEI)	1
NASA task load	1
Generalized Anxiety Disorder Scale (GAD-7)	1
Experiences in Close Relationships–Short Form (ECR-SF)	1
Mental Readiness Form (MRF)	1
Positive and Negative Affect Scale (PANAS)	1
Nested data structure of consultation	1
Swedish Occupational Fatigue Inventory (SOFI) NASA-TLX instrument	1
Speech Content Analysis	1
Interview: Transcript Analysis	1
Profile of Mood States 2	1
Facial Recognition Software: Affectiva	1
Interview: Analysis based on transformative learning framework	1
Self-Report: Qualitative analysis	1

Profile of Mood States (POMS)	1
Watson and Friend's Fear of Negative Evaluation Scale	1
Visual Analog Scale on Stress	1
Fear of Death Scale	1
Visual analog scale on psychological preparation (VASpp)	1
Social Avoidance and Distress Scale (SAD)	1
The Fear of Negative Evaluation Scale (FNE)	1
Test of Self Conscious Affect	1
Experimental Shame Scale (ESS)	1
Heart Rate	1
Blood Pressure	1

References for Appendix

1. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. *Int J Soc Res Methodol*. 2005;8(1):19-32. doi:10.1080/1364557032000119616

2. Creswell JW. *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. 4th ed. SAGE Publications; 2014.

3. Turner TL, Balmer DF, Coverdale JH. Methodologies and study designs relevant to medical education research. *Int Rev Psychiatry*. 2013;25(3):301-310. doi:10.3109/09540261.2013.790310

4. Bilgic E, Gorgy A, Yang A, et al. Exploring the roles of artificial intelligence in surgical education: A scoping review. *Am J Surg*. Published online November 2021:S0002961021006826. doi:10.1016/j.amjsurg.2021.11.023

Chapter 3 Bridging Text

In Chapter 2, I have identified that emotions research in SBE tend to lack deep theoretical integration. Studies that did rely on theoretical models relied on Russel's work on the circumplex model of emotions[44], or Pekrun's CVT[54]. As discussed in Chapter 1, these theories have common grounds in defining what emotions are, including its nature of having multiple components, and requiring attention for emotion generation. Russel and Pekrun's models also share the same categorization system where they both rely on valence and activation (i.e., arousal) axes to describe emotions.

Pekrun's CVT have been used in medical education repeatedly since the early 2010s,[7,112] with recent work focusing on validating the application of CVT for capturing emotions in medical education (including in cross-cultural contexts)[64,72]. While it can be perfectly acceptable for other emotion frameworks to be applied to medical education contexts, the lineage of research the CVT possess can be a strong starting point for researchers entering the realm of emotions research in medical education.

In the following manuscript, I examine emotions in an understudied subject domain: anti-harassment education. I also examine emotions in medical residents, a specialized population. Further, I rely on learning theories including Pekrun's CVT to investigate how our anti-harassment educational intervention was effective, and how emotions relate to our intervention.

Chapter 3 – Anti-Harassment Education for Internal Medicine Residents: Learning Activity Sequence, Productive Failure, and Emotions

Authors: Byunghoon (Tony) Ahn[1], Myriam Johnson[1], Negar Matin[1], Ning-Zi Sun[2,3], Jason M. Harley[1,3,4,5]

[1]*Department of Surgery, McGill University, Montreal, Canada*
[2]*Department of Medicine, McGill University, Montreal, Canada*
[3]*Institute for Health Sciences Education, McGill University, Montreal, Canada*
[4]*Research Institute of the McGill University Health Centre*
[5]*Steinberg Centre for Simulation and Interactive Learning*

Institution: Department of Surgical and Interventional Sciences (Formerly Department of Surgery)

Technology, Knowledge and Learning. Under Review.

Correspondence for Reprints to:
Jason M. Harley, PhD
jason.harley@mcgill.ca

Parts of this work have also been presented as:

Ahn, B.T., Johnson, M., Matin, N., Zhang, X., Han, B., Sun, N., & Harley, J. (2023, August). *Harassment During a Central Line Insertion: A Pilot Study of Bystander Intervention with Videos and Simulation.* Presentation presented at the Association for Medical Education in Europe (AMEE) 2023, Glasgow.

Ahn, B.T., Johnson, M., Matin, N., Sun, N., Derby, N., German, E., Hug, L., Solomon, H., & Harley, J. (2024, April). *Educating Medical Residents on Harassment with Simulations and Videos: Does Order of Educational Activities Matter?* Paper presented at the American Educational Research Association (AERA), Philadelphia.

Ahn, B.T., Johnson, M., Matin, N., Sun, N., Derby, N., German, E., Hug, L., Solomon, H., & Harley, J. (2024, May). *Learning and Simulation Performance in Anti-Harassment Education: Potential Impact of Activity Sequence.* Paper presented at the Richard and

Sylvia Cruess Symposium on Scholarship in Health Sciences Education, Montreal, Canada.

Abstract

Technology rich environments offer potential in delivering effective education to medical trainees, including challenging topics such as anti-harassment education. However, empirical investigations of effective anti-harassment educational interventions are scarce. Our study examined an anti-harassment intervention for medical resident trainees that featured two types of technology rich environments: one featuring animated multimedia videos, and one featuring high-fidelity simulation training. We set out to examine whether the sequential order trainees interacted with these learning activities led to differences in learning and simulation performance. We further examined medical trainees' emotions during these activities, and their potential roles in learning. We drew from the knowledge-learning-instruction framework, the idea of productive-failure, and the control value theory of achievement emotions to conceptualize our research questions and discussion. Our study used a pre-test post-test design with a subsequent follow-up up to 8 months later, where we randomly assigned 52 internal medicine residents to a video-first or a simulation-first group and provided educational activities in different sequences accordingly. We assessed knowledge via a questionnaire, and simulation performance via a performance checklist. Results revealed that our simulation-first group had poorer simulation performance than the video-first group. Both groups showed significant increase in knowledge post-intervention. The simulation-first group had significantly higher knowledge than the video-first group when the intervention was over. Residents reported positive-activating emotions such as enjoyment and curiosity during our intervention. Educators should critically reflect on their use of specific educational technology and learning activity sequences.

Keywords

Medical education, harassment, video instructions, simulation-based education, emotions, activity sequence

Acknowledgements

We would like to thank Elias German and Nicholas Derby with assessing the simulation performance of our participants. We would like to thank Caroline White and her team at MUHC-i-SIM center for providing the simulation equipment, location, and personnel. We would like to thank Xinran (Tippie) Zhang for partaking as the standardized participant, as well as Lena Hug for help facilitating and administering the educational interventions.

Background

Harassment in healthcare education has been a documented issue for more than 30 years. (Neville, 2008; Timm, 2014). Reports from various countries such as Iran (Alimohammadi et al., 2013), Greece (Chrysafi et al., 2017), Canada (Bahji & Altomare, 2019), South Korea (Chang & Jeong, 2021), and the United States (Chadaga et al., 2016) illustrate its global presence. In addition, while such workplace violence continues to plague various disciplines of health sciences—including medicine (Fnais et al., 2014), dentistry (Ullah et al., 2018) and nursing (Clarke et al., 2012)—research suggests a lack of substantial progress in eliminating or even reducing harassment from healthcare education (Phillips, 2016). Considering harassment's potential harm to trainees' well-being such as burnout and suicidal ideation (Hu et al., 2019), as well as withdrawal from programs (Major, 2014), empirical investigations to establish educational interventions that empower trainees to address harassment effectively are sorely needed (Averbuch et al., 2021; Huang et al., 2018). In a response to tackling the issue of harassment, this study extends the tradition of employing technology-rich environments (TRE) in medical education (Duffy et al., 2020; Zheng et al., 2022) by utilizing video-based instructions and high-fidelity simulation training to foster authentic, interactive learning experience.

Prior research has provided support for both video-based education and simulation-based education (SBE) in delivering effective teaching (Lynch, 2020; Valery et al., 2019). Brockfeld and others (2018) demonstrated that video lectures can be as effective as in-person lectures when it comes to educating medical students for their exams—adding to the list of other studies that lend support for video-based education (e.g., Farahani et al., 2020; Schreiber et al., 2010). Video-based education has been documented to offer advantages such as: presenting more information

in a given amount of time (relative to pure verbal communication); clarifying complex, abstract concepts; being compatible with self-paced learning; and being free of requiring traditional classroom spaces (Brockfeld et al., 2018; Hurtubise et al., 2013).

SBE is known for its benefits of providing a non-threatening learning environment that can be designed to match the learner's current skill level, which can be repeated until the learner reaches the desired mastery level (Dutta & Krummel, 2006). These benefits especially shine when simulations can be used to prepare learners for uncommon, yet crucial scenarios learners may encounter in real-world practice (Fanning & Gaba, 2009), such as patient death (DeMaria et al., 2016), or gun violence in the hospital (Brown et al., 2018). All in all, SBE can offer an "emotionally compelling learning experience" (Fraser & McLaughlin, 2019, p.184) that can help learners achieve both technical and non-technical skill acquisition (Blackmore et al., 2018; Tan & Sarker, 2011). While there are various simulation types, ranging from screen-based simulations to virtual-reality headset experiences, high-fidelity mannikins also serve as an essential component in creating TREs that closely mimic clinical settings, allowing for an immersive and interactive learning experience.

We therefore developed an educational intervention featuring these two modalities with the aim of empowering trainees with knowledge to help them recognize, intervene in, and report harassment. Further, we aimed to leverage many of SBE's advantages, including providing a safe environment for trainees (Cook et al., 2018) to practice intervention strategies. This study investigates whether our educational intervention with high-fidelity mannikin simulation and videos were effective for training medical trainees (residents) on bystander harassment intervention. In addition, we examined the potential impact the order of the educational activities

may have had on knowledge and performance outcomes. We used the Knowledge Learning Instruction (KLI) framework (Koedinger et al., 2012) to help us interpret our intervention's effectiveness. Specifically, KLI describes how our learning events' instructional principles target knowledge components to facilitate various learning processes. In addition, we drew upon the concept of "productive failure" (Kapur, 2016; Steenhof et al., 2019), a phenomenon characterized by learners solving problems without sufficient prerequisites (therefore performing poorly), and then receiving instructions and opportunities to consolidate target concepts. Lastly, we refer to Pekrun's Control Value Theory of (CVT) Achievement Emotion to interpret our results in relation to how emotions can impact learning outcomes.

Overall, our study aimed to address the gap in developing and investigating actionable educational responses to the pervasive issue of harassment. This gap is characterized by three factors. First, while research continues to highlight the prevalence of harassment (e.g., Bahji & Altomare, 2019; Fnais et al., 2014; Hu et al., 2019), there is a paucity of empirical studies that critically examine educational curricula designed to combat harassment. For example, while some have called for specific modalities such as SBE to be implemented for combatting harassment (Duchesne et al., 2023; Herweck et al., 2021), exactly how SBE's affordances fare when implemented in a specific learning scenario needs investigation. Second, the empirical investigations of such training tend to omit conceptual frameworks in their work. And, third, these investigations also too often rely on self-report surveys for evaluations (Mazer et al., 2018). In response to this gap, our study presents an innovative simulation and video-based instruction; our simulation is one of the firsts in integrating a harassment intervention component into a medical procedure to reflect authentic clinical environments where harassment may occur. We

further go beyond self-reported satisfaction or perception surveys, but instead focus on measuring knowledge with a questionnaire, and assessing simulation performance via video analysis and checklist. Lastly, we anchor our study and findings to conceptual frameworks to extend the wider education literature on sequences of different educational activities (Chen & Kalyuga, 2020; Lai et al., 2017).

Knowledge Learning Instruction Framework

The KLI framework (Koedinger et al., 2012) uses its proposed taxonomies for knowledge, learning and instruction to match optimal learning processes to specific knowledge components. It further details the instructional principles that can effectively drive these learning processes. In brief, there are three types of knowledge components: 1) Facts, 2) Rules, and 3) Principles. *Facts* are what KLI would describe as "constant-constant knowledge" as one thing is simply associated with another—like a definition corresponding to a word. In our training, knowing that the 4D harassment bystander strategies are "Distract", "Delegate", "Delay", and "Direct" would be an example of a factual knowledge component. *Rules* are described as "variable-constant knowledge", since a single rule (i.e., constant) could apply to wide ranging circumstances (i.e., variable). Knowing to categorize personal attacks as harassment, and thereby distinguishing it from constructive criticism would be an example of a rules knowledge component. Lastly, *principles*, described as "variable-variable knowledge", refers to knowledge related to mental models of how things work. For example, knowing how to choose the appropriate harassment intervention strategy while accounting for patient safety, patient outcome, and professionalism would require a deep understanding of the "principle" of being a proficient doctor.

The KLI framework describes how each of the three types of knowledge best match up with specific learning processes. The framework posits that "memory and fluency building" (e.g., repeatedly testing for memory retention) is most optimal for facts related knowledge. It also states that "induction and refinement" (e.g., providing a timely correction to an erroneous decision) is most optimal for rules related knowledge. It finally states that "understanding and sense making" (e.g., asking the learner to verbally walk through their decisions) is most optimal for principles related knowledge.

Educators and researchers can utilize the KLI framework to dissect their learning objectives and pinpoint the most pertinent knowledge components. This lens offers clarity on which learning processes could be most effective, guiding the selection of the optimal instructional techniques. For example, Stahl et al. (2023) aimed to improve surgical trainees' recognition of surgical anatomy. They identified two major knowledge components crucial for surgical residency: 1) a constant-constant knowledge component (i.e., facts), which entails an immediate comprehension of anatomical visual features akin to effortlessly reading a sentence; and 2) a variable-variable knowledge component (i.e., principles), which encompasses the ability to articulate the perceived anatomical characteristics and link them to relevant concepts and knowledge at the moment. The authors accordingly leveraged two instructional techniques: sense-making and perceptual-fluency intervention (aligned with KLI's "memory and fluency building"). Specifically, their sense-making intervention included a 6-minute video that highlighted visual cues to distinguish the hernia sac from adjacent structures during surgery, while their perceptual-fluency intervention included 30 visual flashcards showcasing images of laparoscopic inguinal hernia repairs.

As demonstrated with Stahl and others' (2023) study, KLI can elucidate what learning experiences specific instructional techniques provide, and what knowledge component is best enhanced through those instructions. While our study does not set out to compare different instructional techniques, we utilize the KLI framework to deconstruct our educational intervention and critically assess the driving factors behind our learners' performance and knowledge outcomes.

Productive Failure

Kapur (2016) illustrates the underlying mechanism of how failure and success can both be either productive, or unproductive. *Productive success* is when the learner immediately achieves their learning objectives and can retain their learning over long term. Productive success can occur when learners are scaffolded slowly with helpful guiding posts withdrawn as they gain expertise. Problem-based learning, and guided inquiry can often lead to productive success. *Unproductive success* is when learning objectives are achieved immediately, but the learned behaviour or knowledge decays quickly. Sole drill-and-practice, and rote memorization can often lead to such unproductive success.

Failures are *unproductive* when the learner fails at a given learning task and achieves none of the learning objectives. Pure discovery learning can often lead to unproductive failures, as learners are expected to uncover concepts solely by themselves, devoid of any guidance or support (Kapur, 2016). In contrast, *productive failures* refer to when learners fail at a given task but achieve their learning objective and the learning is retained long-term. Productive failure involves two phases: a problem-solving phase followed by a consolidation (or instruction) phase. The problem-solving phase affords opportunities for students to generate and explore the

affordances and constraints of multiple solutions to novel, complex problems. The consolidation phase affords opportunities for comparing and contrasting, organizing, and assembling the relevant student-generated solutions into the final, accepted solution. In sum, the aim of productive failure is not immediate problem-solving success, but rather priming students for enhanced learning from subsequent instruction.

Productive failures have been examined in the health professions education space. For example, (Steenhof et al., 2019) investigate whether pharmacy students placed in the productive failure learning group would have better learning outcomes than students placed in the direct instruction only group. The students were to learn how to apply the Cockcroft-Gault equation to determine proper medication doses. While the direct instruction group was taught this equation directly to estimate creatinine clearance, the productive failure group was given the task of estimating the creatinine clearance without the equation—effectively leaving the students to engineer the equation Cockcroft and Gault invented by themselves. All students were then given subsequent instructions where they were given the Cockroft and Gault equation and were given opportunities to practice applying the equation to various problems. The authors reported that the productive failure group had higher performance than their peers when it came to their level of preparedness for future learning (i.e., Preparation for Future Learning; PFL; Mylopoulos et al., 2016).

Yang and others (2019) on the other hand, examined medical students learning about integrated care via simulation training. The students played the role of a geriatric patient or their caregiver, and experienced navigating through the healthcare system by interacting with various healthcare professionals (e.g., nurse, psychologist, pharmacist, etc.). Findings revealed that the

students who “struggled” (i.e., failed) to recall the case information and navigate through the healthcare system, in the end, came to a better understanding of integrated care. It is important to note that subsequent instructions followed post-simulation in the form of reflection writing exercises, group discussions, and facilitated debriefing—which the authors speculate fostered productive failure.

As our study featured a design where some resident learners would be given a task (i.e., simulation) without specific prior training, but with subsequent instructions meant to foster learning, we expected productive learning to be relevant to our results.

Control Value Theory of Achievement Emotions

Emotions are ubiquitous in educational settings (Pekrun, 2019), including in SBE environments (LeBlanc & Posner, 2022). Pekrun’s theory describes achievement emotions as a type of emotions related to the success or failure of an achievement activity (e.g., learning how to intervene in harassment situations). Emotions are elicited based on one’s perceived control over, valuation of, and the outcome of an achievement situation. In other words: how much control do I have over this? Do I care about this? And, did I do well? Emotions are described based on a three-dimensional taxonomy: valence (negative or positive), activation (deactivating or activating, also known as arousal), and object focus (retrospective outcome, concurrent activity, or prospective outcome). These psychological states are important because they can impact learners’ cognitive, motivational, and regulatory processes, and specific types of emotions are associated with high and low learning outcomes (Pekrun, 2019). Specifically, research indicates that positive-activating emotions (e.g., enjoyment) should favor learning

outcomes, while other types of emotions will tend to hinder outcomes, especially negative-deactivating emotions such as hopelessness (Loderer et al., 2020).

An important aspect of the CVT is how the object focus can be anchored to other relevant aspects of the achievement situation. These can include not only the current ongoing learning activity, and the outcome of the activity, but also the contents of the learning material (i.e., topic emotions), and the interpersonal social factors (i.e., social emotions) as well (Pekrun & Stephens, 2012). For example, a student learning about how to intervene in a harassment situation through a simulation may experience the following emotions: 1) enjoyment, as acting out different strategies can be stimulating and engaging; 2) anxiety, as the idea and topic of harassment can reminds one of a prior negative experience. And 3) pride, when the simulation ends, and the learner reflects how they performed better than they initially expected. In sum, during moments of learning, a learner can experience diverse kinds of emotions depending on what aspect of the learning experience the learner is focused on.

While SBE tends to be filled with negative-activating emotions such as anxiety, exploring other aspects of emotions such as object focus is of importance. For example, Madsgaard and others' (2022) review of health profession students' emotional experiences in SBE concluded that a typical SBE is predominantly filled with fear and stress—most studies featured in their review reported the likes of performance anxiety, discomfort, disappointment, and devastation from the learners. Another review that examined medical students and trainees' emotional experiences (Ahn et al., 2023) echoed a similar finding, but suggested that the prevalence of these negative emotions seem to stem from studies focusing on students' reactions to particularly challenging SBE scenarios—like simulated patient death (Fraser et al., 2014) or

managing emergency room acute care as novices (Groot et al., 2020). In other words, if SBE researchers probed different object foci of their leaners (e.g., social, epistemic, topic emotions, etc.), it is possible the SBE literature captured emotions other than just negative-activating emotions from their learners. Our study therefore drew inspiration from the broader educational psychology research has shown the value of discerning learners' object focus to better assess their emotional responses to both the learning subject and the method of instruction (Harley et al., 2019).

In sum, given how emotions are dependent on object focus, our study set out to assess both the topic emotions and the activity emotions (i.e., learning via a simulation, or a video) to examine whether emotions played any significant role in our resident learners' education.

The Current Study

The KLI framework, Kapur's conceptualization of productive failure, and the CVT shaped our hypotheses. First, presenting video instructions before simulation training has traditionally led to high knowledge gains and simulation performance (Kapur, 2016). In contrast, conducting the simulation first may lead to poor performance as the learners must solely rely on prior knowledge. However, productive failure can explain how providing the simulation first can still result in significant improvements in knowledge gains. The simulation exercise serves as the "failure", while the debriefing and the video serve as opportunities for consolidation. Specifically, debriefing requires residents to explain their response to harassment situations (i.e., "prompted self-explanation", "accountable talk"; Koedinger et al., 2012). The videos give worked examples and explanations that highlight the most relevant concepts (i.e., "feature focusing"; Kapur, 2016). Combined, these two educational events provide learning processes for

residents' consolidation. We therefore expected residents to exhibit knowledge gains regardless of the order of educational activities. We further expected residents who participated in the simulation before receiving any instructions to have lower performance score than their peers who were given prior training.

Both the educational videos and the simulation training were developed with best practices and learning theories in mind (e.g., "INACSL Standards of Best Practice," 2016; Mayer, 2019). We therefore expected the actual activity—viewing the educational videos or going through the simulation—to be engaging and therefore elicit activity enjoyment. We expected the topic emotion—learning about harassment—to elicit anxiety or boredom, depending on whether the learners valued the content of our instructions. Based on the CVT's empirically-backed propositions on emotions role in learning (Loderer et al., 2020; Pekrun, 2019), we expected positive-activating emotions such as enjoyment and curiosity to be related to high knowledge levels post-learning.

We proposed the following research questions: RQ1: Did medical residents (i.e., our learners) in the video-first group have a statistically significantly better simulation *performance* when compared to the simulation-first group? RQ2: Did residents' *knowledge* of harassment and related concepts statistically significantly increase from pre to post educational intervention in both groups? In other words, did residents learn from our educational intervention? RQ3: What were the dominant emotions of residents during the A) video instructions, and B) simulation training? RQ4: Were there correlations between the emotional experiences of participants during a learning activity and their subsequent knowledge levels?

Method

Participants

We recruited 60 of the 88 (68.18%) of the first and second-year internal medicine residents from a large North American university from the 2022-2023 academic year (June 2022 – June 2023). From these residents, 52 agreed to participate in our study (Age $M = 27.7$, $SD = 2.65$; 43.5% female). Five residents were unavailable for participation, while three residents declined consent. We employed convenience sampling as we recruited the residents when they were scheduled to receive our training as part of their curriculum. We assured their anti-harassment training regardless of their decision to partake in our research study as it was part of their training that day. An IRB approval was obtained prior to the conduct of the study. Residents could choose which data channel we could collect (i.e., self-report, video recording of simulation).

Educational Intervention

Video Lectures

Our intervention featured 5 PowerPoint screencast animated video lectures we created (approximately 37 minutes collectively). The videos had voice overs to explain the slides and used original illustrations and animations to provide examples of concepts. They covered harassment and related terms, how to identify harassment, and how to provide effective bystander interventions (i.e., 4D strategies by Right To be; Sexual Violence Training Development Team, 2021; *The 5Ds of Bystander Intervention*, n.d.). Our intervention featured four strategies: Distract—shift attention away from victim; Direct—directly confront the harasser; Delay—follow up with the victim post-harassment; and Delegate—call authorities for help. See Fig. 1 for example screenshots.

Simulation Training & Debriefing

Our simulation featured the CAE Blue Phantom Central Line Ultrasound Training Model. Relatedly, our scenario required residents to perform an internal jugular central line insertion on a mannikin while a senior resident (Standardized Participant; SP) was present to supervise and engage in increasingly intensive harassment (e.g., constant interruption, verbal abuse) towards the medical student (SP) who was observing the procedure. The high-fidelity mannikin featured lifelike external landmarks for identify the correct insert site for the central line, in addition to allowing for real punctures with needles, and being compatible with ultrasounds to produce realistic images of vascular access anatomy of the upper thorax and neck. In addition, the simulation participant was provided with the central line insertion kit in addition to other equipment such as simulated anesthetics, gown and gloves, and an ultrasound machine.

One resident participated in the simulation at a time. Simulation debriefing, a facilitated discussions that takes post-simulation to provide structured feedback, was based on the PEARLS framework (Eppich & Cheng, 2015). Both simulation and debriefing were developed per Healthcare Simulation Standards of Best Practice (Watts et al., 2021). Relatedly, we ensured psychological safety of our participants by following these guidelines, as well as consulting various education experts, following ethics guidelines, along with other considerations that are explained in supplementary material A. See Figure 2 for a representative snapshot of our simulation taken from our simulation-walkthrough demonstration video for residents to view after our study was over. See supplementary material B, C, D, and E for details on the simulation scenario progression guideline, debriefing guideline, pre-briefing guideline, and consent forms for participants respectively.

Graduate students and one volunteer undergraduate student, including those who helped develop the script and were working on dissertation or theses from the study, served as standardized participants along with simulation centre staff because of lack of availability of simulation centre staff and/or professional actors due to human and material resource constraints and the ongoing COVID-19 pandemic. Simulation staff and graduate students played different roles at times due to availability, but all received the same training. Trainees and students playing an active role in simulations is not uncommon (e.g., Tyson et al., 2024; Botelho et al., 2024).

Measures

Knowledge

We measured harassment-related knowledge using a questionnaire. We developed 4 versions, each featuring 15 questions with 4 response options. All questions across the 4 versions targeted the same content and had very similar questions and answer choices (i.e., foils; see Fig. 3). We created similar versions to help account for test-retest effect and the ordering effect. We elected not to include more versions, or administer the 4 versions per condition group, to ensure accuracy and feasibility in our study logistics, including survey development and administration.

Simulation Performance

We measured simulation performance via a checklist. It was developed accordingly to the Healthcare Simulation Standards of Best Practice (Watts et al., 2021), and Standards for Educational Psychological Testing (American Educational Research Association, 2011) to support assessment validity. The checklist assessed fulfillment of core elements of different bystander intervention strategies with 3 categories of items. Each category corresponded to 3 of

the 4D strategies relevant to our simulation scenario: direct, distract, and delay (1 strategy—delegate—was irrelevant for our scenario). See Table 1 for details.

The first author and two of the co-authors assessed the simulations via video recordings and our checklist. Based on 12 items, mean Fleiss' κ was at a substantial level (Landis & Koch, 1977) of 0.74 (*SD* = 0.17; Min = 0.50; Max = 0.95). Further, lead evaluator centric interrater reliability was calculated for each checklist item (*M* = 91%; *SD* = 0.8%; Min = 67%; Max = 100%) because the first author had the most expertise in simulation training and was the facilitator for the simulations.

