Presimulation Instruction of Technical Skills to Enhance Simulation-Based Education of Non-Technical Skills: A Convergent Mixed Method Study

Nicholas Robillard, MD, FRCPC, MHPE, Christian Vincelette, RN, PhD, Arnaud Robitaille, MD, Terry Varshney, MD CM, FRCPC, Meghan Andrews, MD, FRCPC, Richard Waldolf, MD, MMEd, Maureen Thivierge-Southidara, MD, MA, Matthew Lineberry, PhD, Rachel Yudkowsky, MD, MHPE, Vicki LeBlanc, PhD, and Ara Tekian, PhD, MHPE

Introduction: Teamwork practice through simulation-based education (SBE) is effective, but optimal instructional design remains uncertain. Preinstruction targeting technical skills (TS) and non-TS (NTS) has shown promise in supporting their respective acquisition through simulation. However, evidence remains limited on whether preteaching TS can enhance NTS acquisition, such as crisis resource management (CRM). This study aims to assess the impact of presimulation instruction of TS on the acquisition of CRM during SBE.

Methods: We used a convergent mixed-method design, combining a quantitative post-test–only control group design with a complementary qualitative component. The intervention group had access to preinstruction of the TS necessary for managing an acutely ill patient, whereas the control group was exposed to a sham video. The main outcome was CRM skills acquisition, as measured by the Ottawa Global Rating Scale (OGRS) after 2 SBE sessions held 3 months apart (T_0 and T_3). Secondary objectives were the intervention's effect on anxiety, cognitive load, and participants' perceptions of the intervention. Quantitative outcomes were assessed with a repeated-measures general linear model. Semistructured interviews were conducted after each simulation, and thematic analyses were performed.

Results: Sixty-four postgraduate year 1 (PGY1) residents were randomized into intervention and control groups. Participants who received preinstruction of TS in addition to SBE of NTS achieved significantly higher overall OGRS scores than those who received SBE of NTS alone. There were no between-group differences in anxiety measures. Qualitative analysis revealed high variability in the intervention's impact on participants, some revealing lower cognitive load, whereas others heightened levels of performance anxiety.

Conclusions: In PGY1 residents, preinstruction of TS may reduce cognitive load during simulation training and enhance CRM skill acquisition at 3 months, although not via anxiety reduction. Responsiveness to the study intervention was variable and highlights the need for further research on the impact of instructional design adaptations on different learner subsets.

Key Words: Crisis resource management, nontechnical skills, technical skills, stress response, cognitive load, presimulation instruction, preteaching

(*Sim Healthcare* 2026;21:1–10)

From the Department of Medicine, Critical Care Service, Centre Hospitalier de l'Université de Montréal (CHUM); Simulation Centre, Académie CHUM, Centre de Recherche en Pédagogie de la Santé, Faculté de Médecine, Université de Montréal; Centre de Recherche du Centre Hospitalier de l'Université de Montréal, Québec (N.R.); Faculty of Nursing, Université de Montréal, Montréal; Centre de Recherche du Centre Hospitalier de l'Université de Montréal, Québec (C.V.); Department of Anesthesiology, Centre Hospitalier de l'Université de Montréal, Québec (A.R.); Department of Pediatrics, Division of Emergency, University of Ottawa, Ontario (T.V.); Department of Anesthesiology and Pain Medicine, Montfort Hospital, University of Ottawa, Ontario (M.A.); Department of Family Medicine and Department of Innovation in Medical Education, University of Ottawa, Ontario; Institut du Savoir Montfort, Ottawa, Ontario (R.W.); Department of Medicine, McGill University, Montréal, Québec, Canada (M.T.-S.); Zamierowski Institute for Experiential Learning, University of Kansas Medical Center and Health System, Kansas City, KS (M.L.); Department of Medical Education, University of Illinois at Chicago, IL (R.Y.); Department of Innovation in Medical Education, University of Ottawa, Ontario, Canada (V.L.); and Department of Medical Education; University of Illinois at Chicago, IL (A.T.).

Reprints: Nicholas Robillard, MD, FRCPC, MHPE, Centre Hospitalier de l'Université de Montréal, 1000 rue Saint-Denis, Montréal, QC, Canada H2X 0C1 (e-mail: n.robillard@umontreal.ca).

ORCID: Nicholas Robillard, 0000-0001-9435-5889
Richard Waldolf, 0009-0008-8472-3164
Maureen Thivierge-Southidara, 0000-0001-9882-3571
Matthew Lineberry, 0000-0002-0177-5305
Rachel Yudkowsky, 0000-0002-2145-7582
Vicki LeBlanc, 0000-0001-5407-6420
Ara Tekian, 0000-0002-9252-1588

Conflicts of interest: The authors declare no conflict of interest.