Emotions

We measured emotions using a modified version of the multiple object foci emotion questionnaire (MOFEQ; Harley et al., 2016, 2019, 2020), which in turn is a modified version of the achievement emotions questionnaire (AEQ; Pekrun et al., 2002). MOFEQ allows for measuring a set of emotions for each object focus; in our case our foci were the topic, and the modality used for the educational activity—effectively representing the activity emotion the participant experienced (i.e., video or simulation). We measured 6 emotions for each object focus: enjoyment, curiosity, anxiety, frustration, boredom, and hopelessness. The MOFEQ uses a 5-point Likert scale that ranges from 1 (Strongly Disagree) to 5 (Strongly Agree); the items differed accordingly to the object focus to be assessed. For example, when residents filled out the MOFEQ after watching the instructional videos, an item assessing topic enjoyment would state: "I enjoyed learning about harassment and related trainings". An item assessing activity enjoyment regarding the video instructions would instead state: "I enjoyed learning with the Anti-Harassment Training Videos".

Prior studies that focused on general higher education with MOFEQ have gathered support for its internal and external validity by confirming the expected directions of correlations between sets of emotions, appraisals, and perceived success (Harley et al., 2019). Further, direct consultation from Pekrun (Harley et al., 2020) on the MOFEQ lends further validity evidence for the MOFEQ.

Procedure

Participants were randomized into two groups: 1) video-first group (vid-first; n=29); and 2) simulation-first group (sim-first; n=23). There were four timepoints in our study (t1, t2, t3, and t4). T1 was the pre-intervention phase, and participants from both groups completed a knowledge questionnaire. Participants then engaged in the first educational activity and completed a knowledge questionnaire and the MOFEQ (t2). After, the participants engaged in their second educational activity, and completed another knowledge questionnaire and the MOFEQ (t3). The vid-first group participants watched the videos for their first educational activity, and then engaged in the simulation for their second. This was reversed for the sim-first group. The entire intervention was under 100 minutes. A follow-up survey occurred at the end of the school year (t4), where the participants filled out the knowledge questionnaire again. We counter-balanced our questionnaires: Vid-first participants used questionnaire versions 1, 2, 3, 4 for t1, t2, t3, t4 respectively; the sim-first participants used versions 2, 4, 1, 3 instead. See Fig. 4 for study overview diagram.

Analyses

We ran an independent t-test to answer our first research question. 15 participants from vid-first and 15 participants from sim-first consented to video recording and assessment of their

simulation. We adjusted statistical outlier datapoints to the next most extreme but non-outlying datapoint (i.e., winsorized; Kwak & Kim, 2017). We winsorized 2 extreme datapoints to adjust outliers. Post-hoc power analysis via *G*Power* with large effect size (d = 0.8; Cohen, 1988) revealed a power of 0.56.

We conducted a repeated measures ANCOVA to answer our second research question. Our preliminary analysis identified a significant difference between the sim-first and vid-first group at t1; to account for this initial difference of knowledge levels, we used t1 knowledge level as the covariate for our analysis when examining timepoint and group difference in knowledge between t2 and t3. We conducted one-way ANCOVAs at t2 and t3 to examine whether there were group differences in knowledge when accounting for the covariate. The analysis included 29 participants from the vid-first group, and 22 participants from the sim-first group. We winsorized 5 datapoints to adjust outliers. Post-hoc power analysis with G*Power large effect size (f = 0.40; Cohen, 1988) revealed a power of 1.00 for repeated measures within-between interaction, and a power of 0.80 for the ANCOVAs.

We conducted descriptive statistics and a Wilcoxon signed-rank test to answer our third research question. For the descriptive statistics, our focus was on identifying salient emotions. Participants rated emotions on a 5-point Likert scale: 1 = Strongly Disagree; 2 = Disagree; 3 = Neutral; 4 = Agree; 5 = Strongly Agree. We considered emotions with mean values of 3.5 and above as salient because this would indicate that, on average, participants agreed with experiencing that emotion rather than reporting being "neutral". We used a series of Wilcoxon tests to identify whether there was a significant difference of dominant emotions between the vid-first and sim-first group, and whether dominant emotions were different between educational

activities. Wilcoxon was used due to our data being skewed. These analyses included 28 to 29 participants from the video-first group, and 22 to 24 participants from the sim-first group depending on missing data. Post-hoc power analysis via G*Power for large effect size ($d_z = 0.80$) revealed a power of 0.96 for the Wilcoxon tests.

We ran Kendall's Tau-b test to answer our fourth research question. The tests included 21 to 24 participants from the sim-first group, and 28 to 29 participants from the vid-first group. We relied on G*Power's bivariate normal model statistical test to estimate power; post-hoc power analysis with a large effect size ($r = 0.50$; Cohen, 1988) revealed a power of 0.67 for analysis with 21 participants, while analysis with 29 participants yielded power of 0.81.

Results

Research Question 1

An independent t-test revealed that the vid-first group ($M = 0.82$, $SD = 0.10$) had significantly higher bystander intervention simulation performance scores than the sim-first group ($M = 0.72$, $SD = 0.13$), $t(28) = 2.29$, $p = .030$, $d = 0.84$.

Research Question 2

We conducted a repeated measures ANCOVA to assess whether residents' knowledge of harassment increased from our educational intervention with t1 knowledge included as a covariate. See Table 2 for descriptive statistics. See Table 3 for tests of within-subject factors and Table 4 for between-subjects effects. There was a significant main effect of time on knowledge scores, $F(1, 47) = 18.91$, $p < .001$, partial $\eta^2 = 0.287$, indicating that knowledge scores statistically significantly increased from t2 to t3 for all participants, regardless of group assignment. The main effect of group assignment was not significant, $F(1, 47) = 3.29$, $p = 0.076$,

partial $\eta^2 = 0.065$, suggesting no *overall* difference in knowledge scores (i.e., pooled knowledge scores of t2 and t3) between the sim-first and vid-first groups when considering the covariate.

There was a significant interaction between time and t1 knowledge levels (i.e., *covariate*), $F(1, 47) = 7.08$, $p = 0.011$, partial $\eta^2 = 0.131$, indicating that the increase in knowledge scores over time was different depending on the residents' timepoint 1 knowledge. To investigate group differences separately at t2 and t3, we followed up the above analysis with ANCOVAs, with t1 knowledge level as the covariate. There was no statistically significant difference in t2 knowledge scores between the groups, $F(1, 48) = 2.98$, $p = 0.09$, partial $\eta^2 = 0.059$. However, even after controlling for t1 knowledge, there was a significant difference between the groups at t3, $F(1,49) = 14.86$, $p < 0.001$, partial $\eta^2 = 0.233$. Pairwise comparison showed that the sim-first group had higher knowledge than their vid-first peers, $\Delta M = 0.175$, 95% CI[0.084, 0.266], $p < 0.001$.

Research Question 3

Dominant Emotions During the (a) Simulation Training and (b) Video Instructions

During the (a) *simulation training*, the *vid-first group* generally reported experiencing enjoyment ($M = 4.07$, $SD = 0.65$) and curiosity ($M = 3.89$, $SD = 0.83$). Similar results were seen for topic emotions, where the group reported an enjoyment level of 3.86 ($SD = 0.83$) and curiosity level of 3.55 ($SD = 0.95$). The *sim-first group* also generally reported experiencing enjoyment ($M = 3.65$, $SD = 0.94$) and curiosity ($M = 3.87$, $SD = 0.92$). Similar results were seen for topic emotions, where the group reported enjoyment level of 3.63 ($SD = 0.92$) and curiosity level of 3.54 ($SD = 0.93$).

For the (b) *video instructions*, the *vid-first group* reported experiencing enjoyment (M = 3.55, SD = 0.95) when learning via the instructional videos. Similarly, the group reported a mean enjoyment of 3.52 (SD = 1.02) for the topic emotion enjoyment. The *sim-first group* did not report any prominent levels of emotions regarding the instructional videos. The highest mean level of emotion was 3.33 (SD = 1.09) for *both* the activity emotion and topic emotion.

See Table 5 for descriptive statistics regarding the *vid-first group*'s emotions, and Table 6 for the *sim-first group*'s emotions.

Differences Between Activity and Topic Emotions During the (a) Simulation Training and (b) Video Instructions

During the (a) *simulation training*, the *vid-first group*'s anxiety activity emotion (Mdn = 3, IQR = 2) was statistically significantly higher than their anxiety topic emotion (Mdn = 2, IQR = 2, $W(28) = 55$, $p = 0.005$, rank biserial correlation = 1.0). For the *sim-first group*, there were no significant difference between activity and topic emotions reported during the simulation.

During the (b) *video instruction*, there were no statistically significant difference between the activity and topic emotions for both groups.

Differences Between Simulation Training and Video Training for (a) Activity Emotions and (b) Topic Emotions

For the (a) *activity emotions*, the *vid-first group*'s reported higher enjoyment during the simulation training (Mdn = 4, IQR = 0) relative to during the video instructions (Mdn = 4, IQR = 1, $W(28) = 5.5$, $p = 0.004$, rank biserial correlation = 0.88). Curiosity was also higher during the simulation training (Mdn = 4, IQR = 0.25) compared to during the instructional videos (Mdn = 3, IQR = 1, $W(27) = 8$, $p < 0.001$, rank biserial correlation = 0.90). Boredom was however, lower

during the simulation training (*Mdn* = 2, *IQR* = 1) when compared to the instructional videos (*Mdn* = 3, *IQR* = 1, $W(27) = 203.5$, $p < 0.001$, rank biserial correlation = -0.94).

The *sim-first group*'s (a) *activity emotions* showed similar results, as boredom was lower during the simulation training (*Mdn* = 2, *IQR* = .5) relative to the simulation training (*Mdn* = 3, *IQR* = 1.5, $W(22) = 91$, $p = 0.001$, rank biserial correlation = 1). Curiosity was also higher during the simulation (*Mdn* = 4, *IQR* = 0) than the videos (*Mdn* = 3, *IQR* = 2, $W(22) = 0$, $p = 0.002$, rank biserial correlation = 1). Lastly, anxiety was higher during the simulation (*Mdn* = 3.5, *IQR* = 2) than during the videos (*Mdn* = 2, *IQR* = 1, $W(21) = 0$, $p < 0.001$, rank biserial correlation = 1).

For the (b) *topic emotions*, both the vid-first and the sim-first group reported significantly lower levels of boredom during the simulation training compared to the instructional videos. The *vid-first group*, reported significantly lower levels of boredom during simulation (*Mdn* = 2, *IQR* = 1) when compared to the videos (*Mdn* = 3, *IQR* = 1, $W(27) = 91$, $p < 0.001$, rank biserial correlation = 1). The *sim-first group* echoed this finding with boredom during the simulation (*Mdn* = 2, *IQR* = 2) also being lower than during the videos (*Mdn* = 3, *IQR* = 2, $W(23) = 87.5$, $p = 0.003$, rank biserial correlation = 0.92).

See Tables S1 to S4 in the supplementary material F to see all Wilcoxon Signed Ranks Test run.

Research Question 4

Kendall's tau-b test did not reveal any correlations between emotions and knowledge levels when applying Bonferroni correction to account for multiple corrections. Without the correction, the sim-first group's enjoyment activity emotion during the video instruction was positively related to the post-video knowledge level, ($\tau b(23) = 0.38$, $p = 0.025$), while the

boredom activity emotion was negatively related to knowledge ($\tau b(23) = -0.37$, $p = 0.031$). Anxiety topic emotion was also negatively related to knowledge ($\tau b(23) = -0.35$, $p = 0.048$). See Tables 7 and 8 to see all correlation tests run.

Discussion

Our study featured four time points that our resident learners went through: t1 (pre-intervention); t2 (post-activity #1), where the vid-first group watched instructional videos and the sim-first group participated in the simulation-training; t3 (post-activity #2), where the learning activities were switched for the two groups; and t4 (follow-up). Keeping in mind our timepoints, sequence of learning activities, t1 knowledge as a covariate in some of our analyses, and the conceptual frameworks we have introduced in mind, we discuss our results below.

Knowledge Gains, Simulation Performance, and Impact of Activity Sequence

Our hypothesis of the vid-first group outperforming the sim-first group in the simulation training was supported. This was expected since the vid-first group's activity sequence—instruction followed by application—is typically recommended to maximize performance (Kapur, 2016). Further, the sim-first residents would have had to solely rely on their limited prior knowledge for their simulation performance.

Our hypothesis of both groups achieving higher levels of knowledge from our intervention was partially supported. It was partially supported because we did not assess differences between t1 and t2 or t1 and t3 as originally planned: only t2 to t3. Our repeated measures ANCOVA revealed that from t2 to t3, our residents showed statistically significant increases in knowledge, despite controlling for the difference in knowledge levels at t1.

The repeated measures analysis by itself suggested that activity sequence had no statistically significant impact on knowledge levels from t2 and t3. However, our follow-up one-way ANCOVAs revealed a statistically significant difference in activity sequence. Considering that t2 knowledge had no significant difference between groups, but t3 knowledge did despite accounting for the covariate, we believe group assignment had as big, or bigger impact relative to the influence the covariate had on later timepoint knowledge. First, the effect size (partial η^2) found for the group effect at t3 was 0.23, relative to the effect size the covariate had at 0.09, indicating the more prominent role group assignment played at t3. Further, this is in contrast to t2, where the covariate's effect size was larger at 0.29, and the insignificant group effect of t2's effect size being 0.06. Overall, we conclude that the sim-first group ended their educational intervention with significantly higher knowledge scores than their vid-first group peers.

The concept of productive failure may explain the potential group difference we observed. The lower simulation performance of the sim-first group is consistent with reports that unguided problem-solving leads to poorer learning (Kapur, 2016). Yet, this seemingly poor performance becomes valuable when combined with subsequent debriefing and instructional videos. We posit that through the simulation, residents recognized gaps in their knowledge, enabling them to focus more effectively on crucial learning elements during subsequent instructions. For example, a sim-first resident, unaware of the "direct" strategy, might find its introduction from the video instructions more enlightening than a resident who started with the video. Realizing how the strategy could have improved their simulation performance, the sim-first resident might experience a deeper understanding and appreciation for the instruction, enhancing their overall learning outcome. Hence, despite initial low gains after the simulation,

the sequencing of activities led to overall more robust increases in knowledge for the sim-first group.

Our educational intervention being successful aligns with KLI's emphasis on providing varied learning experiences, as we did with a blend of video instructions and simulation training (Koedinger et al., 2012). However, our findings point out the framework's limited insight on the potential the sequence of instructional techniques could have on learning. Within our learning context, it seemed promoting procedural skills and targeting more difficult knowledge components (e.g., variable-variable knowledge; Koedinger et al., 2012; knowing how and when to intervene harassment) first, and then following up with simpler conceptual and factual knowledge components and instructions was most effective. We believe KLI would posit that the sim-first approach cultivated a flexible and adaptable foundation of knowledge by fostering inductive reasoning and learning from mistakes. In contrast, the vid-first approach, starting with direct instruction, led to a more rigid integration of knowledge; residents were primed to absorb facts and procedures in a structured manner.

Finally, the KLI framework can help underline how we might have improved the vid-first group's learning, specifically by better matching specific instructional techniques to our teaching objectives. For example, while our videos provided applied examples of utilizing the 4D strategies to intervene, and utilizing the definition of harassment to identify specific behaviours that constitute harassment, we did not provide such examples for making sense of the theory of how harassment occurs in work environments. In hindsight, the lack of such instructional techniques for certain parts of our video instructions may be a source of why the vid-first group had significantly lower knowledge at t3 relative to the sim-first group.

Dominant Emotions During the Intervention

Our hypothesis of observing enjoyment as a dominant emotion from our educational intervention was mostly met. During the simulation activity, both the vid-first and sim-first groups reported enjoyment and curiosity for the activity emotions; topic emotions echoed this finding. During the instructional videos, the vid-first group reported enjoyment regarding the activity and the topic of anti-harassment training. The sim-first group's response, however, was more subdued with their reports being closer to neutral. While both of our educational activities seemed generally enjoyable, our simulation activity seemed to foster better emotional engagement. Specifically, our simulation prompting curiosity is an indication that our intervention fostered trainees' interest in learning more about the topic at hand (Vogl et al., 2021). Further, the general high positive activating emotions may be an indication of successfully providing an immersive TRE experience through the instructional videos and high-fidelity SBE.

Anxiety, while not revealed to be a dominant emotion during the intervention, showed differences in a couple of areas. First, the vid-first group after the simulation reported higher levels of activity emotions anxiety relative to topic emotions anxiety. Second, the sim-first group reported higher levels of anxiety towards the simulation activity relative to the video viewing activity. Our results indicate that our simulation activity has the most potential to elicit anxiety. This is aligned with prior reviews (Ahn et al., 2023; Madsgaard et al., 2022), as well as other research highlighting SBE's affordance of direct hands-on experiential learning that leads to emotionally-charged learning experiences (Rogers et al., 2019).

Our tests for correlations showed limited links between emotions and knowledge levels. The findings were nonetheless aligned with CVT, however. In the sim-first group, enjoyment (a positive-activating emotion) from the video activity correlated positively with knowledge levels. In contrast, boredom (a negative-deactivating emotion) showed a negative relationship with knowledge. Moreover, anxiety (a negative-activating emotion) associated with the video's topic also negatively correlated knowledge levels.

Limitations, Strengths, and Future Directions

This study was limited by a small sample, especially for our follow-up analysis ($N = 30$). Nonetheless with the exception of our t-test for our first research question, we were able to yield acceptable power when set to detect large effect sizes. Future studies with larger samples should enable adequate power for even smaller effect sizes. Our original study design also set out to investigate whether the increase in knowledge levels post-intervention would be retained after a follow-up. However, the attrition rate made our analysis unfeasible, due to only 18 participants of the 51 providing usable data for follow-up analysis. In addition, the group distribution amongst these 18 participants were uneven, with most (72.2%) of them being from the vid-first group, Longitudinal assessment has traditionally been a challenge (Butler & Raley, 2015), and therefore pursuing a larger study with better participant retention rates during follow-up is another future direction. However, to the best of our knowledge, there is scant research empirically testing performance and knowledge through a multi-phased educational intervention to educate medical residents and even less on harassment intervention strategies. We believe our study adds much needed insight in this line of research.

Relatedly, our study only assessed performance once during the intervention. To further expand on our findings regarding productive failure, future studies can strive to include multiple simulations for performance assessments. Our current study's inclusion of performance assessment, however, still answers to the call of going beyond the typical probing of self-perceived success or intervention satisfaction via learner-completed self-report surveys.

Another limitation of our study was t1's knowledge level acting as the covariate for future knowledge levels. While this may be attributable to the difference of prior knowledge in our randomized groups, we believe it is more likely an unintentional difference in the difficulty in the knowledge quiz versions. While we were able to confirm knowledge acquisition and group differences after accounting for t1 knowledge, analysis of the test items via the KLI framework and adjustments of the questionnaire should lead to clearer results in future studies. Specifically, by deconstructing which knowledge component the items are testing, and which instruction methods were present in the learning activities corresponded to that component may help with optimizing even stronger learning outcomes.

Other future directions include examining underlying factors that contribute to productive success or failures such as cognitive and affective factors (Kapur, 2016; Koedinger et al., 2012). Our current analysis of emotions while revealing some relationships with knowledge, did not reveal any clear relationship between emotions and performance. However, our emotions analysis highlights the potential gap in the KLI framework. Specifically, while the framework focuses on matching optimal instructional techniques to specific knowledge component types, it lacks guidance on how educators can design instructional materials that are emotionally engaging. There have been calls for integrating emotions thoughtfully into medical education to

optimize learning (Ahn et al., 2023; Artino et al., 2012; Fraser & McLaughlin, 2019; LeBlanc & Posner, 2022). Future directions therefore include focusing on the roles of emotions in specific instructional techniques the KLI proposes. One specific direction would be to explore multimodal measures of emotions via data channels other than self-reports (e.g., electrodermal activity, facial expression) as shown in the contemporary education literature (Barrett, 2016; Harley, 2016).

Lastly, due to resource constraints, we have relied on simulation center staff and laboratory student members who were part of the research team to participate in the simulation as SPs. Therefore, utilizing professionally trained actors and actresses is a future direction for this research.

Conclusion

In conclusion, our study provides valuable insights into how educators can better approach anti-harassment education for healthcare residents, especially keeping in mind how specific educational technology can be utilized. First, our results provide preliminary evidence that our anti-harassment training with its innovative combination of videos and simulations, was effective at improving residents' knowledge of harassment. We described the likely source of knowledge acquisition by dissecting our learning activities via the KLI framework. We therefore illustrated how educators can systematically identify optimal instructional techniques given the specific knowledge-components sought after.

Second, our results suggest that educational programs seeking to improve residents' knowledge may benefit from carefully situating a performance component (e.g., simulation) prior to in-depth instructions. Our results also support the traditional placement of the

performance component after the instructions yield good performance and knowledge gain. Our study echoes findings similarly found in more general fields of education in regards to productive failure (e.g., Lai et al., 2017; Loibl & Leuders, 2019), and highlight the importance in careful consideration of learning activity sequencing.

Lastly, our study adds support to the ubiquitous nature of emotions in medical education, especially when simulations are involved. Given how our simulation training played a crucial role in ensuring our residents' success in our intervention, various emotions associated with anti-harassment SBE such as enjoyment, curiosity, and anxiety warrant further investigations. This study points to future studies that focuses on follow-ups with better participant retention, along with studies that investigate instructional techniques that better prevents decay of harassment intervention knowledge.

We advise educators to try deconstructing their current instructional materials accordingly to educational frameworks such as the KLI and CVT and test whether the intended instructional activities, and the order that they are placed are indeed delivering the intended learning outcomes and psychological engagement.

Reference

Ahn, B., Maurice-Ventouris, M., Bilgic, E., Yang, A., Lau, C. H.-H., Peters, H., Li, K., Chang-Ou, D., & Harley, J. M. (2023). A scoping review of emotions and related constructs in simulation-based education research articles. *Advances in Simulation*, *8*(1), 22. https://doi.org/10.1186/s41077-023-00258-z

Alimohammadi, H., Zolfaghari Sadrabad, A., Bidarizerehpoosh, F., Derakhshanfar, H., Shahrami, A., & Farahmand Rad, R. (2013). Residents' experiences in dealing with abuse in emergency department: A survey in Iran hospitals. *Emergency Medicine Journal : EMJ*, *30*(9), 732–734. https://doi.org/10.1136/emermed-2012-201715

American Educational Research Association (Ed.). (2011). *Report and recommendations for the reauthorization of the institute of education sciences*. American Educational Research Association.

Artino, A. R., Holmboe, E. S., & Durning, S. J. (2012). Can achievement emotions be used to better understand motivation, learning, and performance in medical education? *Medical Teacher*, *34*(3), 240–244. https://doi.org/10.3109/0142159X.2012.643265

Averbuch, T., Eliya, Y., & Van Spall, H. G. C. (2021). Systematic review of academic bullying in medical settings: Dynamics and consequences. *BMJ Open*, *11*(7), e043256. https://doi.org/10.1136/bmjopen-2020-043256

Bahji, A., & Altomare, J. (2019). Prevalence of intimidation, harassment, and discrimination in resident physicians: A systematic review and meta-analysis. *Canadian Medical Education Journal*. https://doi.org/10.36834/cmej.57019

Barrett, L. F. (2016). Navigating the Science of Emotion. In *Emotion Measurement* (pp. 31–63). Elsevier. https://doi.org/10.1016/B978-0-08-100508-8.00002-3

Blackmore, A., Kasfiki, E. V., & Purva, M. (2018). Simulation-based education to improve communication skills: A systematic review and identification of current best practice. *BMJ Simulation and Technology Enhanced Learning*, *4*(4), 159–164. https://doi.org/10.1136/bmjstel-2017-000220

Botelho, F., Gerk, A., Harley, J. M., & Poenaru, D. (2024). Improving Pediatric Trauma Education by Teaching Non-technical Skills: A Randomized Controlled Trial. *Journal of Pediatric Surgery*, *59*(5), 874–888. https://doi.org/10.1016/j.jpedsurg.2024.01.018

Brockfeld, T., Müller, B., & De Laffolie, J. (2018). Video versus live lecture courses: A comparative evaluation of lecture types and results. *Medical Education Online*, *23*(1), 1555434. https://doi.org/10.1080/10872981.2018.1555434

Brown, R. G., Anderson, S., Brunt, B., Enos, T., Blough, K., & Kropp, D. (2018). Workplace Violence Training Using Simulation. *The American Journal of Nursing*, *118*(10), 56–68. https://doi.org/10.1097/01.NAJ.0000546382.12045.54

Butler, A. C., & Raley, N. D. (2015). The Future of Medical Education: Assessing the Impact of Interventions on Long-Term Retention and Clinical Care. *Journal of Graduate Medical Education*, *7*(3), 483–485. https://doi.org/10.4300/JGME-D-15-00236.1

Chadaga, A. R., Villines, D., & Krikorian, A. (2016). Bullying in the American Graduate Medical Education System: A National Cross-Sectional Survey. *PLOS ONE*, *11*(3), e0150246. https://doi.org/10.1371/journal.pone.0150246

Chang, H. E., & Jeong, S. (2021). Male Nurses' Experiences of Workplace Gender Discrimination and Sexual Harassment in South Korea: A Qualitative Study. *Asian Nursing Research*, *15*(5), 303–309. https://doi.org/10.1016/j.anr.2021.09.002

Chen, O., & Kalyuga, S. (2020). Exploring factors influencing the effectiveness of explicit instruction first and problem-solving first approaches. *European Journal of Psychology of Education*, *35*(3), 607–624. https://doi.org/10.1007/s10212-019-00445-5

Chrysafi, P., Simou, E., Makris, M., Malietzis, G., & Makris, G. C. (2017). Bullying and Sexual Discrimination in the Greek Health Care System. *Journal of Surgical Education*, *74*(4), 690–697. https://doi.org/10.1016/j.jsurg.2016.12.005

Clarke, C. M., Kane, D. J., Rajacich, D. L., & Lafreniere, K. D. (2012). Bullying in Undergraduate Clinical Nursing Education. *Journal of Nursing Education*, *51*(5), 269–276. https://doi.org/10.3928/01484834-20120409-01

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed). L. Erlbaum Associates.

Cook, D. A., Andersen, D. K., Combes, J. R., Feldman, D. L., & Sachdeva, A. K. (2018). The value proposition of simulation-based education. *Surgery*, *163*(4), 944–949. https://doi.org/10.1016/j.surg.2017.11.008

DeMaria, S., Silverman, E. R., Lapidus, K. A. B., Williams, C. H., Spivack, J., Levine, A., & Goldberg, A. (2016). The impact of simulated patient death on medical students' stress response and learning of ACLS. *Medical Teacher*, *38*(7), 730–737. https://doi.org/10.3109/0142159X.2016.1150986

Duchesne, E., Caners, K., Rang, L., & Dagnone, D. (2023). Addressing microaggressions with simulation: A novel educational intervention. *Canadian Journal of Emergency Medicine*, *25*(4), 299–302. https://doi.org/10.1007/s43678-023-00474-6

Duffy, M. C., Lajoie, S. P., Pekrun, R., & Lachapelle, K. (2020). Emotions in medical education: Examining the validity of the Medical Emotion Scale (MES) across authentic medical learning environments. *Learning and Instruction*, *70*, 101150. https://doi.org/10.1016/j.learninstruc.2018.07.001

Dutta, S., & Krummel, T. M. (2006). Simulation: A New Frontier in Surgical Education. *Advances in Surgery*, *40*, 249–263. https://doi.org/10.1016/j.yasu.2006.06.004

Eppich, W., & Cheng, A. (2015). Promoting Excellence and Reflective Learning in Simulation (PEARLS): Development and Rationale for a Blended Approach to Health Care Simulation Debriefing. *Simulation in Healthcare: The Journal of the Society for Simulation in Healthcare*, *10*(2), 106–115. https://doi.org/10.1097/SIH.0000000000000072

Fanning, R., & Gaba, D. (2009). Simulation-Based Learning as an Educational Tool. In *Anesthesia Informatics* (pp. 459–479). Springer New York. https://doi.org/10.1007/978-0-387-76418-4_24

Farahani, S., Farahani, I., Burckhardt, B. B., Schwender, H., & Laeer, S. (2020). Self-Instruction Video Versus Face-to-Face Instruction of Pharmacy Students' Skills in Blood Pressure Measurement. *Pharmacy*, *8*(4), 217. https://doi.org/10.3390/pharmacy8040217

Fnais, N., Soobiah, C., Chen, M. H., Lillie, E., Perrier, L., Tashkhandi, M., Straus, S. E., Mamdani, M., Al-Omran, M., & Tricco, A. C. (2014). Harassment and Discrimination in

Medical Training: A Systematic Review and Meta-Analysis. *Academic Medicine*, *89*(5), 817–827. https://doi.org/10.1097/ACM.0000000000000200

Fraser, K., Huffman, J., Ma, I., Sobczak, M., McIlwrick, J., Wright, B., & McLaughlin, K. (2014). The Emotional and Cognitive Impact of Unexpected Simulated Patient Death. *Chest*, *145*(5), 958–963. https://doi.org/10.1378/chest.13-0987

Fraser, K., & McLaughlin, K. (2019). Temporal pattern of emotions and cognitive load during simulation training and debriefing. *Medical Teacher*, *41*(2), 184–189. https://doi.org/10.1080/0142159X.2018.1459531

Gianakos, A. L., Freischlag, J. A., Mercurio, A. M., Haring, R. S., LaPorte, D. M., Mulcahey, M. K., Cannada, L. K., & Kennedy, J. G. (2022). Bullying, Discrimination, Harassment, Sexual Harassment, and the Fear of Retaliation During Surgical Residency Training: A Systematic Review. *World Journal of Surgery*, *46*(7), 1587–1599. https://doi.org/10.1007/s00268-021-06432-6

Groot, F., Jonker, G., Rinia, M., ten Cate, O., & Hoff, R. G. (2020). Simulation at the Frontier of the Zone of Proximal Development: A Test in Acute Care for Inexperienced Learners. *Academic Medicine*, *95*(7), 1098–1105. https://doi.org/10.1097/ACM.0000000000003265

Harley, J. M. (2016). Measuring Emotions. In *Emotions, Technology, Design, and Learning* (pp. 89–114). Elsevier. https://doi.org/10.1016/B978-0-12-801856-9.00005-0

Harley, J. M., Lajoie, S. P., Tressel, T., & Jarrell, A. (2020). Fostering positive emotions and history knowledge with location-based augmented reality and tour-guide prompts. *Learning and Instruction*, *70*, 101163. https://doi.org/10.1016/j.learninstruc.2018.09.001

Harley, J. M., Liu, Y., Ahn, T. B., Lajoie, S. P., Grace, A. P., Haldane, C., Whittaker, A., & McLaughlin, B. (2019). I've got this: Fostering topic and technology-related emotional engagement and queer history knowledge with a mobile app. *Contemporary Educational Psychology*, *59*, 101790. https://doi.org/10.1016/j.cedpsych.2019.101790

Harley, J. M., Poitras, E. G., Jarrell, A., Duffy, M. C., & Lajoie, S. P. (2016). Comparing virtual and location-based augmented reality mobile learning: Emotions and learning outcomes. *Educational Technology Research and Development*, *64*(3), 359–388. https://doi.org/10.1007/s11423-015-9420-7

Herweck, A. M., Kumnick, A., Perone, H., Muller, C., & Bornstein, K. (2021). "What a Pretty Assistant You Have": Addressing Sex-Based Harassment by Patient Offenders in Undergraduate Medical Training. *Medical Science Educator*, *31*(3), 1187–1191. https://doi.org/10.1007/s40670-021-01289-9

Hu, Y.-Y., Ellis, R. J., Hewitt, D. B., Yang, A. D., Cheung, E. O., Moskowitz, J. T., Potts, J. R., Buyske, J., Hoyt, D. B., Nasca, T. J., & Bilimoria, K. Y. (2019). Discrimination, Abuse, Harassment, and Burnout in Surgical Residency Training. *New England Journal of Medicine*, *381*(18), 1741–1752. https://doi.org/10.1056/NEJMsa1903759

Huang, Y., Chua, T. C., Saw, R. P. M., & Young, C. J. (2018). Discrimination, Bullying and Harassment in Surgery: A Systematic Review and Meta-analysis. *World Journal of Surgery*, *42*(12), 3867–3873. https://doi.org/10.1007/s00268-018-4716-5

Hurtubise, L., Martin, B., Gilliland, A., & Mahan, J. (2013). To Play or Not To Play: Leveraging Video in Medical Education. *Journal of Graduate Medical Education*, *5*(1), 13–18. https://doi.org/10.4300/JGME-05-01-32

INACSL Standards of Best Practice: SimulationSM Simulation Design. (2016). *Clinical Simulation in Nursing*, *12*, S5–S12. https://doi.org/10.1016/j.ecns.2016.09.005

Kapur, M. (2016). Examining Productive Failure, Productive Success, Unproductive Failure, and Unproductive Success in Learning. *Educational Psychologist*, *51*(2), 289–299. https://doi.org/10.1080/00461520.2016.1155457

Koedinger, K. R., Corbett, A. T., & Perfetti, C. (2012). The Knowledge-Learning-Instruction Framework: Bridging the Science-Practice Chasm to Enhance Robust Student Learning. *Cognitive Science*, *36*(5), 757–798. https://doi.org/10.1111/j.1551-6709.2012.01245.x

Kwak, S. K., & Kim, J. H. (2017). Statistical data preparation: Management of missing values and outliers. *Korean Journal of Anesthesiology*, *70*(4), 407. https://doi.org/10.4097/kjae.2017.70.4.407

Lai, P. K., Portolese, A., & Jacobson, M. J. (2017). Does sequence matter? Productive failure and designing online authentic learning for process engineering: Productive failure and designing online authentic learning. *British Journal of Educational Technology*, *48*(6), 1217–1227. https://doi.org/10.1111/bjet.12492

Landis, J. R., & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics*, *33*(1), 159. https://doi.org/10.2307/2529310

LeBlanc, V. R., & Posner, G. D. (2022). Emotions in simulation-based education: Friends or foes of learning? *Advances in Simulation*, *7*(1), 3. https://doi.org/10.1186/s41077-021-00198-6

Loderer, K., Pekrun, R., & Lester, J. C. (2020). Beyond cold technology: A systematic review and meta-analysis on emotions in technology-based learning environments. *Learning and Instruction*, *70*, 101162. https://doi.org/10.1016/j.learninstruc.2018.08.002

Loibl, K., & Leuders, T. (2019). How to make failure productive: Fostering learning from errors through elaboration prompts. *Learning and Instruction*, *62*, 1–10. https://doi.org/10.1016/j.learninstruc.2019.03.002

Lynch, A. (2020). Simulation-based acquisition of non-technical skills to improve patient safety. *Seminars in Pediatric Surgery*, *29*(2), 150906. https://doi.org/10.1016/j.sempedsurg.2020.150906

Madsgaard, A., Smith-Strøm, H., Hunskår, I., & Røykenes, K. (2022). A rollercoaster of emotions: An integrative review of emotions and its impact on health professional students' learning in simulation-based education. *Nursing Open*, *9*(1), 108–121. https://doi.org/10.1002/nop2.1100

Major, A. (2014). To Bully and Be Bullied: Harassment and Mistreatment in Medical Education. *AMA Journal of Ethics*, *16*(3), 155–160. https://doi.org/10.1001/virtualmentor.2014.16.3.fred1-1403

Mayer, R. E. (2019). How Multimedia Can Improve Learning and Instruction. In J. Dunlosky & K. A. Rawson (Eds.), *The Cambridge Handbook of Cognition and Education* (1st ed., pp. 460–479). Cambridge University Press. https://doi.org/10.1017/9781108235631.019

Mazer, L. M., Bereknyei Merrell, S., Hasty, B. N., Stave, C., & Lau, J. N. (2018). Assessment of Programs Aimed to Decrease or Prevent Mistreatment of Medical Trainees. *JAMA Network Open*, *1*(3), e180870. https://doi.org/10.1001/jamanetworkopen.2018.0870

Mylopoulos, M., Brydges, R., Woods, N. N., Manzone, J., & Schwartz, D. L. (2016). Preparation for future learning: A missing competency in health professions education? *Medical Education*, *50*(1), 115–123. https://doi.org/10.1111/medu.12893

Neville, A. J. (2008). In the age of professionalism, student harassment is alive and well. *Medical Education*, *42*(5), 447–448. https://doi.org/10.1111/j.1365-2923.2008.03033.x

Pekrun, R. (2019). Achievement emotions: A control-value theory perspective. In *Emotions in late modernity* (pp. 142–157). Routledge.

Pekrun, R., Goetz, T., Titz, W., & Perry, R. P. (2002). Academic Emotions in Students' Self-Regulated Learning and Achievement: A Program of Qualitative and Quantitative Research. *Educational Psychologist*, *37*(2), 91–105. https://doi.org/10.1207/S15326985EP3702_4

Pekrun, R., & Stephens, E. J. (2012). Academic emotions. In K. R. Harris, S. Graham, T. Urdan, S. Graham, J. M. Royer, & M. Zeidner (Eds.), *APA educational psychology handbook, Vol 2: Individual differences and cultural and contextual factors.* (pp. 3–31). American Psychological Association. https://doi.org/10.1037/13274-001

Phillips, J. P. (2016). Workplace Violence against Health Care Workers in the United States. *New England Journal of Medicine*, *374*(17), 1661–1669. https://doi.org/10.1056/NEJMra1501998

Rogers, T., Andler, C., O'Brien, B., & van Schaik, S. (2019). Self-Reported Emotions in Simulation-Based Learning: Active Participants vs. Observers. *Simulation in Healthcare: The Journal of the Society for Simulation in Healthcare*, *14*(3), 140–145. https://doi.org/10.1097/SIH.0000000000000354

Schreiber, B. E., Fukuta, J., & Gordon, F. (2010). Live lecture versus video podcast in undergraduate medical education: A randomised controlled trial. *BMC Medical Education*, *10*(1), 68. https://doi.org/10.1186/1472-6920-10-68

Sexual Violence Training Development Team. (2021). *Active bystander intervention: Training and facilitation guide : training for preventing and responding to sexual violence in B.C. post-secondary institutions*. BCcampus, BC Open Textbook Project.

Stahl, C. C., Rau, M. A., & Greenberg, J. A. (2023). Teaching advanced surgical anatomy with visual representations: Comparing perceptual fluency and sense making. *Instructional Science*, *51*(4), 661–686. https://doi.org/10.1007/s11251-023-09630-y

Steenhof, N., Woods, N. N., Van Gerven, P. W. M., & Mylopoulos, M. (2019). Productive failure as an instructional approach to promote future learning. *Advances in Health Sciences Education*, *24*(4), 739–749. https://doi.org/10.1007/s10459-019-09895-4

Tan, S. S. Y., & Sarker, S. K. (2011). Simulation in surgery: A review. *Scottish Medical Journal*, *56*(2), 104–109. https://doi.org/10.1258/smj.2011.011098

The 5Ds of Bystander Intervention. (n.d.). Right to Be. https://righttobe.org/guides/bystander-intervention-training/

Timm, A. (2014). "It would not be tolerated in any other profession except medicine": Survey reporting on undergraduates' exposure to bullying and harassment in their first placement year. *BMJ Open*, *4*(7), e005140–e005140. https://doi.org/10.1136/bmjopen-2014-005140

Tyson, L., Skinner, J., Hariharan, B., Josiah, B., Okongwu, K., & Semlyen, J. (2024). Tackling discrimination in medicine head on: The impact of bystander intervention training. *Medical Teacher*, 1–10. https://doi.org/10.1080/0142159X.2024.2316849

Ullah, R., Siddiqui, F., Zafar, M. S., & Iqbal, K. (2018). Bullying experiences of dental interns working at four dental institutions of a developing country: A cross-sectional study. *Work*, *61*(1), 91–100. https://doi.org/10.3233/WOR-182784

Valery, J., Helmi, H., Spaulding, A., Che, X., Prada, G., Chamorro Pareja, N., Moreno-Franco, P., & Stancampiano, F. F. (2019). Video intervention to improve incident reporting among medical trainees. *BMJ Open Quality*, *8*(4), e000706. https://doi.org/10.1136/bmjoq-2019-000706

Vogl, E., Pekrun, R., & Loderer, K. (2021). Epistemic Emotions and Metacognitive Feelings. In D. Moraitou & P. Metallidou (Eds.), *Trends and Prospects in Metacognition Research across the Life Span* (pp. 41–58). Springer International Publishing. https://doi.org/10.1007/978-3-030-51673-4_3

Watts, P. I., Rossler, K., Bowler, F., Miller, C., Charnetski, M., Decker, S., Molloy, M. A., Persico, L., McMahon, E., McDermott, D., & Hallmark, B. (2021). Onward and Upward: Introducing the Healthcare Simulation Standards of Best PracticeTM. *Clinical Simulation in Nursing*, *58*, 1–4. https://doi.org/10.1016/j.ecns.2021.08.006

Yang, S., Chaudhary, Z., Mylopoulos, M., Hashmi, R., Kwok, Y., Colman, S., Yogaparan, T., & Sockalingam, S. (2019). Using simulation to explore medical students' understanding of integrated care within geriatrics. *BMC Medical Education*, *19*(1), 322. https://doi.org/10.1186/s12909-019-1758-9

Zheng, J., Li, S., & Lajoie, S. P. (2022). Diagnosing virtual patients in a technology-rich learning environment: A sequential Mining of Students' efficiency and behavioral patterns. *Education and Information Technologies*, *27*(3), 4259–4275. https://doi.org/10.1007/s10639-021-10772-0

Statements and Declarations

Funding

This study was funded by a Social Sciences and Humanities Research Council of Canada (SSHRC) doctoral fellowship (Application number: 752-2022-1899) awarded to the first author, B.T. Ahn, and a SSHRC grant awarded to the senior author, J.M. Harley, Grant ID: 430-2020-00573. The authors have no other fundings to declare.

Competing Interests

The authors have no relevant financial or non-financial interests to disclose.

Author Contributions

BTA: Study design, lead instructional development, instructional delivery (simulation training facilitation), data collection, data preparation, data analysis, lead manuscript writing.

MJ: Instructional delivery (simulation training standardized participation), data collection, data preparation, support manuscript writing.

NM: Instructional delivery (simulation training standardized participation), data collection, support manuscript writing.

NS: Study design, curriculum integration, instructional development, data collection planning

JH: Funding acquisition, supervision, study design, instructional development, data collection planning, data analysis, support manuscript writing.

Compliance with Ethics Standards

The authors have no potential conflicts of interest to report. This study has been reviewed and approved by the Research Ethics and Compliance in the Faculty of Medicine and Health

Sciences at McGill University (IRB: A07-B71-22B). All participants included in our study have provided informed consent for their data collection and analyses.

Figure. 1 Screenshots from our educational videos. These screenshots are from module 4, which covers the bystander effect and the 4D strategies.

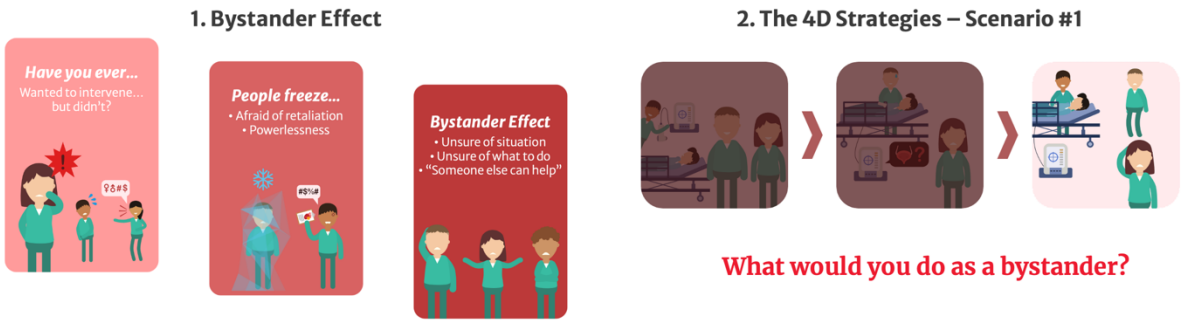


Figure. 2 A Snapshot from our Simulation Demonstration Video. This photo is representative of our simulation scenario. A medical student (SP) is on the left observing, while one of our participant (center) would perform the central line. The senior resident (SP) is on the right constantly harassing the medical student.

Figure 3. Example Items from the Knowledge Questionnaires. Each question targets the same learning content across the four versions of the questionnaire. Each question and answer choice is intentionally similar to ensure consistency

VERSION 1:

10. Greg has made an offensive joke based on Lisa's gender which makes her visibly upset. James keeps quiet until Greg leaves, but offers Lisa to walk with her to the reporting office and talk about what just happened. Which of the following bystander strategies **best** represent James?

- Delegate
- Direct
- Distract
- None of the above

VERSION 2:

10. Greg has made an offensive joke based on Lisa's gender which makes her visibly upset. As a bystander Eric confronts Greg and asks him to explain the joke and questions what the joke implies. Which of the 4D bystander intervention strategy would this be the most related to? Which of the following bystander strategies **best** represent Eric?

- Direct
- Delay
- Distract
- None of the above

Figure 4. Study Flowchart. Both groups filled out knowledge questionnaires, with learning activities in between.

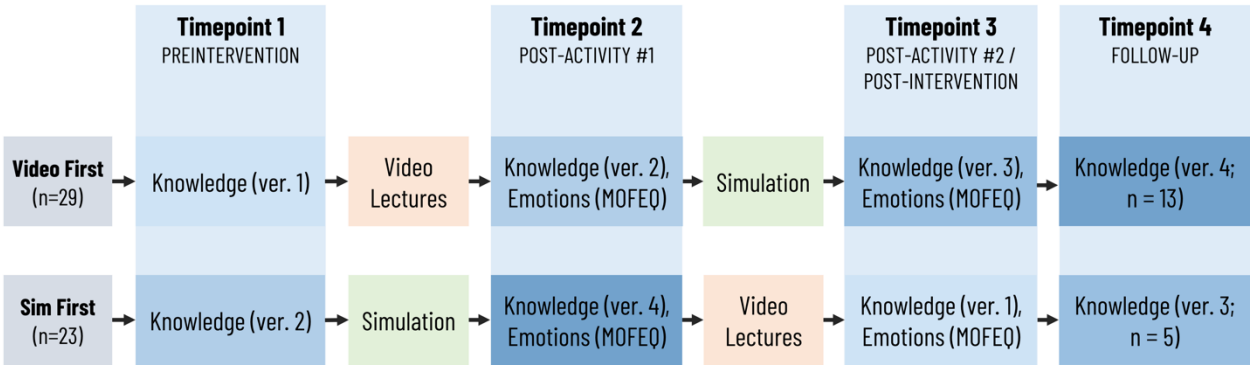


Table 1. Assessment Criteria for Bystander Intervention.

Strategy	Checklist Item	Description	Points
Direct	Confront harasser	Question or challenge the harasser's behaviours or actions as inappropriate or unacceptable.	2
	De-escalation	Use tone and language that does not intend to increase the tension or stakes.	1
	Workplace respect	Emphasize the importance of professionalism, and equity in the workplace and the learning environment.	1
	Resilience	Continues intervention even if initial attempt is unsuccessful.	1
Distract	Shift attention away	Directly engage with the victim or harasser during the harassment (or immediately after). Draw attention away from the harassment.	2
	De-escalation	Use tone and language that does not intend to increase the tension or stakes.	1
	Uplift victim	Raise the victim's self-esteem, specifically aspects that have been the target of harassment.	1
	Resilience	Continues intervention even if initial attempt is unsuccessful.	1
Delay	Check in with victim	Ask the victim how they feel following the harassment.	2
	Offer third-party resources	Refer the victim to organizational offices that can help with reporting harassment or supporting victims.	1
	Problematize the harassment	Ensure that the victim knows that the harasser's actions were unacceptable and inappropriate.	1
	Uplift victim	Raise the victim's self-esteem, specifically aspects that have been the target of harassment.	1
Comprehen siveness	No strategies	No strategies were used.	0
	1 strategy	Just 1 strategy was used.	3
	2 strategies	Just 2 strategies were used.	4
	3 strategies	All 3 strategies were used.	5

Average subtotals of each criterion for final score.

Table 2. Descriptive Statistics for the repeated measures ANCOVA.

	Group	t1 knowledge (covariate)	t2 knowledge	t3 knowledge
N	Vid-first	29	29	29
	Sim-first	22	22	22
Mean	Vid-first	.556	.643	.664
	Sim-first	.397	.619	.812
SD	Vid-first	.161	.211	.205
	Sim-first	.130	.166	.136
Minimum	Vid-first	.200	.130	.130
	Sim-first	.200	.270	.470
Maximum	Vid-first	.730	.930	.930
	Sim-first	.730	.870	1.00

Table 3. Tests of Within-Subject Factors for the Repeated Measures ANCOVA.

Source	Sum of Squares	df	Mean Square	F	Sig.	partial η^2
Timepoint	.201	1	.201	18.911	<.001	.287
Timepoint * Group	.003	1	.003	.28	.599	.006
Timepoint * Timepoint 1 Knowledge	.075	1	.075	7.075	.011	.131
Timepoint * Group * Timepoint 1 Knowledge	$5.77*10^{-5}$	1	$5.77*10^{-5}$	.005	.942	0
Error	.499	47	.011			

Table 4. Tests of Between-Subject Factors for the Repeated Measures ANCOVA.

Source	Sum of Squares	df	Mean Square	F	Sig.	partial η^2
Intercept	1.721	1	1.721	46.535	<.001	.498
Group	.122	1	.122	3.287	.076	.065
Timepoint 1 Knowledge	.379	1	.379	10.236	.002	.179
Group * Timepoint 1 Knowledge	.027	1	.027	.74	.394	.016
Error	1.738	47	.037			

Table 5. Descriptive Statistics for Vid-First Group's Emotions

	Simulation Activity Emotions					
	Enjoyment	Boredom	Frustration	Curiosity	Anxiety	Helpless
N	29	29	29	28	29	29
Mean	4.07	1.97	2.38	3.89	2.79	1.9
Mdn	4	2	2	4	3	2
SD	0.651	0.823	1.05	0.832	1.29	0.976
	Simulation Topic Emotions					
	Enjoyment	Boredom	Frustration	Curiosity	Anxiety	Helpless
N	29	28	29	29	29	29
Mean	3.86	2.14	2.21	3.55	2.31	1.97
Mdn	4	2	2	4	2	2
SD	0.833	0.932	0.902	0.948	1.11	0.981
	Video Activity Emotions					
	Enjoyment	Boredom	Frustration	Curiosity	Anxiety	Helpless
N	29	28	29	29	29	29
Mean	3.55	3.04	2.17	3.28	2.24	1.86
Mdn	4	3	2	3	2	2
SD	0.948	0.922	0.759	0.922	1.02	0.875
	Video Topic Emotions					
	Enjoyment	Boredom	Frustration	Curiosity	Anxiety	Helpless
N	29	28	29	29	29	29
Mean	3.52	2.71	2.07	3.28	2.17	1.86
Mdn	4	3	2	3	2	2
SD	1.02	1.08	0.799	0.996	1.1	0.875

Table 6. Descriptive Statistics for Sim-First Group's Emotions

	Simulation Activity Emotions					
	Enjoyment	Boredom	Frustration	Curiosity	Anxiety	Helpless
N	23	23	23	23	22	23
Mean	3.65	2.09	2.39	3.87	3.09	1.74
Mdn	4	2	2	4	3.5	2
SD	0.935	0.996	1.31	0.92	1.06	0.619
	Simulation Topic Emotions					
	Enjoyment	Boredom	Frustration	Curiosity	Anxiety	Helpless
N	24	24	24	24	24	24
Mean	3.63	2.25	2.33	3.54	2.29	1.88
Mdn	4	2	2	4	2	2
SD	0.924	1.15	1.05	0.932	1.04	0.741
	Video Activity Emotions					
	Enjoyment	Boredom	Frustration	Curiosity	Anxiety	Helpless
N	24	24	24	24	24	24
Mean	3.33	3.17	2.08	3.08	1.71	1.71
Mdn	3	3	2	3	2	2
SD	0.948	0.922	0.759	0.922	1.02	0.875
	Video Topic Emotions					
	Enjoyment	Boredom	Frustration	Curiosity	Anxiety	Helpless
N	24	24	24	24	24	24
Mean	3.33	3.04	2.33	3.17	1.88	1.67
Mdn	3	3	2	3	2	2
SD	1.09	1.08	1.01	1.09	0.797	0.702

Table 7. Correlation Statistics Between Emotions from Simulation and Knowledge

Simulation Activity Emotions		Sim-First Group Knowledge (Post-Sim; t2)	Vid-First Group Knowledge (Post-Sim; t3)
Enjoyment	Kendall's Tau B	-0.137	0.148
	p-value	0.434	0.348
	N	22	29
Boredom	Kendall's Tau B	0.011	-0.145
	p-value	0.951	0.351
	N	22	29
Frustration	Kendall's Tau B	0.094	0.009
	p-value	0.589	0.953
	N	22	29
Curiosity	Kendall's Tau B	0.101	0.045
	p-value	0.57	0.774
	N	22	28
Anxiety	Kendall's Tau B	0.237	0.082
	p-value	0.189	0.583
	N	21	29
Helplessness	Kendall's Tau B	0.159	-0.003
	p-value	0.384	0.984
	N	22	29
Simulation Topic Emotions		Sim-First Group Knowledge (Post-Sim; t2)	Vid-First Group Knowledge (Post-Sim; t3)
Enjoyment	Kendall's Tau B	-0.058	0.041
	p-value	0.736	0.792
	N	23	29
Boredom	Kendall's Tau B	0.028	0.1
	p-value	0.867	0.52
	N	23	28
Frustration	Kendall's Tau B	0.178	-0.012
	p-value	0.296	0.936
	N	23	29
Curiosity	Kendall's Tau B	-0.039	0.104
	p-value	0.819	0.492
	N	23	29
Anxiety	Kendall's Tau B	0.115	0.063
	p-value	0.501	0.677
	N	23	29
Helplessness	Kendall's Tau B	0.081	-0.037
	p-value	0.645	0.809
	N	23	29

Table 8. Correlation Statistics Between Emotions from Video and Knowledge

Video Activity Emotions		Sim-First Group Knowledge (Post-Vid; t3)	Vid-First Group Knowledge (Post-Vid; t2)
Enjoyment	Kendall's Tau B	0.381	0.009
	p-value	0.025	0.952
	N	24	29
Boredom	Kendall's Tau B	-0.366	0.017
	p-value	0.031	0.914
	N	24	28
Frustration	Kendall's Tau B	-0.119	-0.124
	p-value	0.49	0.418
	N	24	29
Curiosity	Kendall's Tau B	0.119	0.237
	p-value	0.479	0.116
	N	24	29
Anxiety	Kendall's Tau B	-0.064	-0.045
	p-value	0.716	0.763
	N	24	29
Helplessness	Kendall's Tau B	-0.104	-0.052
	p-value	0.558	0.731
	N	24	29
Video Topic Emotions		Sim-First Group Knowledge (Post-Vid; t3)	Vid-First Group Knowledge (Post-Vid; t2)
Enjoyment	Kendall's Tau B	0.29	-0.022
	p-value	0.088	0.885
	N	24	29
Boredom	Kendall's Tau B	-0.273	0.058
	p-value	0.106	0.705
	N	24	28
Frustration	Kendall's Tau B	-0.07	-0.044
	p-value	0.681	0.774
	N	24	29
Curiosity	Kendall's Tau B	0.175	0.2
	p-value	0.302	0.185
	N	24	29
Anxiety	Kendall's Tau B	-0.346	-0.098
	p-value	0.048	0.514
	N	24	29
Helplessness	Kendall's Tau B	-0.263	-0.136
	p-value	0.139	0.373
	N	24	29

Supplementary Material A: Psychological Safety

We have deliberately incorporated psychological safety into our anti-harassment educational intervention, especially concerning the harassment bystander simulation component. We ensured such incorporation through the following methods: 1) following the Healthcare Simulation Standards of Best Practice (Watts et al., 2021); 2) consulting various content and education experts; 3) iterative development to script creation, including running through the simulation with graduate students and content and education experts, and adapting the script accordingly, 4) following strict ethics guidelines of our institutional review board; 5) analysis of data from our research findings; and 6) other considerations. We note that while providing education on harassment can indeed be sensitive to our learners, harassment has been widely and extensively reported to compromise healthcare professional trainees' psychological and, in some cases, physical, safety. Failing to equip medical trainees with tools to better manage harassment directed toward themselves or others stands to perpetuate, not protect them, from harassment. Internal Medicine and Surgery residents were selected because they are programs who have identified a need to better equip residents with tools to combat harassment. In this simulation, we focused on bystander interventions as we felt residents getting practice dealing with indirect harassment would be psychologically safer than having them practice managing harassment directed at them.