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's Web site (www.simulationinhealthcare.com).

ISSN: 1559-2332
DOI: 10.1097/SIH.0000000000000893

The report *To Err Is Human: Building a Safer Health System*[1] has led to significant efforts to mitigate human errors, such as the implementation of nontechnical skills (NTS)[2,3] training. Simulation-based education (SBE) literature about NTS, such as crisis resource management (CRM), has since greatly expanded.[4] SBE has become a widely adopted resource for acquiring these skills.[3–7] Despite the effectiveness of SBE,[8,9] the optimal instructional design for NTS training remains uncertain.

Meta-analyses assessing different SBE instructional designs provide guidance[10,11] but are not specific to NTS instruction. Some evidence suggests that NTS acquisition requires preexisting technical skills (TS) capacities,[12,13] which favors teaching TS before a simulation activity to ensure that learners possess the abilities to engage in NTS instruction.

Two conceptual frameworks support the offloading of taskwork-related learning to the presimulation phase, thereby reserving in-person time to the development of NTS (Fig. 1). First, this could reduce the burden imposed on learners' cognitive processes by separating the instructional content. According to cognitive load theory, working memory has a finite capacity.[14,15] Teaching TS and NTS simultaneously may overload learners' working memory and impede learning. Temporally separating instructional content could free up learners' cognitive capacities and align them to the intended NTS learning objectives.[16]

Second, preinstruction of TS could favorably modulate learners' cognitive load by reducing their anxiety. Multiple studies link simulation instruction to heightened levels of anxiety.[17–21] Such negative emotions are associated with reduced cognitive resources available for learning and could hinder working memory.[22,23] Pekrun et al[24] found that anxiety arising from instructional activities is associated with task-irrelevant thinking, resulting in cognitive resources being focused away from learning. Recommendations to develop learning activities that foster positive emotions have been published.[25] Three strategies can be instantiated through the model we propose. First, offloading medical knowledge may align the simulation's range of difficulty to the learner's capacity. Second, providing learners with preparatory information can enhance their perceived control over the learning experience, resulting in higher anticipated chances of success and limiting emergence of anticipatory anxiety.[26–28] Third, a shared understanding of the upcoming learning experience may create an environment that is less performance focused,[29] potentially leading learners to engage in a more mastery-oriented practice[30] rather than demonstrating competency through peer comparisons.[31]

In this study, we assessed whether providing junior medical residents with preparatory instruction on the medical content of a simulation case would allow them to focus more on CRM principles, resulting in increased acquisition of these skills after 2 technology-enhanced simulation sessions held 3 months apart aimed at acquiring CRM competencies.

METHODS

Research Design

We conducted this study at a single institution using a convergent mixed-method design with an embedded experimental model.[32] Quantitative and qualitative phases of the study were conducted simultaneously and merged during the interpretation phase. This trial was double-blinded and used random assignment to experimental groups.

Institutional review board approval was obtained from the Centre Hospitalier de l'Université de Montréal and the University of Illinois Chicago College of Medicine.

Participant Selection

The inclusion criteria required participants to be postgraduate year 1 (PGY1) residents, on an intensive care unit rotation during recruitment, and with an anticipated ability to complete both simulations. PGY1 learners were purposefully[33,34] chosen because they were deemed most likely to benefit from TS instruction. Recruitment was conducted by a research assistant through email and in-person invitations. Exclusion criteria included endocrine disorders, corticosteroid therapy, use of drugs interfering with stress response, pregnancy, and prior knowledge of simulation scenarios.

Participants were randomized into control and intervention groups in a 1:1 fashion. Block randomization was initially planned to alleviate potential group imbalances. However, a clerical error resulted in nonblocked randomization. The procedure was reviewed by a biostatistician and its integrity confirmed.

Trial Procedures

Participants completed 3 procedures: (1) a TS instruction web-based module a week before the first simulation session, (2) a first SBE session ($T = 0$ months, T_0), which comprised a dedicated CRM instruction component followed by an initial skills assessment, and (3) a second SBE session 3 months later ($T = 3$ months, T_3) that only included a CRM skills assessment (Fig. 2).