We first thoroughly incorporated the Healthcare Simulation Standards of Best Practice into our simulation design. The guideline highlights psychological safety as a required component in the overall simulation design. Specifically, the aspect of psychological safety is highlighted when concerning the pre-briefing, facilitation, debriefing, operations, outcomes and

objective, and professional integrity in developing and delivering simulation-based training [114–118]. Our simulation satisfies all of the outlined requirements by:

1) Clearly conveying the learning objectives and expectations during the pre-briefing and acknowledging the simulated nature of the learning environment. This included conveying how the harassment the learner will witness as a bystander is *simulated*, and therefore not real, and simulated for the purpose of practicing the harassment bystander intervention strategies.
2) Providing relevant instructional materials directly related to our simulation training that will help the learners understand and contextualize their simulation experience.
3) Allowing the participants to ask any questions prior to beginning the simulation for them to understand the ground rules.
4) Providing a non-competitive environment and clarifying that their performance in the simulation will not impact their academic or career standings in any ways.
5) Providing the simulation in an environment that allows for privacy, confidentiality, and open discussion.
6) The simulation ended with the facilitator and the involved standardized participants showing positive interactions and clearly acknowledging that any of the negative interactions during the simulation were indeed simulated. For example, the facilitator would always thank the standardized participants (i.e., actors) for their acting. In addition, the standardized participants themselves would acknowledge their acting and show positive interactions between themselves and the learner to illustrate how the simulation staffs are cohesive and no harassment occurred.

7) Providing debriefing that again emphasized the simulated nature of the harassment the learner witnessed, and the learning objectives and outcomes.
8) Explicitly stating in our simulation facilitation documentation an emphasis on the simulated nature of the harassment the learner will witness.
9) All stakeholders, including the simulation centre staff, standardized participants, administrators, facilitators, researchers, and learners were all aware of the simulated nature of the harassment featured in the simulation. We ensured this by hosting a series of meeting, and ensuring all stakeholders had received our simulation facilitation documents that outline all details concerning our simulation.
10) Providing resources such as the institution's student wellness centre that the learners can consult in case of any distress.

Second, we consulted various content and education experts concerning the nature of the harassment featured in our simulation, and how to best align the designed harassment interaction to the learning objectives, simulation fidelity, and psychological safety. Specifically, our simulation design has been vetted by:

1) Dr. Ning-Zi Sun: Associate Professor and Internal Medicine Residency Training Program Director, Department of Medicine, Faculty of Medicine and Health Sciences, McGill University
2) Dr. Jean-Sébastien Pelletier: Assistant Professor and Surgical Foundations Program Director, Department of Surgery, Faculty of Medicine and Health Sciences, McGill University

3) Dr. Liane Feldman: Chair of the Department of Surgery, Faculty of Medicine and Health Sciences, McGill University and Surgeon-in-Chief of the McGill University Health Centre (MUHC)
4) Dr. Gerald Fried: Associate Dean of Educational Technology and Innovation and Director of the Steinberg Centre for Simulation and Interactive Learning
5) Caroline White: Simulation Specialist Advisor at the McGill University Health Centre's Interprofessional Simulation Centre (MUHC-i-Sim)
6) Niki Soilis: Education Manager (at time of consultation) at the Steinberg Centre for Simulation and Interactive Learning, Faculty of Medicine and Health Sciences, McGill University
7) Dr. Melina Vassiliou: Associate Professor and Adar Chair of Surgical Education, Department of Surgery, Faculty of Medicine and Health Sciences, McGill Univeristy
8) Other research laboratory members with medical degrees: Dr. Osamu Nomura, and Dr. Lucia Patino Melo

Additionally, the following content and education experts were consulted with and vetted our approaches to ensuring psychological safety:

1) Dr. Farhan Bhanji, Associate Dean of Education, Faculty of Medicine and Health Sciences, McGill University

Third, we have gone through an iterative development process for our simulation. The lead developer gathered existing resources and guidelines on similar simulation training. He then conducted multiple phases of internal testing and rehearsals within the research team to create an initial draft of the simulation. This draft was circulated within the broader laboratory of the lead

developer for further testing and refinement. Other collaborators, along with content and education experts, were invited to provide feedback and iterate on the initial versions of the simulation. Throughout every major iteration, the team's focus included not only the fidelity of the simulation but also the psychological safety of the prospective participants. All feedback regarding psychological safety was thoroughly evaluated and carefully implemented into the simulation design. After multiple tests of different types of harassment scenarios and subsequent reflections, we decided that, for example, no racially or sexually discriminatory remarks should be made. The principal investigator oversaw the development of the simulations, and actively took part in the rehearsals and deliberation of the simulation. These measures helped ensure the psychological safety of our participants while providing realistic harassment scenarios with a high degree of believability.

Fourth, we have followed the ethics guideline at our institutional review board, with the board reviewing the everything related to our research and education objectives. This included all documents relating to our facilitation documents, simulation scripts, and all other related rationales and documentation for delivering our educational intervention. Following the review board's guidelines, we have deliberately:

1) Proposed rationale for our simulation training—it provides a safe space for trainees to practice the bystander strategies covered by our instructional video series, and provides an opportunity for learners to receive tailored feedback for refining their mastery.
2) Provided consent forms that detailed how our learners are not forced to participate in our intervention in any way. Our educational intervention did require the learner to be

present to start their training, or provide valid reasons for absence as our intervention was part of the *official curriculum* of our learners. However, the participants were allowed to stop at any time, and they had full autonomy of how they interacted in the simulation (i.e., they were not forced to engage in the simulation in any specific way). Furthermore, learners could participate in the intervention, but opt out of the research study.

3) Specify the associated benefits and risks concerning the educational intervention. Specifically, we have specified the possibility of learners experiencing upsetting feelings during their educational intervention. However, we have also explicitly specified that they can at any time stop participating in the intervention, ask questions at any time during the intervention. In addition, the learners have a direct line of contact to the McGill Student Wellness Hub in case of any distress.

Fifth, at the time this document was written we had run 72 residents from surgery related departments, and 69 residents from internal medicine through this simulation. We have not received any concerns or complains about psychological safety from the residents. Conversely, several residents, especially from surgery, suggested that they would like to be challenged more, for example, by not knowing which simulation station would include the harassment training scenario: a measure we had included to maximize psychological safety. In addition, the data we collected on residents' self-reported emotions towards our simulation training supports psychological safety. Specifically, concerning our first internal medicine resident cohort to go through our simulation training, the mean value of the residents' self-reported anxiety was 2.92/5.00 (n=51), where 3 indicates the residents felt "neutral" when asked whether they felt

anxiety towards the simulation activity (5 indicates that they "strongly agreed" that they felt anxiety, while 1 would indicate they "strongly disagreed" that they felt anxiety). On the other hand, mean value of enjoyment was 3.88/5.00 (n=52), reflecting how residents would generally agree that they felt some enjoyment towards the simulation activity. Based on our findings, we believe while the simulation was designed to be challenging, the residents did not report it being emotionally distressing.

Lastly, our other considerations included having our simulation training focus on bystander intervention first, as opposed to dealing with harassment directly targeting the learner. We believed that only after establishing mastery of dealing with harassment that indirectly affects the learner, we should explore how we can scaffold our learners to deal with harassment that directly targets them. This consideration of our learners' prerequisite skillset and knowledge also aligns with the Healthcare Simulation Standards of Best Practice. Further, we carefully avoided any harassment remarks that could be reasonably viewed as discriminatory, including but not limited to disability, race, gender, sex, sexual orientation, or religion. We did this to help minimize any potential triggers.

For the simulation scenario progression outline, please see Supplementary Material B. For the debriefing guideline for simulation facilitators, please see Supplementary Material C. For prebriefing guideline for simulation facilitators, please see Supplementary Material D. For consent form given to the participants for our internal medicine cohorts from the 2022-23 school year, see Supplementary Material E.

Supplementary Material B: Simulation Scenario Progression Outline

Timing (Approx.)	SPs Actions	Expected Intervention	Possible Cues
0-1 min	Junior resident (JR): gets into gown and gloves • Senior resident (SR) helps		
1 min	Medical Student (MS): rushes into the scene ** First harassment scene *e.g., "Nobody asked why you were late. Just stand over there"*	It may be too early to expect interventions	
1-3 min	JR: Prep insertion kit • Take catheter, cap two side ports • Flush saline for each port • Try thumbing the guide wire • Attach the finder needle to the syringe – drop some lidocaine ** Second harassment scene *e.g., "*Sigh… please don't your binder. What's wrong with you"*	Direct: Declare abuse as inappropriate Distract: Makes a joke or comment to ease tension Delegate: Call technician for help	
3-5 min	JR: Locate the site of insertion with the ultrasound. • Apply gel to suspected site Locate the jugular ** Third harassment scene *e.g., "What do you mean, in your opinion. Medicine is based on facts. If you didn't prepare, why are you even here?"*	Direct: Declare abuse as inappropriate Distract: Makes a joke or comment to ease tension Delegate: Call technician for help	Technician: Looks over at the scene
5-6 min	JR: Inject lidocaine to the surrounding area.		

6-8 min	JR: Locate jugular triangle; insert the finder needle guided by the ultrasound. • Apply negative suction pressure as you advance • Try drawing venous blood • Hold and stabilize needle – remove the syringe. Inset wire. • One hand should always be holding the wire ** Fourth harassment scene *e.g., "How do you not know to stay away from the sterile field? Are you an idiot?"*	Direct: Declare abuse as inappropriate Distract: Makes a joke or comment to ease tension Delegate: Call technician for help	Technician: Looks over at the scene
8-10 min	JR: Make NIC incision Then place dilator • Feed it over the wire Insert catheter • Wire should poke out of the middle port • Push the catheter into the skin Remove wire • Place last cap on the middle port • Flush each of the port with saline ** Fifth harassment scene *e.g., "This is dumb. I can't waste my time on this. Good luck with this one."*	Direct: Declare abuse as inappropriate Distract: Makes a joke or comment to ease tension Delegate: Call technician for help	Technician: Looks over at the scene
10-11 min	MS seems shocked. MS stands around for a moment, unsure what to do.	Delay: Check up on MS Delegate: Call technician for help	MS: "That was rough"

Supplementary Material C: Debriefing Guideline for Simulation Facilitators

This simulation debriefing utilizes the PEARLS Healthcare Debriefing Tool

Setting the Scene

- Create a safe context for learning
- State the goal of the debriefing
- Sample phrases:
    - "Ok, so we have 10 minutes for debriefing. Before we start, I'm going to emphasize that this was all a simulation, and nothing said during the simulation was for real. Now, the goal of the simulation was to apply intervention strategies as a bystander to stop harassment, while being engaged in a technical procedure. Let's try to answer in simple one or two sentences to the next following questions."

Reactions

- Explore feelings
- Solicit initial reactions and emotions
- Sample phrases:
    - "How did you feel overall about the simulation?"
        - "What part of the simulation made you feel this way?
    - "When did you realize the medical student was being harassed?"
    - "When you first realized the medical student was being harassed, how did you feel?"
    - "How did you feel about being able to practice being a bystander in a harassment situation?"

Descriptions

- Clarify facts
- Develop shared understanding of case
- Sample phrases:
    - How would you describe the intervention method you've chosen?
        - So, when the medical student started to get harassed, it seems like you chose ____ strategy, since you did ____.

Analysis

- Explore variety of performance domains
- Explore performance domains
- Sample phrases:
    - "What did you think you did well in intervening the harassment?"

    - "If you could re-do this simulation right now, is there anything you'd try differently?"
    - "Do you think you did well on your central line insertion?"

Summary/Application

- Identify take-aways
- Instructor centered (lack of time to be learner-centered)
- Sample phrases:
    - The key learning points were ____, and you can further your mastery by focusing on ____.

Supplementary Material D: Pre-briefing Guideline for Simulation Facilitators

Pre-Briefing and Orientation

Read to the simulation learner:

- "Welcome to the anti-harassment simulation training. I'll be your facilitator for the simulation. You'll be participating in this simulation with standardized participants but otherwise you will be working alone. The objective of this simulation is for you to practice technical skills while effectively applying bystander intervention strategies in a harassment scenario. The technical skill we have chosen for this scenario is a central line insertion."
- "The scenario is as follows: You are a junior resident in a simulation setting, performing an internal jugular central line insertion on a mannikin simulator. With you are a medical student and a senior resident. The medical student is there to observe you and learn, while the senior resident is there to supervise and grade both your and the medical students' performance. The senior resident will harass the medical student. There will be a simulation technician on standby in case of any issues. Think of an intervention strategy you can use to remedy the harassment."
- "For the central line insertion, the patient will already have been positioned, and the tools you need are already gathered. The draping is already complete. You will begin the procedure starting from getting into the gown and gloves, priming the instruments, and using the ultrasound to locate the insertion area. You will not be suturing at the end of the procedure. Instead, if you reach this step of the simulation, consider the central line insertion to be complete. Before you insert the central line, you are expected to inject lidocaine in the surrounding area—please assume the drug takes effect immediately without waiting"
- "Remember, the harassment occurring here is not real, and is simulated for your learning."
- "Before we begin, I'd like to answer questions you have for this simulation."

Supplementary Material E: Wilcoxon Signed Ranks Tests

Table S1. Wilcoxon Singed Ranks Test for Differences of Activity and Topic Emotions for the Simulation Training

Activity Emotions	Topic Emotions	Statistic	p	Mean difference	SE difference	Effect Size
Vid-First Group						
Enjoyment	Enjoyment	24.5	0.071	1	0.104	0.75
Boredom	Boredom	22.5	0.353	-0.5	0.179	-0.318
Frustration	Frustration	17	0.198	0.5	0.132	0.619
Curiosity	Curiosity	32.5	0.04	1	0.146	0.806
Anxiety	Anxiety	55	0.005	1.5	0.137	1
Helpless	Helpless	3	0.577	-0.5	0.11	-0.4
Sim-First Group						
Enjoyment	Enjoyment	24.5	0.851	0	0.172	0.0889
Boredom	Boredom	13.5	0.275	-0.5	0.149	-0.4
Frustration	Frustration	22.5	1	-1.45e−5	0.154	0
Curiosity	Curiosity	25	0.065	1	0.147	0.7857
Anxiety	Anxiety	81	0.012	1.5	0.23	0.7802
Helpless	Helpless	0	0.174	-1.414	0.102	-1

*Effect size is rank biserial correlation.

Table S2. Wilcoxon Singed Ranks Test for Differences of Activity and Topic Emotions for the Video Instructions

Activity Emotions	Topic Emotions	Statistic	p	Mean difference	SE difference	Effect Size
Vid-First Group						
Enjoyment	Enjoyment	12	0.824	2.86e−5	0.1051	0.143
Boredom	Boredom	46	0.052	1	0.1548	0.673
Frustration	Frustration	12	0.233	1	0.076	0.6
Curiosity	Curiosity	5	1	0	0.0702	0
Anxiety	Anxiety	7.5	0.424	1	0.069	0.5
Helpless	Helpless	1.5	1	0	0.0496	0
Sim-First Group						
Enjoyment	Enjoyment	22.5	1	-1.45e−5	0.172	0
Boredom	Boredom	10.5	0.49	1	0.149	0.4
Frustration	Frustration	2.5	0.105	-1	0.154	-0.762
Curiosity	Curiosity	17	0.533	-1	0.147	-0.244
Anxiety	Anxiety	12.5	0.472	-0.5	0.23	-0.306
Helpless	Helpless	4	0.773	0.172	0.102	0.333

*Effect size is rank biserial correlation.

Table S3. Wilcoxon Singed Ranks Test for Differences of Activity Emotions Between the Video Instructions and Simulation Training

Video Activity Emotions	Simulation Activity Emotions	Statistic	p	Mean difference	SE difference	Effect Size
Vid-First Group						
Enjoyment	Enjoyment	5.5	0.004	-1	0.154	-0.879
Boredom	Boredom	203.5	< .001	1.5	0.205	0.938
Frustration	Frustration	24	0.238	-0.5	0.167	-0.385
Curiosity	Curiosity	8	< .001	-1	0.147	-0.895
Anxiety	Anxiety	12	0.01	-1.5	0.196	-0.771
Helpless	Helpless	16	0.821	-4.29e−5	0.116	-0.111
Sim-First Group						
Enjoyment	Enjoyment	12	0.226	-0.5	0.247	-0.467
Boredom	Boredom	91	0.001	2	0.231	1
Frustration	Frustration	26	0.175	-1	0.265	-0.429
Curiosity	Curiosity	0	0.002	-1.5	0.169	-1
Anxiety	Anxiety	0	< .001	-2	0.215	-1
Helpless	Helpless	8	0.299	-1	0.114	-0.429

*Effect size is rank biserial correlation.

Table S4. Wilcoxon Singed Ranks Test for Differences of Topic Emotions Between the Video Instructions and Simulation Training

Video Topic Emotions	Simulation Topic Emotions	Statistic	p	Mean difference	SE difference	Effect Size
Vid-First Group						
Enjoyment	Enjoyment	17.5	0.042	-1	0.151	-0.615
Boredom	Boredom	91	< .001	1	0.14	1
Frustration	Frustration	32.5	0.336	-2.31e−6	0.138	-0.286
Curiosity	Curiosity	13	0.024	-1	0.11	-0.667
Anxiety	Anxiety	30	0.484	-0.196	0.177	-0.231
Helpless	Helpless	16	0.454	-0.5	0.135	-0.289
Sim-First Group						
Enjoyment	Enjoyment	10	0.059	-1	0.141	-0.6364
Boredom	Boredom	87.5	0.003	1.5	0.208	0.9231
Frustration	Frustration	43	0.881	-2.54e−5	0.209	-0.0549
Curiosity	Curiosity	9	0.052	-1	0.179	-0.6727
Anxiety	Anxiety	18	0.096	-1	0.232	-0.5385
Helpless	Helpless	13.5	0.301	-0.5	0.19	-0.4

*Effect size is rank biserial correlation.

References for Supplementary Materials

Bowler, F., Klein, M., & Wilford, A. (2021). Healthcare Simulation Standards of Best PracticeTM Professional Integrity. *Clinical Simulation in Nursing*, *58*, 45–48. https://doi.org/10.1016/j.ecns.2021.08.014

Decker, S., Alinier, G., Crawford, S. B., Gordon, R. M., Jenkins, D., & Wilson, C. (2021). Healthcare Simulation Standards of Best PracticeTM The Debriefing Process. *Clinical Simulation in Nursing*, *58*, 27–32. https://doi.org/10.1016/j.ecns.2021.08.011

McDermott, D. S., Ludlow, J., Horsley, E., & Meakim, C. (2021). Healthcare Simulation Standards of Best PracticeTM Prebriefing: Preparation and Briefing. *Clinical Simulation in Nursing*, *58*, 9–13. https://doi.org/10.1016/j.ecns.2021.08.008

Miller, C., Deckers, C., Jones, M., Wells-Beede, E., & McGee, E. (2021). Healthcare Simulation Standards of Best PracticeTM Outcomes and Objectives. *Clinical Simulation in Nursing*, *58*, 40–44. https://doi.org/10.1016/j.ecns.2021.08.013

Persico, L., Belle, A., DiGregorio, H., Wilson-Keates, B., & Shelton, C. (2021). Healthcare Simulation Standards of Best PracticeTM Facilitation. *Clinical Simulation in Nursing*, *58*, 22–26. https://doi.org/10.1016/j.ecns.2021.08.010

Watts, P. I., Rossler, K., Bowler, F., Miller, C., Charnetski, M., Decker, S., Molloy, M. A., Persico, L., McMahon, E., McDermott, D., & Hallmark, B. (2021). Onward and Upward: Introducing the Healthcare Simulation Standards of Best PracticeTM. *Clinical Simulation in Nursing*, *58*, 1–4. https://doi.org/10.1016/j.ecns.2021.08.006

Chapter 4 Bridging Text

In Chapter 3, I have identified what emotions medical residents experienced in general regarding the topic of anti-harassment education, watching instructional videos, and participating in simulation training. While the finding was that the educational intervention was generally enjoyable and elicited curiosity (i.e., desire to learn more about this topic), the next study delves deeper into emotional experiences of learners by employing a multimodal approach to measuring emotions. Chapter 2 identified only a single emotions SBE study that has employed multimodal analysis for emotions—this study was therefore a valuable addition to the scarce amount of multimodal emotions studies in the SBE literature. In Chapter 4, I showcased how multimodal analysis can go beyond the limitations of unimodal analysis by showcasing how EDA and speech content analyses can cover aspects of emotions self-report measures struggle to capture by itself.

While Chapter 3 did show the expected positive link between positive-activating emotions and knowledge, and the negative link between negative emotions and knowledge, no prominent findings were seen for simulation performance. With a more sophisticated method that goes beyond self-report measures, I use multimodal measures to bring additional insights towards the potential impact of emotions in anti-harassment education in Chapter 4.

I note that in regard to the different discrete emotions captured via the self-report measures in the study, the selection of emotions was intentional to ensure relevance to both the anti-harassment training context and the instructional modalities used (instructional video and high-fidelity simulation). The Multiple Object Foci Emotion Questionnaire (MOFEQ) was specifically chosen to capture these emotions, reflecting the variety of emotional responses expected in technology-rich environments. While the inclusion of additional emotions such as

anger, sadness, and disgust could have provided a more comprehensive emotional profile, the focus on capturing object-specific emotions aligned with the primary objective of the study and its methodological framework. Chapter 4 features a qualitative approach to capturing emotions that has the advantage of capturing a much wide range of emotions that are not restricted by pre-determined list of emotions in a questionnaire. To clarify, Chapter 4 and Chapter 3 are part of the same larger study. In other words, Chapter 4 draws on the same sample (with different inclusion criteria) as Chapter 3 to answer different research questions using different analyses.

Chapter 4 – Multimodal Cluster Analysis of Medical Residents' Emotions During High-Fidelity Harassment Bystander Simulation

Authors: Byunghoon (Tony) Ahn[1], Myriam Johnson[1], Negar Matin[1], Ning-Zi Sun[2,3], Jason M. Harley[1,3,4,5]

[1]*Department of Surgery, McGill University, Montreal, Canada*
[2]*Department of Medicine, McGill University, Montreal, Canada*
[3]*Institute for Health Sciences Education, McGill University, Montreal, Canada*
[4]*Research Institute of the McGill University Health Centre*
[5]*Steinberg Centre for Simulation and Interactive Learning*

Institution: Department of Surgical and Interventional Sciences (Formerly Department of Surgery)

Journal of Computer Assisted Learning. Under Review.

Correspondence for Reprints to:
Jason M. Harley, PhD
jason.harley@mcgill.ca

Parts of this work have also been presented as:

Ahn, B.T., Johnson, M., Matin, N., Delport, A., Han, B., Lee, S.Y., Sun, N., & Harley, J. (2024, August). *Emotional Profiles of Doctors During Intervening Harassment in a Clinical Simulation.* Paper to be presented at the International Conference on Motivation, Bern, Switzerland.

Ahn, B.T., Johnson, M., Matin, N., Delport, A., Han, B., Lee, S.Y., Sun, N., & Harley, J. (2024, May). *How do Residents Intervene Harassment During a Clinical Procedure: A Content Analysis of Simulation Debriefing.* Poster presented at the Surgical & Interventional Sciences Research Day, Montreal, Canada.

Ahn, B.T., Johnson, M., Matin, N., Delport, A., Han, B., Lee, S.Y., Sun, N., & Harley, J. (2024, May). *How do Residents Intervene Harassment During a Clinical Procedure: A Content Analysis of Simulation Debriefing.* Paper presented at the Annual Fraser N. Gurd Surgical Research Forum, Montreal, Canada.

Ahn, B.T., Johnson, M., Matin, N., Delport, A., Han, B., Lee, S.Y., Sun, N., & Harley, J. (2024, April). *Capturing Emotions with Content Analysis: How do Residents Deal with*

Harassment While Placing a Central Line? Poster presented at the International Congress on Academic Medicine (ICAM), Vancouver, Canada.

Abstract

Background: High fidelity simulations can be an effective tool for anti-harassment education. While emotions have been identified as crucial in simulation-based education, their role in anti-harassment education within medical training remains underexplored.

Objectives: We aimed to investigate emotional profiles of medical residents during harassment bystander simulation training via hierarchical clustering based on multimodal emotions data.

Methods: 27 internal medicine residents with complete data sets that were part of a larger study were recruited. Emotions were captured through self-report surveys, an electronic bracelet that records electrodermal activity, and speech content analysis based on the residents' simulation debriefing. The study involved residents performing a simulated central line insertion while a simulated harassment took place that they could use to practice intervening harassment.

Results: Our cluster analysis revealed three equal-sized groups: "Emotionally Balanced, Minimal Arousal", "Positive, Spiked Arousal", and "Negative High Arousal". The clusters had distinct levels of self-report emotions and electrodermal activity. Content analysis revealed distinct emotions, and sources of emotions between the clusters. Post-hoc analysis revealed that the "Emotionally Balanced, Minimal Arousal" group showed a higher propensity for directly confronting the harasser, indicating a composed emotional state conducive to focusing on simulation objectives.

Conclusions: Our findings reveal the varied emotional profiles that can be expected in simulation-based medical education and underscore the value of a multimodal approach to understanding these dynamics. Furthermore, the study highlights the criticality of recognizing

the sources of emotions and promoting effective emotion regulation strategies, especially in authentic learning environments where emotional responses are complex and impactful.

Lay Summary

What is current known about this topic?

- Emotions can play an important role in learning.
- Emotions are ubiquitous in learning settings, especially when it involves emotionally charged topics.
- Emotions can be expressed in multiple ways (e.g., facial expressions, sweat, etc.) and therefore can be measured in different ways.

What does this paper add?

- This paper provides a rich description of what emotions internal medicine residents experienced when tasked with a clinical procedure simultaneously with intervening harassment.
- This paper also exemplifies a combined approach of cluster analysis and speech content analysis to identify various emotional profiles of leaners.
- This paper showed how different emotional profiles may lead to different harassment intervention behaviours in clinical settings.

Implications for practice/or policy

- There are ample opportunities for simulation-based education environments and other technology rich environments to integrate technology that enable multimodal analysis of emotions.
- Simulation-based educators should strive for an emotion-aware practice of designing and delivering simulation trainings.
- As emotions can be associated with specific harassment response behaviours, emotion regulation is a future avenue of research.

Key Words

Medical Residents, Simulation-Based Education, Emotions, Electrodermal Activity, Content Analysis, Cluster Analysis

Background

Emotions can significantly impact learner's cognitive, motivational, and self-regulatory processes (Pekrun, 2006). Such significance has been acknowledged across various contexts, from traditional K-12 education (Cox, 2023; Pekrun, 2011) to higher education (Mendzheritskaya & Hansen, 2019), and even medical education (Artino et al., 2012; Duffy et al., 2020). Recent research advancements have emphasized studying emotions in more authentic, emotionally intense environments (Duffy et al., 2020; Harley, Jarrell, et al., 2019; Pekrun, 2006). Simulation-based education (SBE) offers such authentic environments, as it replicates key aspects of situations and experiences learners are meant to face (Dutta & Krummel, 2006; Fanning & Gaba, 2009). Anti-harassment education, in particular, such as bystander intervention training, can benefit from SBE (Duchesne et al., 2023; Herweck et al., 2021). Unlike group discussions or video-based instructions, SBE places learners in scenarios that elicit true-to-life emotions.

This study features a high-fidelity mannikin to capture the authentic emotional experiences of medical resident trainees during their simulation training. To capture these realistic emotions, we supplemented the traditional method of self-report surveys by capturing the audio of the trainee's post-simulation discussions (i.e., debriefing), and measuring psychophysiological data via a digital bracelet. We further captured video footage of the simulation to assess the trainee's performance and behaviours. We have applied these methodologies to a pressing educational topic in medicine: harassment.

Harassment, or mistreatment, involves repeated vexatious behaviour aimed at creating a threatening learning and working environment (Raver & Nishii, 2010). Victims' deeply

emotional responses to harassment are well-known (Scarduzio et al., 2018), and many medical trainees know that harassment is unacceptable (Gostlow et al., 2018). However, there are barriers to properly reporting and responding to harassment, including fear of hierarchy in medicine (Chung et al., 2018). Given medical training's emotional intensity (McConnell, 2019), and how harassment can occur during high-stakes clinical procedures, it is clear that emotions play a substantial role in learning about and addressing harassment.

To address the gap in examining emotions and its potential role in learning how to address harassment, our study aimed to provide a detailed examination of medical residents' emotional profiles during a bystander harassment simulation training. We adopted a multimodal approach, combining self-report, psychophysiological data, and speech content data to move beyond traditional reliance on self-report in simulation-based education research (Ahn et al., 2023). Our objectives were to identify the emotional profiles typical of such activities and explore how emotions might relate to simulation performance. We further aimed to exemplify how various analytics from technological tools can aid in our research objectives. Our study was grounded in Pekrun's Control Value Theory (CVT) of Achievement Emotions (Pekrun, 2006, 2019), supplemented by literature focusing on the use of EDA in studying emotions in learning contexts.

Control Value Theory of Achievement Emotions

Pekrun's CVT defines achievement emotions as a type of emotions linked to the success or failure of a task (Pekrun, 2006, 2019), tasks such as successfully managing harassment without escalating the situation. These emotions are categorized using a three-dimensional taxonomy: valence (positive or negative), activation (activating or deactivating, also referred to

as arousal), and object focus (focused on past outcome, present activities, or future outcomes). These three characteristics of achievement emotions are shaped by one's perceived control, value, and the results of an achievement activity. Prior research has shown positive-activating emotions such as enjoyment to foster learning, while negative-deactivating emotions such as hopelessness to impede it instead (Loderer et al., 2020; Pekrun, 2006). Other emotions have more mixed findings, although negative-activating emotions (e.g., anger, anxiety) will tend to hinder learning if not resolved during learning (Pekrun, 2006).

CVT indicates that the objective focus—the source of an emotion—can be the learning activity, its results, subject matter, and the social dynamics involved (Pekrun & Stephens, 2012). For example, in a harassment intervention simulation, a medical trainee may feel anxiety from a senior resident's verbal abuse, frustration due to harassment-related distractions to the medical procedure, and relief upon successfully completion of the simulation. In sum, a learner's emotional experience can be varied, shaped by the particular aspects of the learning process they are engaging with.

While reporting learners' discrete emotions such as enjoyment or anxiety is common, identifying *what* elicits these emotions is important. Prior studies in educational psychology, for example, have distinguished emotions elicited from the modality (e.g., mobile learning app) and the topic of the learning material by using instruments specialized for surveying emotions from varying object foci (e.g., multiple object foci emotion questionnaire; MOFEQ; Harley et al., 2016, 2020; Harley, Liu, et al., 2019). Self-report measures, however, can only target pre-determined object foci set by the researchers. Text or speech content analysis can overcome this

limitation as the learners' description of their own emotional experience can reveal potential antecedents to the described emotions.

CVT also underscores the merit of multimodal analyses. Specifically, the CVT conceptualizes emotions as a multi-componential psychological phenomena which include affective, cognitive, physiological, motivational, and expressive processes (Pekrun, 2011). For example, a medical resident's anxiety during a challenging simulation can manifest as feeling nervous (affective), frantically recalling the subsequent steps (cognitive), sweaty palms (physiological), determination to succeed (motivational), and lip-biting (expressive process). Therefore, researchers can not only use self-report measures to probe what learners felt during learning (affective), but they can also measure other facets of learners' emotions, such as their facial expressions, or their physiological arousal level (Harley, 2016). We discuss the latter below focusing on electrodermal activation (EDA).

EDA and Emotional States

EDA represents the fluctuations of electrical activity of the skin, influenced by the sympathetic nervous system's control on the sweat gland (Boucsein, 2012b; Turpin & Grandfield, 2007). As an indirect indicator of sympathetic nervous system activity, EDA can reveal the activation dimension of one's emotional experience (Azevedo et al., 2016; Harley, 2016). Empirical studies (Huber & Bannert, 2023; Törmänen et al., 2021, 2023) have demonstrated how EDA can represent the emotional activation dimension of emotions, with a recent systematic review showing support for EDA's relation to self-reported levels of activation (Roos et al., 2021).

Education studies primarily use exosomatic methods for measuring EDA (Dawson et al., 2007; Horvers et al., 2021). These methods measure skin conductance (SC; amount of sweat which helps conduct electricity) and works by applying a small electrical current to the surface of the skin with electrodes. Various wearable devices such as bracelets (Regalia et al., 2023) and rings (Torniainen et al., 2015) can measure SC in micro Siemens (μS; Turpin & Grandfield, 2007). SC itself is composed of tonic and phasic components. The tonic component, also known as skin conductance level (SCL), consists of slowly varying activity and provides a moving baseline for individual signals. It is relatively stable over short periods, and reflects the general arousal state or readiness to act (R. A. Cohen, 2014). The phasic component, also known as skin conductance response (SCR), encompasses the faster changing elements of the EDA signal and represents reactive responses. Unlike tonic EDA, phasic activity levels can fluctuate heavily, forming "peaks" or "spikes" in the data. These peaks can represent the response to a specific stimulus or event (e.g., senior resident yells at a medical student's mistake), although they may form without an identifiable eliciting stimulus (Braithwaite et al., 2013; Horvers et al., 2021). Tonic and phasic levels can reflect different physiological processes. For example, tonic levels can be more associated with slower, hormonal activities, in contrast to faster neural activities that are more frequently associated with phasic levels (R. A. Cohen, 2014).

Given that general EDA primarily reveals overall arousal level, prior studies have focused on specific EDA features such as tonic and phasic levels to distinguish general arousal from engagement or reactiveness (Harley, 2016; Horvers et al., 2021). For example, Van Bruinessen and others (2016) examined educating oncology patients with video clips. They analyzed both tonic and phasic EDA and interpreted the increased tonic levels of their learners as

their video clips eliciting general arousal. They further reported that their learners' phasic levels were higher than expected in the beginning, and slowly declined as the study continued; the study concluded that the higher phasic level may indicate the learners' alertness and readiness to respond to the videos, while the slow decline of phasic levels can be explained by habituation to the videos (Van Bruinessen et al., 2016).

Other studies echoed similar methods and findings. Harley, Jarrell and others' study (2019) on medical students in a diagnostic reasoning simulation analyzed learners' tonic levels and phasic peaks, depending on what EDA bracelet the learners wore. The authors revealed phasic peaks positively predicting learners' diagnostic efficiency. The study referenced the Yerkes-Dodson law (R. A. Cohen, 2011) to explain how phasic peaks may indicate optimal levels of arousal for good simulation performance (Harley, Jarrell, et al., 2019). A study from Li and Lajoie (2021) looked at simulated aviation training, and utilized facial expression recognition software to infer valence, and EDA measures to infer arousal levels of their learners. Their results revealed that phasic EDA was a positive predictor for performance during easy tasks. The authors concluded that phasic activity should indicate engagement, with low phasic activities during easy tasks potentially reflecting boredom and disengagement, and therefore lead to reduced performance. Consistent with these findings, studies in primary school (Törmänen et al., 2021), and nursing education (Dubovi, 2022), also emphasized phasic EDA's potential for identifying emotional engagement.

Multimodal Analysis of Emotions

Deeper understanding of emotions requires multimodal approaches (Azevedo & Gašević, 2019), which include data channels relating to contextual, behavioural, and physiological factors

of learners (Molenaar et al., 2023). For example, researchers can account for multiple aspect of the learners' emotions by examining self-reported emotions (contextual), verbal responses during an interview (behavioural), and EDA throughout learning (physiological). The calls for multimodal approaches for emotions continues (Harley, 2016; Noroozi et al., 2020; Singh Tomar et al., 2023), with examples such as Duffy et al. (2016), and Paloniemi et al. (2022) underscoring how EDA coupled with other data such as self-reports can minimize drawbacks respective data channels have. Specifically, while self-reports have disadvantages such as being susceptible to social desirability bias, or reporting errors, EDA data does not have these problems. Conversely, while EDA by itself cannot indicate the valence of a learner, self-reports and content analyses can provide insight on the valence of the EDA results.

Emotional Profiles

Various studies have aimed to identify the emotional profiles of learners as certain emotional states can lead to better learning outcomes (Loderer et al., 2020; Pekrun, 2019). Specifically, as CVT predicts learners predominantly experiencing emotions such as enjoyment should have better learning outcomes relative to their peers feeling emotions such as anxiety, researchers have investigated how learners can be grouped by their emotional experiences to identify different emotional profiles from learning activities. Researchers may then examine whether learners with a certain emotional profile show differences in learning. For example, Ahn and Harley (2020) descriptively categorized pre-service teachers by their dominant facial expressions representing emotions (e.g., anger, happiness, sadness, etc.) while they learned about queer history. They found that while learners expressing enjoyment (i.e., happiness) had good learning gains, learners with anger as their dominant emotion had the highest learning gains. Lee

and Chei (2020) used latent profile analysis to categorize undergraduate students into four emotional groups (moderate, positive, negative, and ambivalent) when taking online courses. They found that learners with positive emotional profiles had better learning outcomes, while the negative profile group had poorer outcomes.

Other studies have employed cluster analyses—a statistical method used to discover ways of grouping cases of data based on specific similarities (Antonenko et al., 2012; Landau & Chis Ster, 2010)—to identify various emotional profiles of learners. For example, Jarrell and colleagues' study (2016) on medical students in a diagnostic simulation used k-means clustering to identify 3 types of emotional profiles in their learners: positive emotion, negative emotion, and low emotion. Robinson and others (2017) on the other hand, examined undergraduate students learning about physiology and anatomy. They used a two-step clustering approach, where the first step involved a hierarchical clustering method, and the second step the k-means clustering. They identified 4 types of emotional profiles: positive, deactivated, negative and moderate-low.

The above studies while providing rich insights on differing emotional profiles of learners, they solely rely on self-report measures. We note that to our best knowledge little to no empirical studies have utilized multimodal analyses to examine emotional profiles in anti-harassment training.

Current Study

Our study examined medical residents' emotional profiles in a harassment bystander simulation training via multimodal analyses to addresses the current gaps in research. Specifically, we targeted a subject domain that is emotionally charged by nature but is seldom studied. We also utilized multimodal analyses by employing self-report and EDA measures to

illustrate the emotional profiles of our learners. We extended our clustering with content analysis, a qualitative method used to interpret meaning from text data, and made sense of learners' experiences, thoughts, and feelings (Bengtsson, 2016). Our content analysis therefore allowed further contextualization of the emotional profiles we identified via clustering, similar to how other studies have used of content analysis to identify emotions of their learners (Ansakorpi et al., 2017; Toivonen et al., 2017).

To achieve our study objectives, we investigated the following research questions: RQ1: How can medical residents be clustered by their emotions during their harassment bystander simulation activity? RQ2: What kind of emotional experiences does each cluster represent? RQ3: Do residents have different simulation performance when grouped by emotional experiences clusters?

Method

Participants

As part of a larger study, we recruited 60 of the 88 (68.18%) first and second-year internal medicine residents from a large North American university to participate in an educational simulation training activity we designed. Of the 60 residents who participated in our educational activity 52 provided consent to have at least one data channel collected (typically self-report) for research purposes and 35 participants provided consent for all data channels to be collected. For this article we required several data channels which we were able to secure from 27 participants (Male = 14; Female = 12; M_{age} = 28.0, SD_{age} = 2.85) of the 35 participants that consented to all relevant data channels. This number included data loss from recording problems and aligns with sample sizes typically found in health professions education research (Cook &

Hatala, 2015), despite having greater data requirements. Indeed, multimodal analyses are often constrained by the necessity of having complete data across all relevant channels—our study's requirement of four distinct data types (self-reported emotions, video recordings of simulation activities, audio recordings of simulation debriefings, and EDA data) meant any data loss or non-consent in these areas reduced our sample size. Over-all, our recruitment, especially for analyses with four different data channels, was successful given the highly specialized and small nature of our target population and medical residents' extremely limited availability for in-person activities.

Participants were scheduled for anti-harassment training as part of their curriculum and were assured that they would receive our training regardless of their study participation. Our approach combined purposeful and convenience sampling. We purposefully recruited our target training audience (internal medical residents), while inviting each participant to the study as they arrived for their training (Etikan, 2016; Leech & Donovan, 2023).

As per protocol of our larger study, 14 participants first watched the video first, while the remaining 13 participated in the simulation activity first. Participants arrived individually at our training site and began with the first assigned learning activity post-consent.

Educational Intervention

Video Lectures

Our educational intervention consisted of 5 screencast video lectures (total duration of approximately 37 minutes). Videos featured original illustrations and animations, along with voice overs to explain various concepts related to harassment, and effective bystander strategies. We adopted the 4D strategies by Right To Be (Sexual Violence Training Development Team,

2021; *The 5Ds of Bystander Intervention*, n.d.) for our content for the bystander intervention strategies. The 4D strategies refer to four different approaches bystanders can have when they witness harassment: 1) Direct: taking immediate action to confront the situation or the harasser directly; 2) Distract: Creating a diversion to defuse the situation and redirect the focus; 3) Delegate: Seeking assistance from another person, such as an authority figure; and 4) Delay: Checking in with the victim after the incident to provide support. The 4D strategies also guided our design of the simulation and debriefing.

Simulation Training & Debriefing

We used the CAE Blue Phantom Central Line Ultrasound Training Model for our simulation. This high-fidelity mannikin featured realistic external anatomical landmarks, as well as artificial skin that allows for real punctures with syringes and scans from ultrasounds to produce lifelike vascular access anatomy images. Further, we provided central line insertion kits, as well as other real and simulated equipment including simulated anesthetics, gown and gloves, and an ultrasound machine. Our simulation site was equipped with two different channels of video recording to ensure clear view of the simulation participants.

Our scenario's objective was for the learner to perform a medical procedure (central line insertion) while being an active bystander in a harassment situation. Our learner would play the role of the junior resident tasked with completing the central line insertion, while 2 standardized participants (actors) acted as the harasser and the victim. The harasser was a senior resident tasked with supervising our learner on the central line, while the victim was a medical student who was present to observe the procedure to learn. The harassment entailed constant interruptions and verbal abuse which increased in intensity as the simulation progressed. The

learner could intervene during the harassment in various ways as they worked on the central line insertion. One resident participated in the simulation at a time. We relied on the PEARLS framework (Eppich & Cheng, 2015) for the simulation debriefing. Simulation debriefings are structured, reflective process that usually follows the simulation activity where facilitators and participants discuss, analyze, and consolidate learning points from the simulation experience (Fanning & Gaba, 2009; Torr & Clauson, 2019). Both simulation and debriefing were developed according to Healthcare Simulation Standards of Best Practice (Watts et al., 2021). See Figure 1 for a snapshot from our simulation.

Measures

Self-Reported Emotions

We collected self-reported emotions via a modified version of the MOFEQ (Harley et al., 2016, 2020; Harley, Liu, et al., 2019), which is a modified version of the achievement emotions questionnaire (AEQ; Pekrun et al., 2002). The MOFEQ differs from the AEQ by having one item per emotion similar to other emotion measures (Duffy et al., 2020; Harley et al., 2015), in addition to having the question items repeat for multiple object foci (i.e., specific targets of emotions; Pekrun, 2006). The MOFEQ allows one to measure a set of emotions for each relevant object focus as opposed to asking them how they feel in a more open-ended manner. In this analysis we asked participants to report how they felt about learning during simulation training. In the larger study we also asked participants to report how they felt about learning during video instruction, but this data was outside the scope of the present research objectives and scope. We measured 6 emotions: enjoyment, curiosity, anxiety, frustration, boredom, and hopelessness. The MOFEQ uses a 5-point Likert scale that ranges from 1 (Strongly Disagree) to 5 (Strongly

Agree). Example items include: “The Anti-Harassment Simulation Training made me feel anxious”, and “I enjoyed learning during the Anti-Harassment Simulation Training”.

The current study’s MOFEQ resembles Harley and others’ later work with the instrument (Harley, Liu, et al., 2019), and differs from it by measuring anger instead of annoyance, and having adapted wording to focus on the simulation modality instead of a digital learning app. Harley and others have investigated the MOFEQ’s internal and external validity by confirming the expected directions of correlations between related sets of emotions, appraisals, and perceived success (Harley, Liu, et al., 2019). In addition, they have referenced prior research on the related AEQ (Harley et al., 2016; Pekrun, 2011; Pekrun et al., 2002) to provide convergent validity, along with direct consultation with Pekrun (Harley et al., 2020) for providing evidence of content validity for the MOFEQ.

Electrodermal Activity

EDA was recorded using Empatica E4 electronic bracelets (Empatica Inc., Cambridge, MA, USA) following established protocols (Garbarino et al., 2014; Schuurmans et al., 2020). The E4 wristband EDA sensor operated at a 4 Hz, measuring skin conductance in μS via external electrodes (Boucsein et al., 2012; Dawson et al., 2007; Edelberg, 1967). Participants continuously wore the E4 bracelets during their anti-harassment training, capturing EDA data across activities, including during filling out questionnaires, viewing videos, and simulation activities including the debriefing.

EDA signal processing involves multiple steps, including noise and movement artifact removal, variable extraction, and appropriate analyses (Boucsein, 2012a; Horvers et al., 2021; Posada-Quintero & Chon, 2020). We accordingly conducted a visual examination of all EDA

recordings with the Ledalab package (version V3.4.9) via MATLAB (version R2023a) for anomalies or data loss (Benedek & Kaernbach, 2010a, 2010b; Horvers et al., 2021). We applied a Butterworth low-pass filter with 1Hz cut-off and manually corrected artifacts by spline interpolation (Benedek & Kaernbach, 2010a, 2010b; Cowley et al., n.d.; Dindar et al., 2020; Geršak et al., 2020). We then employed Continuous Decomposition Analysis (CDA), known for its robustness with low sensitivity to noise and artifacts, to extract both phasic and tonic components (Benedek & Kaernbach, 2010a, 2010b; Dindar et al., 2020). We standardized both the phasic and tonic components of the EDA data against each participant's baseline EDA measurements to account for natural variability in individual EDA levels (Dindar et al., 2020; Harley, Jarrell, et al., 2019; Horvers et al., 2021; Törmänen et al., 2023). Following this standardization, we segmented the EDA data into four phases experienced by the learners: watching videos, completing questionnaires, simulation activity, and the simulation debriefing.

Standardizing EDA data produced z-scores (4 z-scores per second of training). To interpret the scores, we established specific thresholds to categorize EDA levels as low, moderate, or high. A z-score at or below -0.5 indicated low EDA level and represented the lowest 30.85% of EDA readings observed with a learner's entire training (Curtis et al., 2016). A z-score ranging from -0.5 to 0.5 signified a moderate EDA level, encompassing the middle 38.30% of EDA readings for that participant's range of EDA. Lastly, a z-score of 0.5 or higher denotes a high EDA level, indicating that the learner's EDA was in the highest 30.85% of readings within one's own EDA range. This categorization allowed us to determine whether a learner's EDA was comparatively low, moderate, or high at any given point during the educational activities.

Including EDA data from the entire duration of the training allowed us to obtain more contextualized EDA our learners (Collins et al., 2019; Horvers et al., 2021; Khan et al., 2019). As this current study focused on the simulation activity, we therefore computed mean phasic and tonic z-scores each learner experienced during the simulation.

Simulation Debriefing Audio

We recorded audio of the simulation debriefing via tablet computers in a separate interview area. The debriefing was conducted by the simulation facilitator (TA or NM), and the recording was used for content analysis. Details on content analysis is shown in the Analyses section. See supplementary material A for more details on the debriefing.

Simulation Performance

We measured simulation performance via a checklist. It was developed accordingly to the Healthcare Simulation Standards of Best Practice(Watts et al., 2021), and Standards for Educational Psychological Testing(American Educational Research Association, 2011) to support assessment validity. Specifically, our assessment adheres to standards regarding test specifications, test administration, and scoring procedures. Future work will include psychometric analyses and cross-validations to further validity evidence. The checklist assessed 3 categories of the 4D bystander intervention strategies: direct, distract, and delay (1 strategy—delegate—was irrelevant for our scenario). See Table 1 for details.

The first author and two trained research assistants assessed the simulations via video recordings and our checklist. Mean Fleiss' κ for 12 items was 0.74 (SD = 0.17; Min = 0.50; Max = 0.95), a substantially high score (Landis & Koch, 1977). Lead evaluator centric interrater reliability was also calculated for each checklist item (M = 91%; SD = 8%; Min = 67%; Max =

100%) because the first author had the most expertise in simulation training and was the facilitator for the simulations.

Procedure

The participants arrived at the simulation center one by one. We told each participant which activity they would first engage in. All participants were asked to fill out a set of questionnaires (including the MOFEQ) after either watching the instructional videos or participating in the simulation training. We then asked the participants to engage in their second learning activity; those who watched the instructional videos first would engage in the simulation training, and vice versa. After the second learning activity, we asked again all participants to fill out a set of questionnaires, including the MOFEQ. We then dismissed the participants from our intervention. The entire training was under 100 minutes. See Figure 2 for study diagram.

Analyses

Hierarchical Cluster Analysis

We ran clustering analysis to answer our first research question. We chose hierarchical clustering given our small sample size (Antonenko et al., 2012; Meyers et al., 2017). The analysis was carried out in JASP, with the dendrograms created with SPSS 29. Using the Euclidean distance measure, we standardized each datapoint prior to clustering to account for variances (Antonenko et al., 2012). We chose the average linkage measure to avoid overwhelming cluster counts (Antonenko et al., 2012) and skewed clustering from outlying cases (Yim & Ramdeen, 2015). The analysis included 4 variables relating to the simulation activity: positive activating emotions levels, negative activating emotion levels, mean phasic levels, and mean tonic levels. The positive and negative activating emotions were based on MOFEQ, where

the former was the mean score of enjoyment and curiosity, while the latter was the mean score of anxiety and frustration.