The web-based module was completed by the intervention and control groups a week before the first simulation. The web platform had 2 separate interfaces: The intervention group had access to the trial intervention, and the control group had access only to a sham simulation laboratory orientation video. The intervention consisted of three 10- to 15-minute narrated video presentations covering the medical knowledge required to complete the taskwork of the upcoming simulated acute care situation (eg, recognizing and managing anaphylaxis, treating unstable atrial fibrillation—including cardioversion—and managing respiratory distress). A pre-/post-test confirmed that the learning material was watched by the intervention group. The test was not administered to the control group because of concerns about compromising blinding, as providing the questionnaire could inadvertently prime participants regarding the simulation's content.

Both groups then participated in a baseline SBE CRM session (T_0) that included a 40-minute video about CRM principles,[35] followed by a technology-enhanced (SimMan 3G, Laerdal) simulation exercise. The case consisted of anaphylaxis proceeding to unresponsive shock, respiratory insufficiency, and eventual unstable atrial fibrillation. The scenario

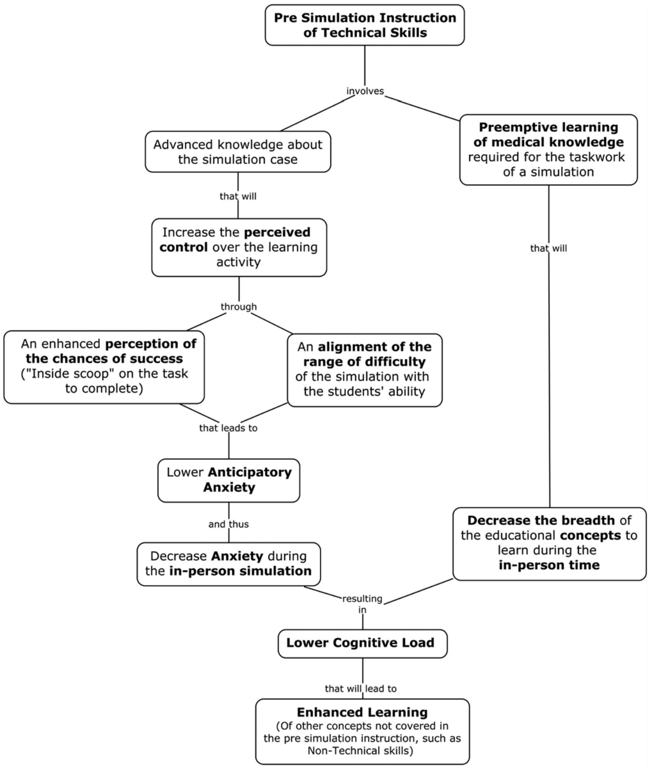

FIGURE 1. Concept map of study hypotheses.