Interpretation of Dendrograms

Hierarchical clustering produces dendrograms, a chart that illustrates how each data point merges to form clusters. Researchers can then use their expertise on relevant theories, guidelines from prior studies, and visual inspections of the cluster distances to identify and explain the clusters (Meyers et al., 2017). We interpreted our dendrograms by relying on the elbow method plotted based on Akaike information criterion (AIC) and Bayesian information criterion (BIC); both agreed that the 3-cluster solution was optimal. Within sum of squares values for all 3 clusters were all under 15, indicating fairly compact clusters. See Figure 3 for the elbow method plot. See Figure 4 for the dendrogram produced.

Simulation Debriefing Content Analysis

To help answer our second research question, we conducted content analysis based on audio recordings from the residents' simulation debriefing. We relied on Bengtsson's (2016) methodology for our content analysis, where we followed the stages of decontextualization, recontextualization, categorization, and compilation. Following the stages led two of the authors (TA and SY) and 2 additional research assistants to orthographically transcribe the simulation debriefing audio recordings and engage in an open coding process. We then collaboratively generated a coding book, basing our initial codebook on MacQueen and others' work (1998, 2008) which accounts for multiple coders.

With the aim of exploring our learners' emotions during the simulation activity, we identified discrete emotions and possible object foci (i.e., the source) for the emotions. Our

debriefing included three questions relevant to our aim: 1) how did you (i.e., the resident) feel overall about the simulation? 2) how did you feel when you realized the harassment was happening? 3) how did you feel about the opportunity to practice being a bystander in a harassment scenario?

All four members took part in initial formal coding to improve interrater reliability. Interrater reliability calculation was based on MacQueen and others' work (2008), where we calculated agreement on phrase highlights and emotion/object foci codes between the lead (TA) and support reviewers (SY and two research assistants). Highlights could have more than one emotion or object foci codes assigned. The team reached the recommended agreement rate of 85% by the third trial. After, SY acted as the adjudicator, while TA and the assistants coded the transcripts. 25 of the 27 debriefings were coded; one lacked any significant codes, while another suffered data loss. Our final interrater reliability was 78.32% (*SD* = 6.91%). Our adjudicator (SY) made final decisions on how codes were applied to resolve discrepancies between coders. We used Taguette (Rampin & Rampin, 2021) for our coding process.

Post-hoc Analyses

We ran a one-way ANOVA to answer our third research question. We specifically looked at our clustering memberships and whether there were group differences in simulation performance between the emotional profile groups. Further, we ran Fisher's exact tests to see whether sim-cluster membership had differences when it came to specific intervention behaviours. Post-hoc power analysis via G*Power indicated the ANOVA would have a power of 0.40 for detecting large effect sizes ($f = 0.40$; Cohen, 1988). Post-hoc analysis for the Fisher's

exact tests revealed powers between the three clusters as 0.92, 0.19, and 0.39, where 0.92 was for the power for comparing cluster 1 and 2. Power of 0.39 was for comparing cluster 2 and 3.

Results

Research Question 1: Hierarchical Clustering Overall Results

Our hierarchical clustering yielded a three-cluster solution—we've labelled them as clusters 1 to 3. Table 2 summarizes the cluster information. The cluster sizes were equal, with all three clusters having 9 members. The explained proportion within-cluster heterogeneity ranged from 0.32 to 0.36, and therefore the within sum of squares amongst the clusters were also consistent with a range of 13.00 to 14.88. The cluster silhouette scores were on low, but none indicated lacking substantial structures as they all exceeded 0.25 (Struyf et al., 1996), ranging from 0.32 to 0.40.

We labelled the clusters based on the cluster means (see Figure 4). We again note that as the clustering variables are standardized, the cluster means represent the *relative* value compared to other clusters. Cluster 1 was labelled as "Emotionally Balanced, Minimal Arousal", as it had neutral levels of both positive and negative emotions, accompanied by low phasic and tonic EDA. Cluster 2 was labelled as "Positive, Spiked Arousal", as it had high positive emotions, low negative emotions, high phasic EDA, and neutral tonic EDA. Lastly, cluster 3 was labelled as "Negative, High Arousal", as it had low positive emotions, high negative emotions, and high levels of both phasic and tonic EDA.

Research Question 2: Emotional Profiles Represented by Clusters

We used descriptive statistics to summarize each cluster's emotional experience. Further, we incorporated our content analysis to identify prevalent emotions among learners and the sources of those emotions. See Figure 5 for cluster group means.

We stress that while clustering results represent relative cluster differences due to standardized clustering variables, the current section reports raw self-report scores, and EDA adjusted to each person's variation (i.e., standardized—the EDA standardization is separate from it being standardized again for the clustering procedure).

Cluster 1: Emotionally Balanced, Minimal Arousal

Cluster 1 (n = 9) reported high enjoyment (M = 4.11, SD = 0.33) and moderate curiosity (M = 3.67, SD = 0.71). Anxiety (M = 3.44, SD = 1.01) and frustration (M = 2.89, SD = 1.27) were on the moderate to low side. This cluster reported moderate levels of phasic (M = 0.114, SD = 0.25) and tonic EDA (M = -0.17, SD = 0.41).

Table 3 details the number of participants that indicated specific emotions during the simulation debriefing, as identified by the content analysis. Regarding cluster 1, the majority (> 50%) indicated: Anxiety (n = 9; 100%), gratefulness (n = 6; 66.67%), confusion (n = 6; 66.67%), and compassion (n = 6; 66.67%). The sources for most of these emotions were consistent. Specifically, learners indicating compassion all expressed compassion for the victim (as opposed to expressing compassion towards the harasser, as learner T134 did from cluster 2). Likewise, learners indicating gratefulness all indicated such emotions towards the opportunity to learn and practice bystander strategies. The object foci regarding confusion was split between the general simulation tasks (n = 4; 44.44%) and intervening during the harassment (n = 4; 33.33%), while anxiety had the greatest number of varied sources.

All cluster 1 members (n = 9; 100%) reported anxiety. One of the two most common object foci was the overall simulation (n = 6), which according to our codebook meant anxiety towards *both* performing the central line insertion and intervening during the harassment. For example, participant O250 stated, “Um. I felt it was, it was stressful. It was uncomfortable. Um, an uh I guess stress, stress, anxiety as well. Because you're kind of like I guess balancing two problems, you're putting in the central line, and to have this situation [the harassment] going on at the same time […] ”. The other most common object focus was the harassment perpetrator (i.e., the senior resident; n = 6). An example quote comes from O339: “you automatically worry how the senior resident would be with you if she was talking that way about the medical student”. Other object foci included the harassment intervention task (n = 3), the harassment in general (n = 2), and the victim (n = 1).

Cluster 2: Positive, Spiked Arousal

Cluster 2 (n = 9) reported very high enjoyment (M = 4.44, SD = 0.73) and curiosity (M = 4.44, SD = 0.53). Anxiety (M = 1.89, SD = 1.05) and frustration (M = 1.22, SD = 0.44) were generally very low. They showed high phasic EDA (M = 0.95, SD = 0.35), and moderate tonic EDA (M = 0.64, SD = 0.44).

Table 3 lists prevalent cluster 2 emotions from their content analysis: Confusion (n = 6; 85.71%); anxiety (n = 5; 71.43%); and gratitude (n = 5; 71.43%). We again note that two of the nine learners’ data could not be analyzed due to either data loss, or unclear indications of any emotions. The object focus for gratitude were consistently towards the opportunity to practice bystander strategies. However, anxiety sources varied, and included general harassment (n = 3), intervening during the harassment (n = 1), the overall simulation tasks (n = 1), and the harasser

(n = 1). Sources for confusion also varied, and included the simulation tasks in general (n = 2), the harasser (n = 2), general harassment (n = 2), and intervening in the harassment (n = 1). An example quote from O256 illustrates anxiety towards the general harassment, with them stating when asked what they felt as the harassment began, "so I felt a bit uncomfortable. Uhm. At that point, do you really understand like what was going on? And they just felt that the environment was, was not comfortable". As per our codebook, confusion often indicated uncertainty and challenge in completely understanding a situation. For example, T342 indicated confusion towards both on intervening the harassment, and towards the perpetrator by stating, "[…] when there was the direct interaction between the senior residents and medical student, I kind of was like in the middle in terms of hierarchy, I wasn't sure how to, to react to respect hierarchy […]".

Cluster 3: Negative, High Arousal

Cluster 3 (n = 9) showed moderate enjoyment (*M* = 2.89; *SD* = 0.78) and curiosity (*M* = 3.44; *SD* = 1.01), while having higher levels of anxiety (*M* = 3.89; *SD* = 0.78) and frustration (*M* = 3.78; *SD* = 0.67). This cluster also had high levels of phasic (*M* = .99; *SD* = 0.40) and tonic EDA (*M* = 0.92; *SD* = 0.44).

Table 3 shows prevalent cluster 3 emotions from their content analysis: Anxiety (n = 8; 88.89%); gratitude (n = 6; 66.67%); compassionate (n = 6; 66.67%); and frustration (n = 5; 55.56%). Like other clusters, gratitude and compassion's object foci were consistently the opportunities to practice, and the victim, respectively. The most common sources of anxiety were the general simulation (n = 6), followed by intervening in the harassment (n = 4), the general harassment (n = 3), the harasser (n = 2), the victim (n = 1), and the medical procedure (n = 1). Confusion's sources were intervening in the harassment (n = 3), the victim (n = 2) and the

harasser (n = 2). The general harassment, the general simulation, and the opportunity to practice each had 1 learner indicating them as source for confusion.

The most common sources for frustration were the general simulation (n = 3) and the harasser (n = 3). O161's quote illustrates the former: "And uh… you know, and I'm annoyed by the fact that I have to step in and teach, and teach like basic human decency and respect. Right, to somebody, while I'm also trying to do a central line". An example quote for frustration towards the harasser comes from O223: "I felt annoyed [the harassment] was going on in the middle of a patient procedure where there was no... The student was not kind of in the wrong. […] and then the other team member was creating patient safety issues by speaking kind of to the student across the room, while the procedure is going on. Kind of jeopardizing patient safety […]". Other sources for frustration included the medical procedure (n = 2), the opportunity to practice (n = 2), and the general harassment (n = 1).

Research Question 3: Emotional Profiles and Simulation Performance

Our ANOVA for the simulation performance between the 3 sim-clusters was not statistically significant, $F(2, 15.1) = 1.97$, $p = 0.173$. See Table 4 for simulation performance for the clusters. However, Fisher's exact tests revealed significant group differences between clusters regarding specific harassment bystander intervention behaviours. Specifically, while there were no significant differences in whether learners engaged in the "distract" (i.e., create a diversion; $p = .751$) or "delay" (i.e., check up on victim; $p = 0.489$) strategies, there was a significant difference in the "direct" strategy (i.e., directly confront the harasser; $p = 0.005$). Subsequent pairwise Z-tests revealed a significant difference between cluster 1 (Emotionally Balanced, Minimal Arousal) and cluster 2 (Positive, Spiked Arousal), $z = -3.3$, $p < 0.001$. See

supplementary material B for full set of pairwise comparisons. As Table 5 shows, all but one learner (n = 8; 88.89%) from cluster 1 engaged in the direct strategy, while all but one learner did not use direct for cluster 2.

Discussion

Emotional Profiles of Medical Residents

Our hierarchical clustering revealed three equal sized clusters, each with 9 learners. Cluster 1, "Emotionally Balanced, Minimal Arousal", showed moderate positive and negative self-report emotions, and exceedingly low phasic and tonic EDA relative to the other clusters. We interpret this cluster as representing those that were able to keep a calm emotional profile conducive to focusing on the simulation objectives. We base this on the following. First, as our *content analysis* indicates, all cluster 1 learners did experience anxiety. However, as indicated by their low EDA, they were not overwhelmed from this anxiety. While low phasic values may indicate lesser task-valuation or disengagement (Harley, 2016; Pekrun, 2006), the high self-reported enjoyment level, the compassion towards the victim shown in the *content analysis*, and cluster 1 descriptively having the highest simulation performance indicate otherwise. Therefore, while cluster 1 learners certainly did experience emotions, they stayed composed and stayed engaged with the simulation.

We believe an explanation to this finding would be that these learners were better at regulating their emotions (i.e., managing one's emotional experience and responses; Harley et al., 2019). Specifically, we believe these learners may have engaged in *cognitive change*, where one chooses to re-frame the situation and therefore experience more desirable type or intensity of emotions (e.g., transforming anxiety from the general simulation into a focus on the learning

opportunity and practice, as seen in the residents' expression of gratitude and compassion). Therefore, while these learners did experience anxiety and other emotions, their regulatory processes may have quickly dampened the intensity of activating emotions, effectively leading to low EDA. We conclude that cluster 1 is the embodiment of the mantra: “Keep calm and carry on”.

Cluster 2, “Positive, Spiked Arousal”, showed exceedingly high levels of positive, but very low levels of negative emotions. Additionally, they showed high phasic EDA, but relatively more modest tonic EDA. The mix of self-report emotions and EDA suggests cluster 2’s potential as effective learners, aligning with CVT and empirical findings that positive emotions enhance, and negative emotions impede foster learning outcomes (Loderer et al., 2020; Pekrun, 2006). Their modest tonic EDA indicated manageable arousal levels, while high levels of phasic EDA showed the learners were responsive to various stimuli during the simulation activity. Cluster 2’s simulation debriefing *content analysis* indicated that gratefulness to practice the harassment intervention strategies was common (n = 5), congruent with their overwhelmingly positive self-report emotions. However, an equal number of leaners (n = 5) indicated anxiety in the *content analysis*, while the predominant emotion in the analysis was confusion. These negative emotions stemmed from varied sources, which included the harassment in general, the harasser (senior resident), the overall simulation tasks, and performing the harassment intervention.

We believe while cluster 2 learners enjoyed the simulation, their experience marked by anxiety, confusion, and uncertainty suggests not all was positive. We believe while some of the EDA may be attributable to the positive side of the simulation, the uncertainty on how to help the victim, how to respond to the senior resident and prevent the harassment while respecting the

hierarchy, and trying to make sense of why such harassment would occur, may all contribute to the increased arousal. In other words, “Positive, Spiked Arousal” learners were more sensitive to the harassment, especially relative to the “Emotionally Balanced, Minimal Arousal” learners.

Our last cluster, cluster 3 was labelled as “Negative, High Arousal”. This cluster’s self-report showed relatively low positive emotions, but high negative emotions when compared to other clusters. They had the highest levels of *both* phasic and tonic EDA from all the clusters. Their emotional profile, characterized by prevalent anxiety, aligns with the typical emotional experiences captured in the SBE literature (B. Ahn et al., 2023; Madsgaard et al., 2022). Our *content analysis* showed that over half of this cluster expressed frustration during the simulation debriefing, consistent with their self-report and EDA data. The learners indicated frustration towards juggling the simulation’s two objectives, and how the need to intervene during the harassment made the central-line insertion more difficult. While harassment certainly elicited anxiety in the cluster 3, the prevalence of frustration suggests these learners’ focus was more on completing the medical procedure. They therefore may have found the harassment intervention secondary, or even a ‘nuisance’.

In sum, we had three evenly distributed emotional profiles of residents when they engaged in our harassment bystander simulation activity: 1) Those who kept calm and carried on with the scenario’s objectives; 2) those who enjoyed the simulation but were uncertain regarding what to do; and 3) those who tended to have limited interest in the harassment and wanted to focus on the central line.

Emotional Profiles and Simulation Performance

There was no difference between the clusters when it came to the simulation performance scores. Descriptively, emotionally balanced and calm learners (cluster 1) scored highest (M = .84; SD = 0.11), while those who enjoyed the simulation but indicated uncertainty and anxiety (cluster 2) scored lowest (M = 0.73; SD = 0.13).

When we looked at *how* the learners intervened in the harassment, we found statistically significant differences between the clusters regarding whether learners confronted the harasser. As Table 5 showed, the "Emotionally Balanced, Minimal Arousal" (cluster 1) residents tended to directly intervene, while "Positive, Spiked Arousal" (cluster 2) residents tended to avoid such direct confrontations. This aligns with our discussion on emotional profiles, where we concluded cluster 2 was more sensitive to harassment as shown by their EDA and content analysis emotions. Cluster 2 learners therefore may have preferred the less risky strategies such as the Distract and the Delay. Conversely, Cluster 1 learners appeared relatively unphased by the simulation objectives, and were not overwhelmed by anxiety. They could thus focus on the medical procedure and intervene the harassment, being comfortable enough to challenge the senior resident's harassment.

Given cluster 3's perception of harassment being distracting and annoying, one might expect these learners to directly confront the harassment as well. However, cluster 3 showed a split between engaging in the direct strategy or refraining from it. We believe that the pressure of hierarchy imposed by the senior resident may have discouraged direct confrontation, as indicated by the very high EDA levels of cluster 3 had relative to other clusters, especially cluster 1.

Although a larger sample might reveal performance score differences amongst emotional profile groups, the variation in the residents' chosen solutions based on their emotions is

noteworthy. While mastery of harassment bystander intervention hinges on *when* and *how* to apply various strategies as opposed to always explicitly confronting the harasser's behaviour, we conclude a calm emotional profile might be desirable in SBE, especially for non-technical skills acquisition involving professionalism or communication skills.

Implications for Methodology and Practice

Our study advances the methodological approach to studying emotions in education. First, we have demonstrated how multimodal analysis's efficacy in encompassing contextual, behavioural, and physiological factors of learners. The combination of self-report measures, phasic and tonic EDA, and speech content analysis allowed us to identify overall emotions, degree of arousal and engagement, and potential sources of emotions in relation to our SBE: a feat that surpasses unimodal analysis's capabilities. Second, our study illustrates the benefits of deeply integrating theoretical frameworks to explore emotions and their impact on SBE. Anchoring our approach to CVT enabled linking complex multimodal data for evidence-based interpretations. Our study thus addresses the SBE literature's call for more theory-driven research (B. Ahn et al., 2023).

Our study furthermore underscores the importance of recognizing diverse individual needs in SBE practice. Our cluster analysis showed an equal distribution of learners into three distinct emotional profiles with equal memberships. Two profiles differed significantly in their harassment intervention strategies, particularly in direct confrontation. Therefore, simulation educators and researchers should be cognizant of learners' varied emotional profiles and how emotions influence responses to harassment in clinical settings.

Limitations, Strengths, and Future Directions

Recognizing the merits of multimodal analysis, and especially in the context of simulation centers where there is wealth of technology to capture various data channels relevant to emotions, researchers may be inclined to include as many data channels as possible. Such approaches, however, can become exponentially complex for each additional measure (Harley, 2016). While our study is limited to relying on EDA (vs. adding other data such as heart rate), a concise self-report measure (vs. adding more items per emotion), and debriefing content analysis (vs. additionally analyzing vocal features), it also serves as a strength by integrating three complementary data sources to formulate digestible and theoretically grounded understanding of emotions. Subsequent research can investigate alternative data features or channels to explore other components of emotions. For instance, future research with more standardized simulations might examine phasic peak frequencies or use channels like heart rate variability.

Another limitation and future direction concern our simulation debriefing. Our study, based on medical residents' curriculum rather than a lab experiment, tailored debriefing questions to the simulation's learning objectives. Therefore, the debriefing number and depth of questions were constrained, as the facilitators often lacked the time to ask follow up questions. Future directions therefore include incorporating a more comprehensive qualitative methodology. More in-depth qualitative methods such as interpretative phenomenological analysis (Behrens et al., 2021; Smith & Fieldsend, 2021) complemented by follow-up interviews could deepen our understanding of emotions medical education. However, we acknowledge the suitability of content analysis within our study's integration into medical residents' teaching schedules. This approach enabled us to transcend numerical data, offering more comprehensive, contextualized emotional profiles of our learners.

A larger sample size, while not ensuring more concrete clustering, could enable inclusion of additional variables or reveal more nuanced clusters within dominant groups. Further, larger samples can yield larger clusters, it should enhance power for post-hoc analyses. The current lack of power leads to various limitations regarding generalizability, and capturing potential other effects in our analyses. However, our innovative multimodal analyses to form emotional profiles nonetheless provides preliminary findings that can aid educators and researchers on how emotions can impact anti-harassment training. Further, our sampling approach offers various merits. A small, purposeful sampling allowed us to feasibly focus on a niche population (Indrayan & Mishra, 2021) situated in an emotionally authentic environment (Duffy et al., 2020), and therefore secure ecological validity. In sum, while securing large samples in highly specialized population such as medical residents remain challenging, future directions include securing a larger cohort to add support to our findings.

EDA data processing is another strength our study has. First, we standardized EDA data pre-analysis to reflect each participants' baseline. Second, we applied an appropriate filter and artifact corrections. Third, we extracted specific data features (i.e., tonic and phasic EDA) for a more nuanced analysis of EDA. Our rigorous data processing and specific selection of EDA features show how educators and researchers can identify the various emotional profiles of their learners.

Conclusion

Anti-harassment training, especially when provided as a simulation that involves clinical procedures can be filled with emotions. Harassment and mistreatment are inherently negatively charged experiences, while missteps during central line insertions (e.g., puncturing the artery)

can lead to severe complications (Ortega & Barash, 2010). We investigated the emotional profiles of medical residents in a harassment bystander simulation via cluster analysis. Technology can extend data analysis much beyond the typical self-report survey, and allow for a multimodal approach. We incorporated self-report and EDA data, in addition to content analysis to provide rich descriptions of medical residents' emotions.

Our findings highlighted how educators and researchers can identify diverse emotional profiles in SBE, specifically in regard to anti-harassment education. Our learners were evenly divided into three clusters: 1) "Emotionally Balanced, Minimal Arousal": those who were generally calm despite reporting anxiety; 2) "Positive, Spiked Arousal": those who indicated a high degree of enjoyment during the simulation but indicated uncertainty on what to do; and 3) "Negative, High Arousal": those who tended to find the harassment a nuisance and wanted to focus on the medical procedure. Calm and collected residents tended to directly confront the harasser, especially compared to those who enjoyed the simulation but indicated uncertainty and anxiety towards how to intervene. Future studies that rely on other combinations of multimodal analysis of emotions can further illuminate how medical trainees feel and perform in simulation environments.

Conflict of Interest Statement

The authors do not have any conflict of interest to declare.

References

Ahn, B., Maurice-Ventouris, M., Bilgic, E., Yang, A., Lau, C. H.-H., Peters, H., Li, K., Chang-Ou, D., & Harley, J. M. (2023). A scoping review of emotions and related constructs in simulation-based education research articles. *Advances in Simulation*, *8*(1), 22. https://doi.org/10.1186/s41077-023-00258-z

Ahn, B. "Tony," & Harley, J. (2020). Facial expressions when learning with a Queer History App: Application of the Control Value Theory of Achievement Emotions. *British Journal of Educational Technology*, *51*(5), 1563–1576.

American Educational Research Association (Ed.). (2011). *Report and recommendations for the reauthorization of the institute of education sciences*. American Educational Research Association.

Ansakorpi, H., Sumelahti, M.-L., & Kaasila, R. (2017). Medical students' experience of emotions and success in neurological studies – What do they tell us? *BMC Medical Education*, *17*(1), 68. https://doi.org/10.1186/s12909-017-0905-4

Antonenko, P. D., Toy, S., & Niederhauser, D. S. (2012). Using cluster analysis for data mining in educational technology research. *Educational Technology Research and Development*, *60*(3), 383–398. https://doi.org/10.1007/s11423-012-9235-8

Artino, A. R., Holmboe, E. S., & Durning, S. J. (2012). Can achievement emotions be used to better understand motivation, learning, and performance in medical education? *Medical Teacher*, *34*(3), 240–244. https://doi.org/10.3109/0142159X.2012.643265

Azevedo, R., & Gašević, D. (2019). Analyzing Multimodal Multichannel Data about Self-Regulated Learning with Advanced Learning Technologies: Issues and Challenges. *Computers in Human Behavior*, *96*, 207–210. https://doi.org/10.1016/j.chb.2019.03.025

Azevedo, R., Taub, M., Mudrick, N., Farnsworth, J., & Martin, S. A. (2016). Interdisciplinary Research Methods Used to Investigate Emotions with Advanced Learning Technologies. In M. Zembylas & P. A. Schutz (Eds.), *Methodological Advances in Research on Emotion and Education* (pp. 231–243). Springer International Publishing. https://doi.org/10.1007/978-3-319-29049-2_18

Behrens, C. C., Driessen, E. W., Dolmans, D. H., & Gormley, G. J. (2021). 'A roller coaster of emotions': A phenomenological study on medical students lived experiences of emotions in complex simulation. *Advances in Simulation*, *6*(1), 24. https://doi.org/10.1186/s41077-021-00177-x

Benedek, M., & Kaernbach, C. (2010a). A continuous measure of phasic electrodermal activity. *Journal of Neuroscience Methods*, *190*(1), 80–91. https://doi.org/10.1016/j.jneumeth.2010.04.028

Benedek, M., & Kaernbach, C. (2010b). Decomposition of skin conductance data by means of nonnegative deconvolution. *Psychophysiology*. https://doi.org/10.1111/j.1469-8986.2009.00972.x

Bengtsson, M. (2016). How to plan and perform a qualitative study using content analysis. *NursingPlus Open*, *2*, 8–14. https://doi.org/10.1016/j.npls.2016.01.001

Boucsein, W. (2012a). *Electrodermal activity* (2nd ed). Springer.

Boucsein, W. (2012b). Principles of Electrodermal Phenomena. In W. Boucsein, *Electrodermal Activity* (pp. 1–86). Springer US. https://doi.org/10.1007/978-1-4614-1126-0_1

Boucsein, W., Fowles, D. C., Grimnes, S., Ben-Shakhar, G., Roth, W. T., Dawson, M. E., & Filion, D. L. (2012). Society for psychophysiological research ad hoc committee on electrodermal measures. Publication recommendations for electrodermal measurements. *Psychophysiology*, *49*(8), 1017–1034.

Braithwaite, J. J., Watson, D. G., Jones, R., & Rowe, M. (2013). A guide for analysing electrodermal activity (EDA) & skin conductance responses (SCRs) for psychological experiments. *Psychophysiology*, *49*(1), 1017–1034.

Chung, M. P., Thang, C. K., Vermillion, M., Fried, J. M., & Uijtdehaage, S. (2018). Exploring medical students' barriers to reporting mistreatment during clerkships: A qualitative study. *Medical Education Online*, *23*(1), 1478170. https://doi.org/10.1080/10872981.2018.1478170

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed). L. Erlbaum Associates.

Cohen, R. A. (2011). Yerkes–Dodson Law. In J. S. Kreutzer, J. DeLuca, & B. Caplan (Eds.), *Encyclopedia of Clinical Neuropsychology* (pp. 2737–2738). Springer New York. https://doi.org/10.1007/978-0-387-79948-3_1340

Cohen, R. A. (2014). Electrophysiology of Attention. In R. A. Cohen, *The Neuropsychology of Attention* (pp. 159–210). Springer US. https://doi.org/10.1007/978-0-387-72639-7_9

Collins, J., Regenbrecht, H., Langlotz, T., Said Can, Y., Ersoy, C., & Butson, R. (2019). Measuring Cognitive Load and Insight: A Methodology Exemplified in a Virtual Reality

Learning Context. *2019 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, 351–362. https://doi.org/10.1109/ISMAR.2019.00033

Cook, D. A., & Hatala, R. (2015). Got power? A systematic review of sample size adequacy in health professions education research. *Advances in Health Sciences Education*, *20*(1), 73–83. https://doi.org/10.1007/s10459-014-9509-5

Cowley, B., Fantato, M., Jennett, C., Ruskov, M., & Ravaja, N. (n.d.). *Learning When Serious: Psychophysiological Evaluation of a Technology- Enhanced Learning Game*.

Cox, R. D. (2023). Emotions in Postsecondary Teaching and Learning. In L. W. Perna (Ed.), *Higher Education: Handbook of Theory and Research* (Vol. 37, pp. 1–37). Springer International Publishing. https://doi.org/10.1007/978-3-030-66959-1_8-2

Curtis, A., Smith, T., Ziganshin, B., & Elefteriades, J. (2016). The Mystery of the Z-Score. *AORTA*, *04*(04), 124–130. https://doi.org/10.12945/j.aorta.2016.16.014

Dawson, M. E., Schell, A. M., & Filion, D. L. (2007). The electrodermal system. *Handbook of Psychophysiology*, *2*, 200–223.

Dindar, M., Järvelä, S., & Haataja, E. (2020). What does physiological synchrony reveal about metacognitive experiences and group performance? *British Journal of Educational Technology*, *51*(5), 1577–1596. https://doi.org/10.1111/bjet.12981

Dubovi, I. (2022). Cognitive and emotional engagement while learning with VR: The perspective of multimodal methodology. *Computers & Education*, *183*, 104495. https://doi.org/10.1016/j.compedu.2022.104495

Duchesne, E., Caners, K., Rang, L., & Dagnone, D. (2023). Addressing microaggressions with simulation: A novel educational intervention. *Canadian Journal of Emergency Medicine*, *25*(4), 299–302. https://doi.org/10.1007/s43678-023-00474-6

Duffy, M. C., Lajoie, S., & Lachapelle, K. (2016). Measuring Emotions in Medical Education: Methodological and Technological Advances Within Authentic Medical Learning Environments. In S. Bridges, L. K. Chan, & C. E. Hmelo-Silver (Eds.), *Educational Technologies in Medical and Health Sciences Education* (Vol. 5, pp. 181–213). Springer International Publishing. https://doi.org/10.1007/978-3-319-08275-2_10

Duffy, M. C., Lajoie, S. P., Pekrun, R., & Lachapelle, K. (2020). Emotions in medical education: Examining the validity of the Medical Emotion Scale (MES) across authentic medical learning environments. *Learning and Instruction*, *70*, 101150. https://doi.org/10.1016/j.learninstruc.2018.07.001

Dutta, S., & Krummel, T. M. (2006). Simulation: A New Frontier in Surgical Education. *Advances in Surgery*, *40*, 249–263. https://doi.org/10.1016/j.yasu.2006.06.004

Edelberg, R. (1967). *Electrical properties of the skin In: Brown CC, editor., ed. Methods in Psychophysiology*.

Eppich, W., & Cheng, A. (2015). Promoting Excellence and Reflective Learning in Simulation (PEARLS): Development and Rationale for a Blended Approach to Health Care Simulation Debriefing. *Simulation in Healthcare: The Journal of the Society for Simulation in Healthcare*, *10*(2), 106–115. https://doi.org/10.1097/SIH.0000000000000072

Etikan, I. (2016). Comparison of Convenience Sampling and Purposive Sampling. *American Journal of Theoretical and Applied Statistics*, *5*(1), 1. https://doi.org/10.11648/j.ajtas.20160501.11

Fanning, R., & Gaba, D. (2009). Simulation-Based Learning as an Educational Tool. In *Anesthesia Informatics* (pp. 459–479). Springer New York. https://doi.org/10.1007/978-0-387-76418-4_24

Garbarino, M., Lai, M., Tognetti, S., Picard, R., & Bender, D. (2014). Empatica E3—A wearable wireless multi-sensor device for real-time computerized biofeedback and data acquisition. *Proceedings of the 4th International Conference on Wireless Mobile Communication and Healthcare - "Transforming Healthcare through Innovations in Mobile and Wireless Technologies."* 4th International Conference on Wireless Mobile Communication and Healthcare - "Transforming healthcare through innovations in mobile and wireless technologies," Athens, Greece. https://doi.org/10.4108/icst.mobihealth.2014.257418

Geršak, V., Vitulić, H. S., Prosen, S., Starc, G., Humar, I., & Geršak, G. (2020). Use of wearable devices to study activity of children in classroom; Case study—Learning geometry using movement. *Computer Communications*, *150*, 581–588. https://doi.org/10.1016/j.comcom.2019.12.019

Gostlow, H., Vega, C. V., Marlow, N., Babidge, W., & Maddern, G. (2018). Do Surgeons React?: A Retrospective Analysis of Surgeons' Response to Harassment of a Colleague During Simulated Operating Theatre Scenarios. *Annals of Surgery*, *268*(2), 277–281. https://doi.org/10.1097/SLA.0000000000002434

Harley, J. M. (2016). Measuring Emotions. In *Emotions, Technology, Design, and Learning* (pp. 89–114). Elsevier. https://doi.org/10.1016/B978-0-12-801856-9.00005-0

Harley, J. M., Bouchet, F., Hussain, M. S., Azevedo, R., & Calvo, R. (2015). A multi-componential analysis of emotions during complex learning with an intelligent multi-agent system. *Computers in Human Behavior*, *48*, 615–625. https://doi.org/10.1016/j.chb.2015.02.013

Harley, J. M., Jarrell, A., & Lajoie, S. P. (2019). Emotion regulation tendencies, achievement emotions, and physiological arousal in a medical diagnostic reasoning simulation. *Instructional Science*, *47*(2), 151–180. https://doi.org/10.1007/s11251-018-09480-z

Harley, J. M., Lajoie, S. P., Tressel, T., & Jarrell, A. (2020). Fostering positive emotions and history knowledge with location-based augmented reality and tour-guide prompts. *Learning and Instruction*, *70*, 101163. https://doi.org/10.1016/j.learninstruc.2018.09.001

Harley, J. M., Liu, Y., Ahn, T. B., Lajoie, S. P., Grace, A. P., Haldane, C., Whittaker, A., & McLaughlin, B. (2019). I've got this: Fostering topic and technology-related emotional engagement and queer history knowledge with a mobile app. *Contemporary Educational Psychology*, *59*, 101790. https://doi.org/10.1016/j.cedpsych.2019.101790

Harley, J. M., Pekrun, R., Taxer, J. L., & Gross, J. J. (2019). Emotion Regulation in Achievement Situations: An Integrated Model. *Educational Psychologist*, *54*(2), 106–126. https://doi.org/10.1080/00461520.2019.1587297

Harley, J. M., Poitras, E. G., Jarrell, A., Duffy, M. C., & Lajoie, S. P. (2016). Comparing virtual and location-based augmented reality mobile learning: Emotions and learning outcomes.

Educational Technology Research and Development, *64*(3), 359–388. https://doi.org/10.1007/s11423-015-9420-7

Herweck, A. M., Kumnick, A., Perone, H., Muller, C., & Bornstein, K. (2021). "What a Pretty Assistant You Have": Addressing Sex-Based Harassment by Patient Offenders in Undergraduate Medical Training. *Medical Science Educator*, *31*(3), 1187–1191. https://doi.org/10.1007/s40670-021-01289-9

Horvers, A., Tombeng, N., Bosse, T., Lazonder, A. W., & Molenaar, I. (2021). Detecting Emotions through Electrodermal Activity in Learning Contexts: A Systematic Review. *Sensors*, *21*(23), 7869. https://doi.org/10.3390/s21237869

Huber, K., & Bannert, M. (2023). What happens to your body during learning with computer-based environments? Exploring negative academic emotions using psychophysiological measurements. *Journal of Computers in Education*, *10*(1), 189–215. https://doi.org/10.1007/s40692-022-00228-w

Indrayan, A., & Mishra, A. (2021). The importance of small samples in medical research. *Journal of Postgraduate Medicine*, *67*(4), 219. https://doi.org/10.4103/jpgm.JPGM_230_21

Khan, T. H., Villanueva, I., Vicioso, P., & Husman, J. (2019). Exploring relationships between electrodermal activity, skin temperature, and performance during. *2019 IEEE Frontiers in Education Conference (FIE)*, 1–5. https://doi.org/10.1109/FIE43999.2019.9028625

Landau, S., & Chis Ster, I. (2010). Cluster Analysis: Overview. In *International Encyclopedia of Education* (pp. 72–83). Elsevier. https://doi.org/10.1016/B978-0-08-044894-7.01315-4

Landis, J. R., & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics*, *33*(1), 159. https://doi.org/10.2307/2529310

Lee, J.-Y., & Chei, M. J. (2020). Latent profile analysis of Korean undergraduates' academic emotions in e-learning environment. *Educational Technology Research and Development*, *68*(3), 1521–1546. https://doi.org/10.1007/s11423-019-09715-x

Leech, N. L., & Donovan, C. (2023). Mixed methods sampling and data collection. In *International Encyclopedia of Education(Fourth Edition)* (pp. 485–490). Elsevier. https://doi.org/10.1016/B978-0-12-818630-5.11047-4

Li, T., & Lajoie, S. (2021). Predicting aviation training performance with multimodal affective inferences. *International Journal of Training and Development*, *25*(3), 301–315. https://doi.org/10.1111/ijtd.12232

Loderer, K., Pekrun, R., & Lester, J. C. (2020). Beyond cold technology: A systematic review and meta-analysis on emotions in technology-based learning environments. *Learning and Instruction*, *70*, 101162. https://doi.org/10.1016/j.learninstruc.2018.08.002

MacQueen, K. M., McLellan, E., Kay, K., & Milstein, B. (1998). Codebook Development for Team-Based Qualitative Analysis. *CAM Journal*, *10*(2), 31–36. https://doi.org/10.1177/1525822X980100020301

MacQueen, K. M., McLellan-Lemal, E., Bartholow, K., & Milstein, B. (2008). Team-based codebook development: Structure, process, and agreement. *Handbook for Team-Based Qualitative Research*, *119*, 119–135.

Madsgaard, A., Smith-Strøm, H., Hunskår, I., & Røykenes, K. (2022). A rollercoaster of emotions: An integrative review of emotions and its impact on health professional

students' learning in simulation-based education. *Nursing Open*, *9*(1), 108–121. https://doi.org/10.1002/nop2.1100

McConnell, M. (2019). Learning with emotions: The relationship between emotions, cognition, and learning in medical education. *Archives of Medicine and Health Sciences*, *7*(1), 11. https://doi.org/10.4103/amhs.amhs_46_19

Mendzheritskaya, J., & Hansen, M. (2019). The role of emotions in higher education teaching and learning processes. *Studies in Higher Education*, *44*(10), 1709–1711. https://doi.org/10.1080/03075079.2019.1665306

Meyers, L. S., Gamst, G., & Guarino, A. J. (2017). *Applied multivariate research: Design and interpretation* (Third Edition). SAGE.

Molenaar, I., Mooij, S. D., Azevedo, R., Bannert, M., Järvelä, S., & Gašević, D. (2023). Measuring self-regulated learning and the role of AI: Five years of research using multimodal multichannel data. *Computers in Human Behavior*, *139*, 107540. https://doi.org/10.1016/j.chb.2022.107540

Noroozi, O., Pijeira-Díaz, H. J., Sobocinski, M., Dindar, M., Järvelä, S., & Kirschner, P. A. (2020). Multimodal data indicators for capturing cognitive, motivational, and emotional learning processes: A systematic literature review. *Education and Information Technologies*, *25*(6), 5499–5547. https://doi.org/10.1007/s10639-020-10229-w

Ortega, R., & Barash, P. (2010). Ultrasound-Guided Internal Jugular Vein Cannulation. *N Engl j Med*.

Paloniemi, S., Penttonen, M., Eteläpelto, A., Hökkä, P., & Vähäsantanen, K. (2022). Integrating Self-Reports and Electrodermal Activity (EDA) Measurement in Studying Emotions in

Professional Learning. In M. Goller, E. Kyndt, S. Paloniemi, & C. Damşa (Eds.), *Methods for Researching Professional Learning and Development* (Vol. 33, pp. 87–109). Springer International Publishing. https://doi.org/10.1007/978-3-031-08518-5_5

Pekrun, R. (2006). The Control-Value Theory of Achievement Emotions: Assumptions, Corollaries, and Implications for Educational Research and Practice. *Educational Psychology Review*, *18*(4), 315–341. https://doi.org/10.1007/s10648-006-9029-9

Pekrun, R. (2011). Emotions as Drivers of Learning and Cognitive Development. In R. A. Calvo & S. K. D'Mello (Eds.), *New Perspectives on Affect and Learning Technologies* (pp. 23–39). Springer New York. https://doi.org/10.1007/978-1-4419-9625-1_3

Pekrun, R. (2019). Achievement emotions: A control-value theory perspective. In *Emotions in late modernity* (pp. 142–157). Routledge.

Pekrun, R., Goetz, T., Titz, W., & Perry, R. P. (2002). Academic Emotions in Students' Self-Regulated Learning and Achievement: A Program of Qualitative and Quantitative Research. *Educational Psychologist*, *37*(2), 91–105. https://doi.org/10.1207/S15326985EP3702_4

Pekrun, R., & Stephens, E. J. (2012). Academic emotions. In K. R. Harris, S. Graham, T. Urdan, S. Graham, J. M. Royer, & M. Zeidner (Eds.), *APA educational psychology handbook, Vol 2: Individual differences and cultural and contextual factors.* (pp. 3–31). American Psychological Association. https://doi.org/10.1037/13274-001

Posada-Quintero, H. F., & Chon, K. H. (2020). Innovations in Electrodermal Activity Data Collection and Signal Processing: A Systematic Review. *Sensors*, *20*(2), 479. https://doi.org/10.3390/s20020479

Rampin, R., & Rampin, V. (2021). Taguette: Open-source qualitative data analysis. *Journal of Open Source Software*, *6*(68), 3522. https://doi.org/10.21105/joss.03522

Raver, J. L., & Nishii, L. H. (2010). Once, twice, or three times as harmful? Ethnic harassment, gender harassment, and generalized workplace harassment. *Journal of Applied Psychology*, *95*(2), 236–254. https://doi.org/10.1037/a0018377

Regalia, G., Resnati, D., & Tognetti, S. (2023). Sensors on the Wrist. In *Encyclopedia of Sensors and Biosensors* (pp. 1–20). Elsevier. https://doi.org/10.1016/B978-0-12-822548-6.00130-8

Robinson, K. A., Ranellucci, J., Lee, Y., Wormington, S. V., Roseth, C. J., & Linnenbrink-Garcia, L. (2017). Affective profiles and academic success in a college science course. *Contemporary Educational Psychology*, *51*, 209–221. https://doi.org/10.1016/j.cedpsych.2017.08.004

Roos, A.-L., Goetz, T., Voracek, M., Krannich, M., Bieg, M., Jarrell, A., & Pekrun, R. (2021). Test Anxiety and Physiological Arousal: A Systematic Review and Meta-Analysis. *Educational Psychology Review*, *33*(2), 579–618. https://doi.org/10.1007/s10648-020-09543-z

Scarduzio, J. A., Sheff, S. E., & Smith, M. (2018). Coping and Sexual Harassment: How Victims Cope across Multiple Settings. *Archives of Sexual Behavior*, *47*(2), 327–340. https://doi.org/10.1007/s10508-017-1065-7

Schuurmans, A. A. T., De Looff, P., Nijhof, K. S., Rosada, C., Scholte, R. H. J., Popma, A., & Otten, R. (2020). Validity of the Empatica E4 Wristband to Measure Heart Rate

Variability (HRV) Parameters: A Comparison to Electrocardiography (ECG). *Journal of Medical Systems*, *44*(11), 190. https://doi.org/10.1007/s10916-020-01648-w

Sexual Violence Training Development Team. (2021). *Active bystander intervention: Training and facilitation guide : training for preventing and responding to sexual violence in B.C. post-secondary institutions*. BCcampus, BC Open Textbook Project.

Singh Tomar, P., Mathur, K., & Suman, U. (2023). Unimodal approaches for emotion recognition: A systematic review. *Cognitive Systems Research*, *77*, 94–109. https://doi.org/10.1016/j.cogsys.2022.10.012

Smith, J. A., & Fieldsend, M. (2021). Interpretative phenomenological analysis. In P. M. Camic (Ed.), *Qualitative research in psychology: Expanding perspectives in methodology and design (2nd ed.).* (pp. 147–166). American Psychological Association. https://doi.org/10.1037/0000252-008

Struyf, A., Hubert, M., & Rousseeuw, P. (1996). Clustering in an Object-Oriented Environment. *Journal of Statistical Software*, *1*(4). https://doi.org/10.18637/jss.v001.i04

The 5Ds of Bystander Intervention. (n.d.). Right to Be. https://righttobe.org/guides/bystander-intervention-training/

Toivonen, A. K., Lindblom-Ylänne, S., Louhiala, P., & Pyörälä, E. (2017). Medical students' reflections on emotions concerning breaking bad news. *Patient Education and Counseling*, *100*(10), 1903–1909. https://doi.org/10.1016/j.pec.2017.05.036

Törmänen, T., Järvenoja, H., & Mänty, K. (2021). Exploring groups' affective states during collaborative learning: What triggers activating affect on a group level? *Educational*

Technology Research and Development, *69*(5), 2523–2545. https://doi.org/10.1007/s11423-021-10037-0

Törmänen, T., Järvenoja, H., Saqr, M., Malmberg, J., & Järvelä, S. (2023). Affective states and regulation of learning during SOCIO-EMOTIONAL interactions in secondary school collaborative groups. *British Journal of Educational Psychology*, *93*(S1), 48–70. https://doi.org/10.1111/bjep.12525

Torniainen, J., Cowley, B., Henelius, A., Lukander, K., & Pakarinen, S. (2015). Feasibility of an electrodermal activity ring prototype as a research tool. *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 6433–6436. https://doi.org/10.1109/EMBC.2015.7319865

Torr, J. W., & Clauson, K. A. (2019). Using Patients and Other Forms of Simulation in Teaching Clinical Skills. In *Encyclopedia of Pharmacy Practice and Clinical Pharmacy* (pp. 205–217).

Turpin, G., & Grandfield, T. (2007). Electrodermal Activity. In *Encyclopedia of Stress* (pp. 899–902). Elsevier. https://doi.org/10.1016/B978-012373947-6.00139-2

Van Bruinessen, I. R., Van Den Ende, I. T. A., Visser, L. N. C., & Van Dulmen, S. (2016). The impact of watching educational video clips on analogue patients' physiological arousal and information recall. *Patient Education and Counseling*, *99*(2), 243–249. https://doi.org/10.1016/j.pec.2015.08.022

Watts, P. I., Rossler, K., Bowler, F., Miller, C., Charnetski, M., Decker, S., Molloy, M. A., Persico, L., McMahon, E., McDermott, D., & Hallmark, B. (2021). Onward and Upward:

Introducing the Healthcare Simulation Standards of Best PracticeTM. *Clinical Simulation in Nursing*, *58*, 1–4. https://doi.org/10.1016/j.ecns.2021.08.006

Yim, O., & Ramdeen, K. T. (2015). Hierarchical Cluster Analysis: Comparison of Three Linkage Measures and Application to Psychological Data. *The Quantitative Methods for Psychology*, *11*(1), 8–21. https://doi.org/10.20982/tqmp.11.1.p008

Figure 1. A photo of our simulation training in action. The medical student (left; standardized participant) is being verbally harassed by the senior resident (right; standardized participant). The learner (centre) takes on the role of the junior resident that is tasked with performing an internal jugular central line insertion, all the while intervening in harassment that occurs.

Figure 2. Study diagram that shows when learning activities and measurements took place.

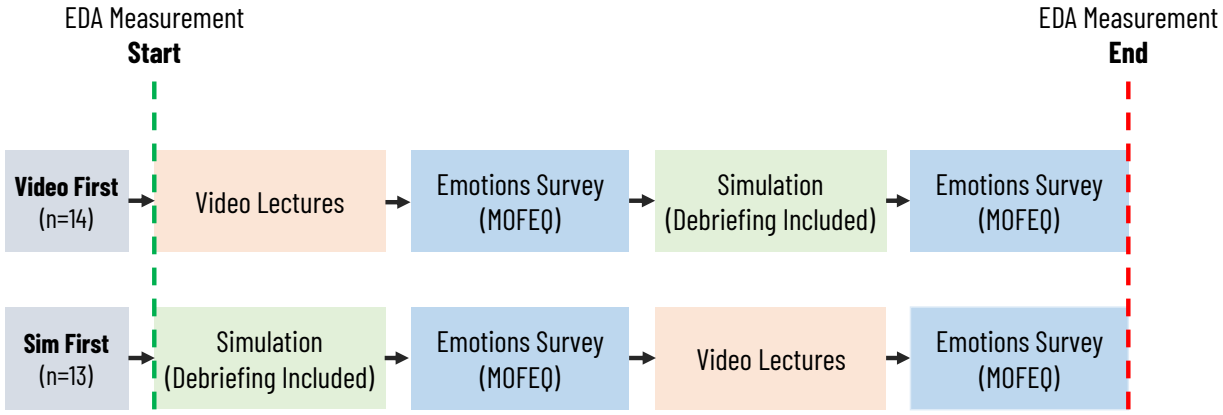


Figure 3. Elbow method plots for the clustering analysis. Recommended cluster count according to both BIC and AIC indicates a 3-cluster solution.

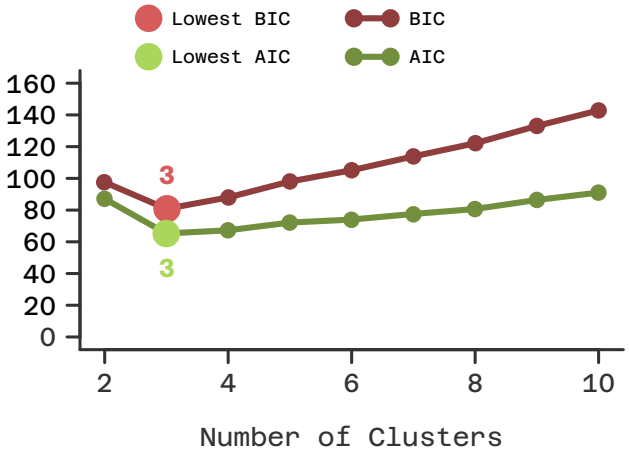


Figure 4. Dendrogram of hierarchical cluster analysis for the simulation activity.

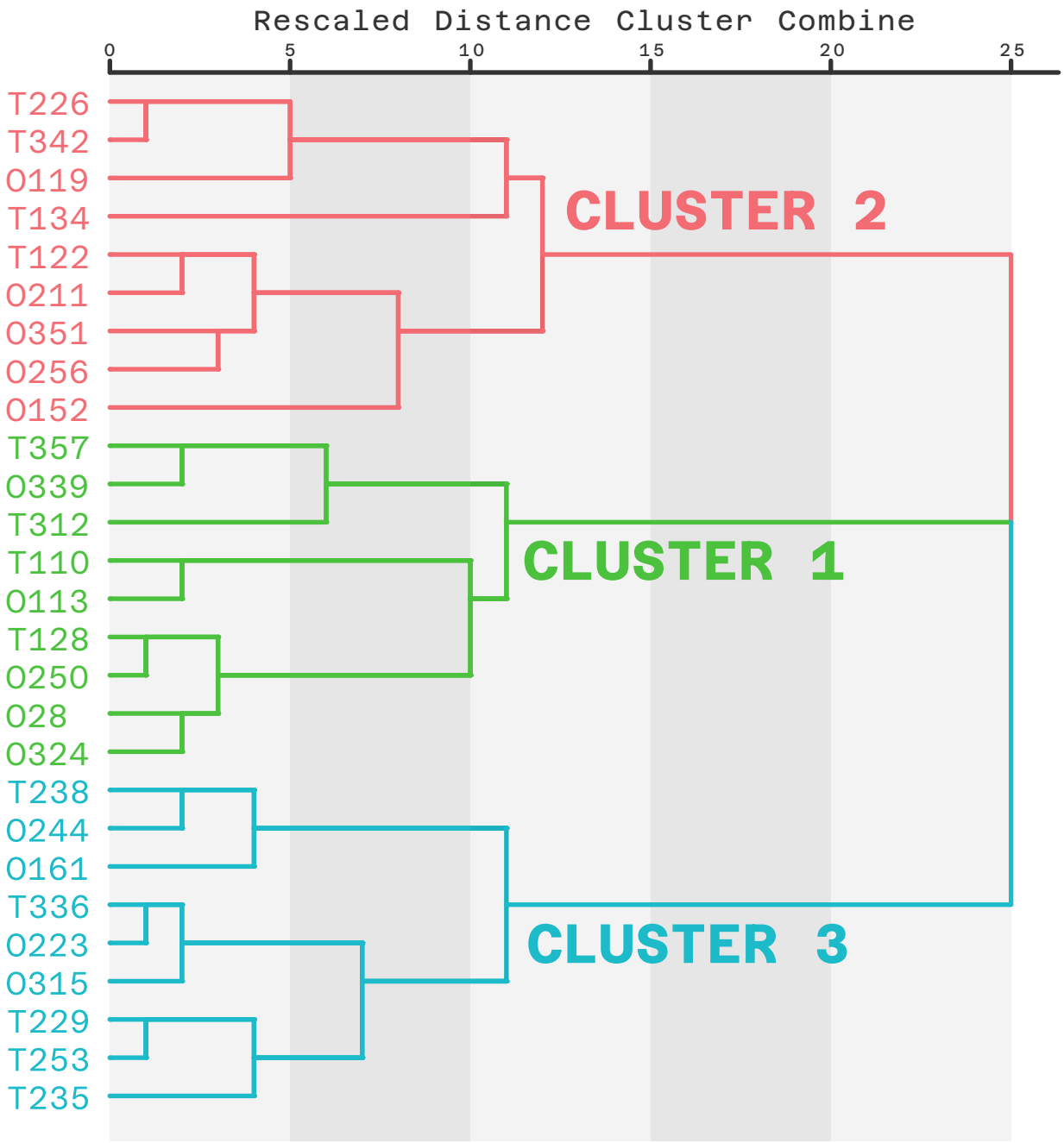


Figure 5. Cluster Group Means.