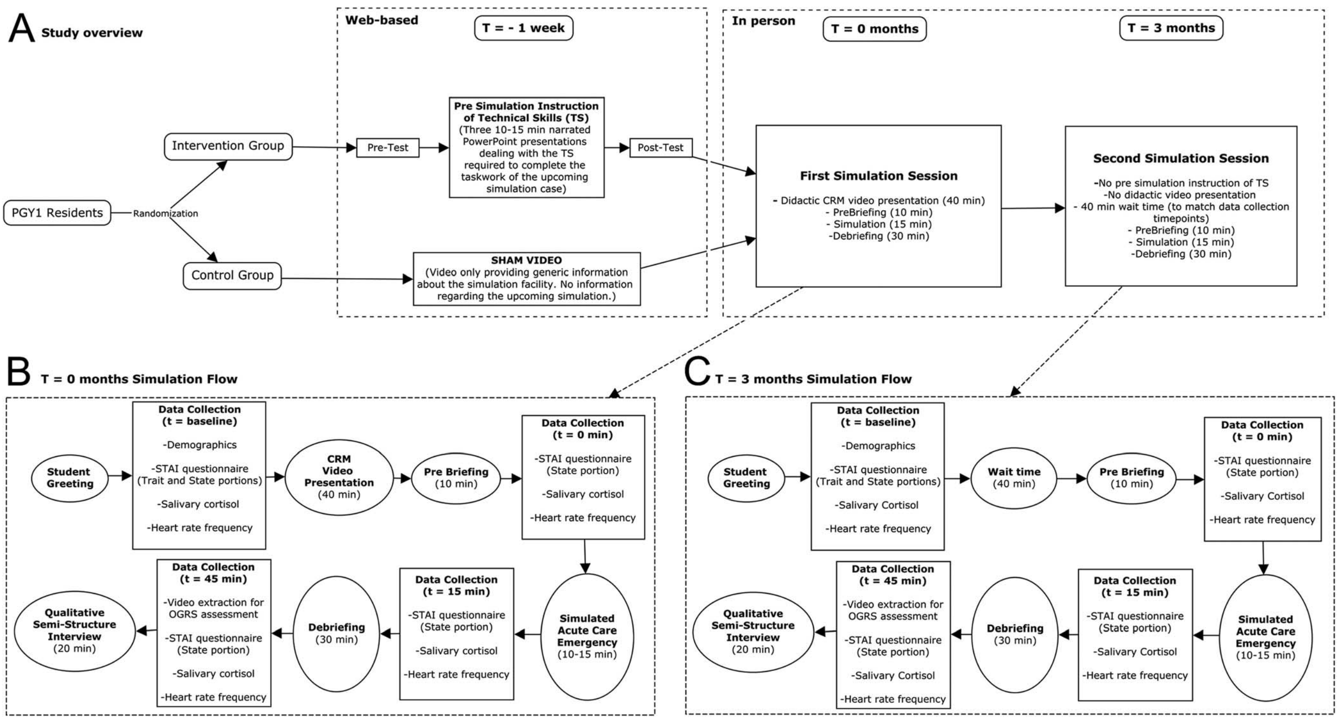

FIGURE 2. Figure 2. Study flow of study procedures. t indicates timing.

required residents to evaluate a patient on the hospital ward as if they were part of our institution's rapid response team. Both groups' video recordings from this initial simulation session were subsequently reviewed to obtain the first assessment measure of CRM skills.

The second assessment of NTS competencies came 3 months later with a different case. This session (T_3) included a simulation exercise consisting of a patient with pulmonary embolism who developed shock, respiratory insufficiency, and unstable bradyarrhythmia (see Tables, Supplemental Digital Content 1 [http://links.lww.com/SIH/B204], highlighting patient states according to time markers).

Both simulation sessions included prebriefing, simulation practice, and debriefing. Instructional components were standardized across participants. Each participant operated individually in both simulation cases, supported by 2 embedded participants who played the roles of a nurse and a respiratory therapist, completing the team. Embedded participants were trained during the trial pilot, receiving individualized feedback after each of the 12 simulations. Simulation cases were developed with specific time markers and evolved regardless of the participant's interventions (see Tables, Supplemental Digital Content 1 [http://links.lww.com/SIH/B204], highlighting patient states according to time markers). Standardized time-based interventions by the embedded participants were also incorporated to induce participants to use specific CRM principles. The simulation ended after participants reached the final patient state and transferred the case to an intensivist played by the instructor. Two simulation instructors (N.R., A.R.) oversaw the simulation sessions and consistently used the same debriefing technique.[36,37] Debriefings were standardized across experimental groups, timed, and could not exceed 30 minutes. They were strictly focused on CRM learning objectives (see Tables, Supplemental Digital Content 2 [http://links.lww.com/SIH/B205] for a debriefing guide outlining key CRM principles and their expected timing).

Trial Outcomes

The primary outcome was the difference in overall Ottawa Global Rating Scale (OGRS) score between intervention and control across the 3-month study period.[35] Secondary quantitative outcomes were measured at both the T_0 and T_3 simulation sessions. They included (1) differences in mean scores on OGRS subcategories; and (2) differences in stress response measured by salivary cortisol,[38,39] State-Trait Anxiety Inventory (STAI)[40] scores, and heart rate (HR) measurements.

Two exploratory qualitative outcomes were also integrated into the trial. The first involved a deductive assessment of the interventions' impact—or lack thereof in the control group—on learners' cognitive load and self-reported anxiety; the second inductively assessed unaccounted mechanisms for the effect of this new instructional design.

Instrumentation and Data Collection Procedures

CRM skills were assessed using the OGRS scale,[35] which includes 6 domains: overall performance, leadership, problem-solving, communication, resource utilization, and situational awareness. Each is rated on a 7-point Likert scale. The validity evidence supporting the use of the OGRS consists of content validity, response process, internal consistency and relationship to other variables, in accordance with Messick's modern framework of validity evidence for assessment purposes.[35,41] All simulations were video recorded. Recordings were extracted following each session to enable subsequent assessment of participants' CRM skills by independent raters. Both simulations (T_0 and T_3) were reviewed by 3 raters who underwent training and were blinded to group allocation (see text document, Supplemental Digital Content 3 [http://links.lww.com/SIH/B206], describing the rater training process). Assessing OGRS scores during the baseline T_0 session was considered important for characterizing learners' skill trajectories, particularly in distinguishing between immediate and delayed improvements. A calibration session was held halfway through the ratings to ensure consistency and reliability in the raters' assessments.