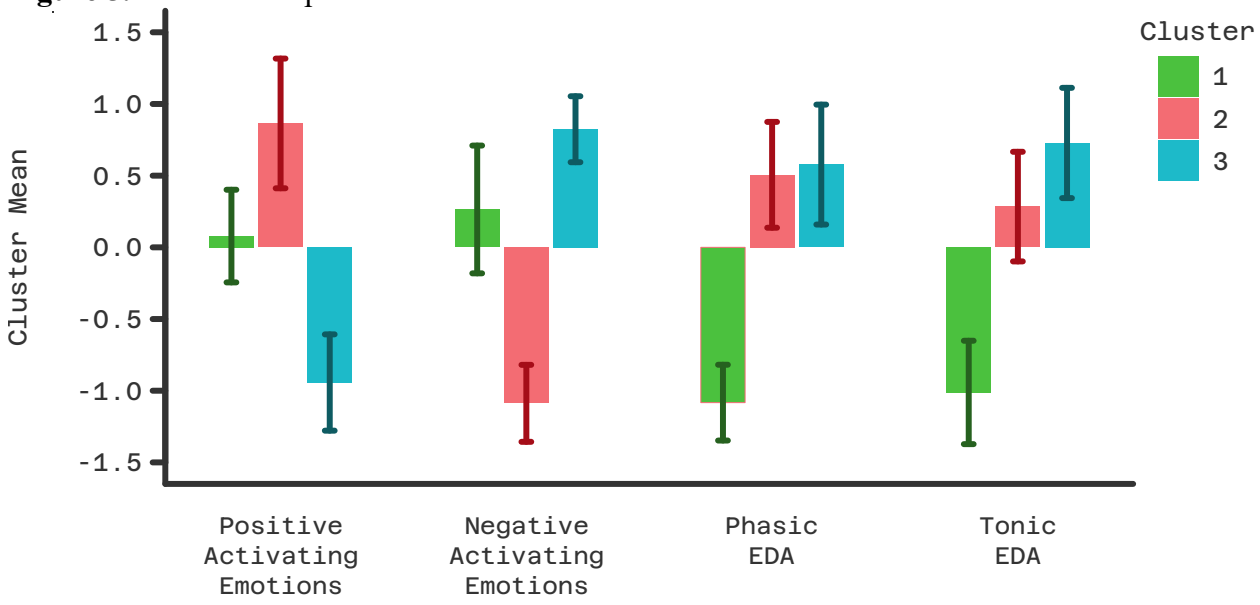


Table 1. Assessment Criteria for Bystander Intervention.

Strategy	Checklist Item	Description	Points
Direct	Confront harasser	Question or challenge the harasser's behaviours or actions as inappropriate or unacceptable.	2
	De-escalation	Use tone and language that does not intend to increase the tension or stakes.	1
	Workplace respect	Emphasize the importance of professionalism, and equity in the workplace and the learning environment.	1
	Resilience	Continues intervention even if initial attempt is unsuccessful.	1
Distract	Shift attention away	Directly engage with the victim or harasser during the harassment (or immediately after). Draw attention away from the harassment.	2
	De-escalation	Use tone and language that does not intend to increase the tension or stakes.	1
	Uplift victim	Raise the victim's self-esteem, specifically aspects that have been the target of harassment.	1
	Resilience	Continues intervention even if initial attempt is unsuccessful.	1
Delay	Check in with victim	Ask the victim how they feel following the harassment.	2
	Offer third-party resources	Refer the victim to organizational offices that can help with reporting harassment or supporting victims.	1
	Problematize the harassment	Ensure that the victim knows that the harasser's actions were unacceptable and inappropriate.	1
	Uplift victim	Raise the victim's self-esteem, specifically aspects that have been the target of harassment.	1
Comprehen siveness	No strategies	No strategies were used.	0
	1 strategy	Just 1 strategy was used.	3
	2 strategies	Just 2 strategies were used.	4
	3 strategies	All 3 strategies were used.	5

Average subtotals of each criterion for final score.

Table 2. Cluster information for the 3-cluster solution.

Cluster	1	2	3
Size (number of members in the cluster)	9	9	9
Explained proportion within-cluster heterogeneity	.324	.361	.315
Within sum of squares	13.379	14.879	12.998
Silhouette score	.339	.316	.394

Table 3. Number and percentage of learners indicating certain emotions during the simulation debriefing. Cluster 2 has originally 9 members, but 2 of the learners' data was excluded for the content analysis. * added when the majority of cluster members reported the emotion.

	Cluster 1 (n = 9)		**Cluster 2 (n = 7)**		**Cluster 3 (n = 9)**	
Emotions	Number of reporting learners	% of reporting learners	Number of reporting learners	% of reporting learners	Number of reporting learners	% of reporting learners
Anger	1	11.1%	1	14.3%	0	0.0%
Anxiety	9*	100.0%*	5*	71.4%*	8*	88.9%*
Ashamed	1	11.1%	0	0.0%	2	22.2%
Bored	0	0.0%	1	14.3%	0	0.0%
Compassion	6*	66.7%*	3	42.9%	6*	66.7%*
Confusion	6*	66.7%*	6*	85.7%*	4	44.4%
Curiosity	1	11.1%	0	0.0%	1	11.1%
Enjoyment	3	33.3%	3	42.9%	2	22.2%
Fear	2	22.2%	2	28.6%	2	22.2%
Frustrated	2	22.2%	3	42.9%	5*	55.6%*
Grateful	6*	66.7%*	5*	71.4%*	6*	66.7%*
Hopeful	1	11.1%	1	14.3%	1	11.1%
Sad	2	22.2%	0	0.0%	2	22.2%
Surprise	2	22.2%	1	14.3%	3	33.3%

Table 4. Simulation performance of the clusters.

Cluster	N	Mean	SD
S1: Emotionally Balanced, Minimal Arousal	9	.839	.1112
S2: Positive, Spiked Arousal	9	.729	.1334
S3: Negative, High Arousal	9	.767	.0763

Table 5. Number of residents employing the direct strategy (i.e., direct confrontation) grouped by the clusters.

Cluster	Did not use "Direct"	"Used Direct"
1: Emotionally Balanced, Minimal Arousal	1	8
2: Positive, Spiked Arousal	8	1
3: Negative, High Arousal	4	5

Supplementary Material A: Simulation Debriefing

Debriefing

The debriefing for the simulation activity utilized the Promoting Excellence and Reflective Learning in Simulation (PEARLS) healthcare debriefing tool [119]. This debriefing framework has the following main components: 1) Setting the Stage: establish a psychologically safe environment for discussion; 2) Reactions Phase: Encourage learners to share emotional and cognitive reactions to the simulation; 3) Description Phase: Provide factual recap of key events during the simulation and establish a common understanding with the learners on what has transpired; 4) Analysis Phase: Guide the learners through reflective discussions about the simulation, such as best practices and areas of improvement; and 5) Application to Practice: Discuss how the practiced and learned skill can transfer to other scenarios.

Keeping the PEARLS framework in mind, our debriefing sheet for our simulation facilitation covered the following.

Setting the Scene

We aimed to create a safe context for learning and state the goal for the debriefing. Suggested phrases included: 1) I need to emphasize that the harassment from the simulation is not real; and 2) The goal of the simulation was to try our intervention strategies as a bystander, while being engaged in a technical procedure.

Reactions

We aimed explore feelings of our learners. We further aimed to solicit initial reactions and emotions from the learners. Suggested phrases included: 1) How did you feel overall about the simulation—what part of the simulation what you feel this way? 2) When did you realize the

medical student was being harassed—when you realized this, what were you feeling? 3) How did you feel about being able to practice being a bystander in a harassment situation?

Description

We aimed to clarify facts, and develop shared understanding of the scenario. Suggested phrases included: 1) How would you describe the different intervention methods you tried? 2) I agree with your description; 3) I also think you've shown [strategy name], since I observed you doing [behaviour that indicates strategy; 4] So, when the medical student started to get harassed, it seems like you chose [strategy name], since you did [behaviour that indicates strategy].

Analysis

We aimed to explore various performance domains of the learners. Suggested phrases included: 1) So you've applied the [strategy name]. Do you think it was effective? 2) What else did you think you did well in intervening in the harassment? 3) If you could re-do this simulation right now, is there anything you'd try differently? 4) I think you could also consider [specific behaviour] for the [strategy name]. 5) It's difficult to say that there is an absolute correct answer, but I think it's helpful to consider how we tend to intervene and how we can do better with the intervention techniques we are comfortable with; 6) Do you think you did well on your central line insertion—do you think the central line insertion was more difficult due to the harassment?

Summary and Application

We aimed to identify key learning points from the simulation activity. Suggested phrases included: 1) The key learning points were to try out bystander intervention strategies while being engaged in a technical procedure; and 2) You can further your master of [strategy name] by focusing on [specific behaviour].

Supplementary Material B: Pairwise Comparisons

This section details the full pairwise comparisons run for the Fisher's exact test evaluating the group difference in whether learners engaged in the "direct" strategy (i.e., directly confront the harasser; $p = .005$). Pairwise comparisons using Z-tests were carried out. Table B1 below show comparisons between the clusters. We note that there was also a significant difference between clusters 2 and 3 (Negative, High Arousal), z = 2, P = .046, but this significance fades once Bonferroni correction is applied (P = .138) to account for inflated type 1 error.

Table B1. Simulation performance difference between the clusters.

Z tests	test statistic	p	p_{bonf}
Between clusters 1 and 2	-3.3	< .001	< .003
Between clusters 1 and 3	-1.58	.114	.342
Between clusters 2 and 3	2	.046	.138

Chapter 5: Discussion

5.1 General Findings

5.1.1 Theory Driven Emotions Research

My dissertation first highlighted the current state of emotions research in the context of medical SBE. While medical educators acknowledge the importance of theoretical integrations in research,[7,110,120] it is clear researchers need to further normalize theory driven research. There is context for the current field's tendency for lacking proper theoretical integration when studying emotions. Physicians, educators, or anyone unfamiliar with formal, scientific conceptualization of emotions can and do use the term "emotion" to successfully converse in everyday situations.[8,56]

However, the merit of integrating theories into emotion research is clear in three areas as it provides insights towards: 1) what emotions and emotional components researchers can examine in the first place; 2) what measurements can be used to capture emotions and emotional components; and 3) how to interpret complex emotions data. First, I believe the lack of theory driven research can lead to a tunnel visioned approach to only the most salient emotions or aspects of emotions. Just as past emotions research I have described in the introduction had overly focused on math anxiety, atheoretical endeavors in SBE research may lead to a neglect of studying other emotions or sources of emotions. Although typical SBE scenarios are filled with fear and anxiety, other emotions ranging from enjoyment in successfully completing the scenario, disappointment in underperforming certain aspects of the procedure, or frustration towards specific obstacles placed in the simulation can all play a role in revealing how SBE can

be effective. A theoretical conceptualization of emotions can introduce new researchers to a whole spectrum of emotions, as showcased by the CVT.[57]

Second, by conceptualizing emotions as a multicomponent construct via emotion theories, it becomes clear what aspects of emotions researchers can capture in their studies. While self-report measures are traditionally the most intuitive, approachable, and inexpensive to administer,[32] other data channels can cover the drawbacks of such self-reports by bringing new dimensions of emotions into context. From my experience, simulation centers are filled with advanced technologies, from audio/video capture devices to computer or virtual reality simulators. In addition, wearable devices for measuring physiological responses continue to advance, with more innovative and approachable devices available for researhcers.[121] Therefore, integration of theories can help awaken SBE researchers to seeing the potential SBE environments offer in providing diverse emotions data along with a better understanding of the role of emotions in learning.

Finally, theories can provide insight concerning how to interpret complex emotions data, especially when handling data from multiple channels or data concerning multiple emotions. For example, relying on the CVT can allow researchers to make evidenced-backed inferences on: 1) the valuation and control a learner has towards a learning situation; 2) the sources and antecedents of emotions learners experienced; 3) implications these emotions have towards learning; 4) alternative explanations in cases where findings do not align well with prior empirical works.[28,54]

5.1.2 Anti-Harassment Education: Activity Sequence, Learning, and Emotions

My research showed the effectiveness of the anti-harassment educational intervention I where I was the lead developer. Specifically, knowledge gains were evident from residents that have completed our training. However, when accounting for simulation performance, there was a group difference between the two groups—the sim-first group showed higher levels of knowledge gain than their peers. I explained the results through linking the impact of activity sequence with the idea of productive failure. Specifically, the concept of productive failure was illustrated by the sim-first group, which engaged in challenging simulations before instructional videos, leading to initial struggles but ultimately higher knowledge gains compared to the vid-first peers. This sequence allowed the sim-first group to identify knowledge gaps early, making subsequent video instruction more effective by focusing their learning on these gaps, exemplifying the benefits of experiencing failure before receiving direct instruction.

The sequence of activities, and therefore general instructional design, also had an impact on learners' emotions. While we found that our learners generally reported wanting to learn more about the topic of anti-harassment and enjoyed the training, the sim-first trainees reported more neutral levels of positive emotions when it came to viewing the instructional videos. I believe that the simulation activity was more engaging for the learners, as expected, based on the experiential learning SBE offers relative to video-based education.[122,123] In addition, the anchoring effect (i.e., the influence of first impressions on later opinions) may explain how the vid-first group still reported high enjoyment levels for the videos, unlike their sim-first peers.[124] Since the sim-first group encountered the more engaging activity first, the residents' initial standard for what counts as enjoyable and curiosity-inducing is set higher than those who

reviewed the videos first. In other words, the video instructions by themselves are enjoyable to the residents—when compared side by side to the simulation, they may be deemed as less engaging.

The sim-first group, however, reported higher levels of anxiety towards the simulation activity relative to the vid-first group. This would also make sense as the vid-first trainees learned what specific strategies they could use in harassment interventions and were provided with examples of harassment and intervention scenarios from the instructional videos. The sim-first trainees, however, lacked this knowledge, and therefore must have had more uncertainty concerning what to expect or how to successfully meet the simulation objectives; this uncertainty can easily turn into anxiety.[125] These findings highlight how learning design can impact the emotional experiences of learners.

5.1.3 The Emotional Experience of Anti-Harassment Education

My empirical studies showed the different ways of approaching emotions research, specifically by showcasing how multimodal emotions data can illustrate the different emotional landscapes of learners relative to solely relying on self-reports. Specifically, by just examining the emotions self-report data, the reasonable conclusion to make is that our harassment intervention tends to elicit positive-activation emotions. This conclusion is quite different from the conclusion I made when I looked at three different sources of emotions data.

I believe the discrepancies in the findings are due to the characteristics of self-report measures. Specifically, self-report measures in my study relied on recall of emotional experiences, as opposed to reporting ongoing experiences. Memory recall may be prone to misattribution or misremembering what discrete emotions were experienced, as well as the

intensity of those emotions, and even the sources of the emotions. In addition, it would be up to the learners to interpret the prompts of the self-report questionnaires to some extent. When the prompt asks how the learner felt towards the simulation activity, the learner will decide, subjectively, which aspect of the simulation to focus on as they recall emotion episodes. For example, in our anti-harassment training, a learner could have focused on the following aspects of the simulations to the exclusion of others: the mannikin's silicone skin being more difficult to work with than anticipated; the senior resident being unable to provide any procedural help despite the senior resident helping being the norm in real situations; certain harassment intervention strategies seemingly not being effective; and so on. Focusing on just the end results, many may have said that the simulation was enjoyable, while during most of the simulation it was filled with confusion and anxiety as shown in my content analysis. While probing emotions via speech content analysis is a promising way to address these shortcomings of self-report measures, self-report questionnaires that focus on specific aspects of the learning experience (e.g., ask emotions regarding *just* the use of the mannikin) may be another solution.

The discrepancies in the findings also come from the fact that the addition of EDA and content analysis simply offer more dimensions of emotions. Further, EDA data is not retrospective, but rather preserves the temporal nature of emotions. The content analysis from the simulation debriefing provided insights to the sources of emotions, and other diverse sets of emotions not captured by self-report measures. Combining all these data channels allowed me to draw a more nuanced picture of the anti-harassment training experience.

5.1.4 Emotions, Learning, and Simulation Performance

The first anti-harassment study we conducted (chapter 3) revealed limited relationships between learning and emotions. The findings were still congruent with the CVT, and prior reviews that have outlined the tendency of emotions' impact on learning.[10,28] Specifically, enjoyment of the video instruction was positively linked to knowledge levels, while boredom and anxiety was negatively linked to knowledge levels. I believe the engaging animations and illustrations, and carefully crafted presentations of the videos would have fostered enjoyment in watching the videos. Further, residents may have found examples of harassment types and harassment scenarios relatable. Studies external to medical education have found that relatable content can elicit positive emotions.[126,127] Ergo, it stands to reason those who connected with the content experienced more positive emotions, which could have fostered enhanced cognitive engagement and motivation.[54]

However, relatability does not guarantee positive emotions,[126] and in fact, emotional engagement to an emotionally charged or sensitive topic such as harassment may elicit negative emotions, such as anxiety.[51,111] These emotions can be helpful in learning, but requires them to be resolved during learning, and not hinder cognitive processes needed for learning.[28] I posit that for some of the learners, the relatable content to harassment was *too* relatable, evoking specific past memories or knowledge regarding harassment experiences. Recalling harassment related experiences and feeling anxiety would be a distraction to learning, and therefore it makes sense to see anxiety linked to lower knowledge levels.

Finally, some residents may not have valued anti-harassment education, or did not appreciate the difficulty level of the learning materials. Medical education has traditionally

focused more on technical skills acquisition,[128] and some residents may subscribe to this tradition and find training related to non-technical skills such as managing harassment scenarios simply less valuable, therefore boring.[54] Further, I believe that our video instructions' learning content is easy relative to what medical residents would have had to learn throughout their academic career—the lack of challenge may also contribute to the feeling of boredom. As boredom relates to disengagement, it makes sense to observe the negative link to knowledge levels.[54]

The second anti-harassment study we conducted (chapter 4) showed that emotions may play a role in how residents choose to intervene during the harassment simulation. To clarify, I did not find any significant difference in terms of simulation performance scores (assessed via checklist) between three distinct emotional profile groups identified via cluster analysis. However, I did find a difference in the specific intervention response the residents showed depending on what kind of emotional profile they had. Those who were relatively less aroused as shown by their EDA chose to directly intervene. Those who showed higher levels of positive emotions, but showed a high degree of responses from the harassment situation did not directly confront the harasser. I believe this finding not only underscores how different emotional profiles can lead to different simulation performances, but also highlights the natural next step in studying emotions: emotion regulation.[53,129] This is especially so as the residents that seemed relatively calm to their peers, still reported anxiety and other negative emotions during the simulation debriefing. I believe these residents may have certain emotional regulation strategies that allowed them to have a different emotional profile relative to their peers. Specifically, as suppressing emotion expression (a type of response modulation) would not lead to lower physiological activities,[20] other strategies such as attentional deployment (i.e., shift attention to a

different object focus, therefore eliciting different emotions) or cognitive change (i.e., re-appraise the value and control of the situation and therefore elicit different emotions) may be more appropriate for explaining the results.

It is worth clarifying that the CVT's propositions and associated findings concerning how positive-activating emotions can often lead to desirable performance and learning outcomes is not a guarantee of positive emotions always leading to a good educational outcome. The context and type of emotional responses, as well as ones ability to regulate these emotions, are crucial in determining their impact on learning and performance. In our study, residents experiencing confusion and anxiety along with positive emotions showed that the source and nature of the emotions are important factors. Additionally, learners with a calm emotional profile still experienced emotions, but may have effectively regulated them to focus on cognitive processes beneficial for learning. Accordingly, both the presence of certain emotions and the strategies used to manage them are essential for achieving optimal educational outcomes.

5.2 Contributions

While there are is similar knowledge synthesis work that has documented the emotional experience of learners in medical education[74] or general SBE in healthcare education,[73] my dissertation is the first to systematically investigate how emotions and related constructs are conceptualized and measured in SBE research, and provide a critical view of what the current gaps are regarding the lack of theory-driven research. I believe this dissertation contributes to the slow, but firm turn towards deeper integration of relevant theories in emotions research in medical education and SBE research.

This dissertation is also, to my knowledge, the first to map out the emotions of medical residents during their anti-harassment training. Further, I have not only examined the more typical video-based instructions, and associated emotions, but also answered the call for SBE in anti-harassment education[108], and have provided rich descriptions of medical residents' emotional experience. I have also shown how instructional designs will elicit different sets of emotions, underscoring the importance of medical educators being aware of their learners' affect when designing interventions.

This endeavor highlights several methodological contributions. First, it shows the merit of multimodal analysis in SBE, something I have identified to be rare.[130] Relatedly, my research demonstrated how cluster analysis can be used with not just unimodal data, but also multimodal data to provide additional insights into identifying emotional profiles of learners. Further, my research has shown how multi-method approaches that integrate both quantitative and qualitative data can provide further context in emotions research, particularly towards object foci.

I lastly highlight how my dissertation shows the importance of looking at varied sets of emotions and object foci. It may seem at first glance, reading literature on the CVT and related empirical and review studies, that educators should primarily focus on eliciting positive emotions from learning environments. However, given the nature of certain learner characteristics (i.e., medical residents and their daily tasks), the emotionally charged subject domain (i.e., harassment), and the complex nature of emotions and their ties to various psychological processes, it makes sense not to always fixate on asking whether the learning experience was positive. Research should also consider various sources and aspects of emotions and ask how they tie into learners' behaviour and performance.

5.3 Limitations and Future Directions

The initial knowledge synthesis study my dissertation set out to include aimed to survey the entire medical education literature, instead of focusing on just SBE. While this is one limitation, as I have stated in Chapter 2, this may also serve as a strength of my scoping review. First, the initial aim of taking on the entire medical education literature was formulated without concrete knowledge of just how vast the literature was at the time. As our expert librarian pointed out, and as aligned with Arksey and O'Malley's framework[131] for conducting scoping reviews, such an oversized scope would have diluted the results of our review, and may have made it more difficult to extract meaningful conclusions or to identify precise research gaps. Second, also aligned with the framework, is the iterative process of conducting such a knowledge synthesis, where research questions and scopes are adjusted as researchers acquire a deeper understanding of the literature.[131]

Nonetheless surveying how emotions are conceptualized in other subfields of medical education, or other related healthcare fields can serve as a future direction. While my study focused on not only medical students, but also medical interns, residents, and fellows across all spectrums of medicine, future work could focus on a subset of the population, or subfield within medicine or surgery. Further, while my study was limited to published empirical articles, future work could embrace the gray literature to capture more nuanced view of how emotions are conceptualized by medical educators.

Sample size can be a limitation for some of the analyses in my empirical studies. However, as emphasized in Chapter 4, I believe our recruitment efforts were successful, and even analyses with smaller samples met the typical size found in health professions education

research.[132] Future directions include securing a larger sample size through potentially conducting a multicenter or multi-year study, or recruiting related populations that may have more participants available (e.g., nursing students). Such research should provide insights towards not just the generalizability of my findings, but can also enhance power of analyses, or allow analyses requiring larger samples.

Partially related to sample size, future studies can expand on the limitations I had in terms of follow-up. Specifically, my follow-up was not feasible due to only 18 participants being available from the overall 51 participants we analyzed during the intervention. The field needs more studies that examine how knowledge and performance is (not) retained overtime.[133–135] Aside from securing more participants for follow-up investigations, future anti-harassment education research could test the potential shift of knowledge and simulation performance relating to harassment interventions. As deliberate practice is a core part of SBE and achieving expertise,[134,136] the need for repeated anti-harassment education also serves as an educational need for medical trainees.

Further, future studies could examine the potential difference in emotions between simulations and real-life situations. Specifically, simulations may not fully capture the intensity and complexity of emotions experienced in real-life situations, where stakes and consequences are higher. Participants might perceive less risk in a simulated setting, leading to behaviors that differ from those in real-world contexts. For example, intervening in simulated harassment may feel less intimidating than confronting real harassment, affecting the authenticity of the emotional responses and actions taken. Additionally, the knowledge that a scenario is simulated could influence participants' engagement and emotional investment. This context might result in

lower levels of stress, empathy, and other psychological and behavioral responses compared to real-life experiences. Furthermore, different psychological processes might be at play in simulated versus real environments, potentially leading to divergent outcomes. Thus, while simulations offer valuable insights, it is crucial to consider these limitations and strive for methods that enhance the ecological validity of such research. Future studies should aim to bridge the gap between simulated and real-life investigations to ensure findings are more generalizable. One such method is recruiting professional actors to serve as the senior resident and medical student.

Though not a limitation, we have opted to focus on self-report survey, EDA, and simulation debriefing speech content data. Future studies could examine different data channels, such as heartrate for inferring valence[67] or voice features for inferring activation[137] (e.g., pitch and amplitude). Behavioural analysis from videos can also serve as a data source for emotions, as Törmänen and others (2022) have shown by conducting valence coding of learners' interactions.[138] While other data channels such as brain imaging or facial expressions remain a challenge especially in SBE environments, future technological advancements and a possible reduction in masking frequency for COVID-19 may allow for feasible implementations of such measures.

As I have pointed out through my chapters, the next natural step in emotions research after examining generated emotions is emotion regulation. Therefore, knowledge synthesis examining how emotion regulations are conceptualized by residents therefore remains an important future direction. Further, future research can shift the focus from instructional designs to individual differences in trainees, specifically their emotion regulation processes, and their

roles in learning and performance. Other factors such as prior experience with harassment, and demographic variables including ethnicity and gender should also be included in future research. Finally, educational interventions can include evidence-based content on how to form effective emotion regulation strategies for better performance in challenging situations such as intervening harassment.

Chapter 6: Conclusion

My dissertation has focused on how emotions can be conceptualized and studied in the context of anti-harassment education for medical trainees. Emotions are complex, and require adequate tools, such as theoretical frameworks and evidence-based measurements to properly study. This complex nature stems from its multicomponent characteristics, and its relation to other related but distinct constructs and processes. Simulation training environments are generally filled with fear and anxiety but have ample room for other emotions to be captured. These emotions can change based on how the simulations are designed, but also when the simulation is provided in relation to other learning activities. Emotional profiles can be made not just from self-report surveys, but from other multiple channels to provide rich, nuanced summaries of learners' emotional experiences. Emotions can have an impact on how trainees behave and perform in simulations.

The development and testing of the instructional videos and simulation activity I was the lead developer of is a firm step towards addressing the urgent issue of harassment in health professions training. While multimodal approaches for emotion identification is the future of SBE research, emotion regulation represents another important avenue of research. Effectiveness inter and intrapersonal management of emotions will play a crucial role in addressing harassment and anti-harassment education.

Chapter 7: References

Automatic citation updates are disabled. To see the bibliography, click Refresh in the Zotero tab.