Salivary cortisol, a minimally invasive and widely used physiological marker,[38,42] was chosen to assess stress response because of its established role in reflecting Hypothalamic-Pituitary-Adrenal axis activation.[43,44]

The STAI questionnaire measured subjective anxiety through 2 components: State (capturing momentary anxiety) and Trait (assessing baseline stress levels). Both contain 20 items rated on a 4-point Likert scale and have been extensively validated in stress research.[40]

Lastly, HR was used as an additional physiological stress metric. Studies have shown significant HR increases during simulations, correlating with anxiety levels.[45] HR was manually measured by a research assistant over a 1-minute period.

Salivary cortisol, STAI, and HR data were collected at 4 timepoints during both simulations: baseline (t = baseline), presimulation (t = 0), postsimulation (t = 15), and postdebriefing (t = 45).

A 20-minute qualitative semistructured interview was conducted after each debriefing with an interview guide for both the control and intervention groups (see text document, Supplemental Digital Content 4 [http://links.lww.com/SIH/B207], providing the interview guide) and transcribed verbatim.

All data collection procedures were standardized and conducted by a research assistant.

Sample Size and Quantitative Data Analysis

A sample size of 50 participants was necessary to detect a 10% absolute difference on the OGRS between groups with a 2-sided alpha of 0.05 and a power of 80%.[35] Anticipating a 20% attrition rate, we increased the sample size to 64 participants, 32 in each group.

The use of pooled estimates of the 3 raters was initially planned to assess between-group differences in overall OGRS scores. However, this plan was revised when Rater 1 appeared to have misunderstood rating directives during the calibration session. Sources of variance arising from the raters were assessed with a repeated-measures general linear model (GLM). If Rater 1 was identified as the source of significant variance in OGRS scores, we planned a sensitivity analysis to exclude this rater.

Descriptive statistics for participants' characteristics, as well as primary and secondary outcomes, are presented. The primary and secondary quantitative outcomes were analyzed using a repeated measures GLM, incorporating 2 within-subject factors: simulation timing (T_0 vs. T_3) and rater (Rater 1, Rater 2, and Rater 3). The between-subject factor was study group allocation, with 2 levels: intervention and control. For other continuous variables, we used Student t tests to assess between-group differences. Discrete variables were analyzed with chi-square or Fisher's exact tests. Our analysis treated Likert scale data as continuous interval data.[46,47] The level of significance was set at 0.05. Analyses were conducted with SPSS software version 29 (IBM Corp, Armonk, NY).

Qualitative Data Analysis

This study was anchored in pragmatic epistemology with a qualitative description methodology.[48] A thematic analysis method was applied to provide inductive insights to validate the trial's conceptual framework.[49,50] Data from both the intervention and control groups were included to enhance the credibility of our findings by offering contrasting perspectives that contribute to a more comprehensive understanding of the phenomenon under investigation. The Braun and Clarke's 6-step method was used to analyze the data.

Two investigators (N.R., C.V.) first familiarized themselves with the data, and then independently coded 15% of the data corpus. After they agreed on the coding structure, the principal investigator (N.R.) coded the entire corpus. Data were managed with NVivo 12 software. Data queries were conducted after every 25% increment of completed data corpus analysis. These queries, a feature of the NVivo data management software, enable exploration of coded data to uncover patterns, relationships, or distributions that may not be immediately apparent (eg, between-group differences in the cognitive load theme). They provided a broader analytical perspective and informed iterative refinement of the thematic framework. Memos were generated throughout the process to keep track of these queries and other important markers of analysis.

Once initial codes were generated for the entire data corpus, codes and categories were iteratively reviewed and refined, allowing major themes to emerge from the data. We then conducted axial coding to establish links between themes to provide a rationale for the study findings. The supporting evidence for each theme was reviewed by both investigators (N. R., C.V.) to assess for coherence.

RESULTS

The study included 64 participants. Two were lost to follow-up, leaving 31 in each group (Fig. 3).

Participant Characteristics

Participant characteristics were similar in both groups (Table 1). Slight differences were noted in specialty types, with more internal medicine and fewer anesthesia residents in the intervention group.

Quantitative Results

Primary Outcome: Overall OGRS Scores

The repeated-measures GLM revealed a statistically significant within-subject difference related to the raters, $F(1.61, 93.43) = 4.93$, $P = 0.014$. Pairwise comparisons showed that Rater 1, who misunderstood rating directives, was a major source of this variance.

We therefore conducted the same analysis, excluding Rater 1 assessments. A significant between-subject difference favored the intervention group, $F(1, 58) = 5.46$, $P = 0.023$, with an estimated marginal mean difference of 0.492 ($P = 0.023$)

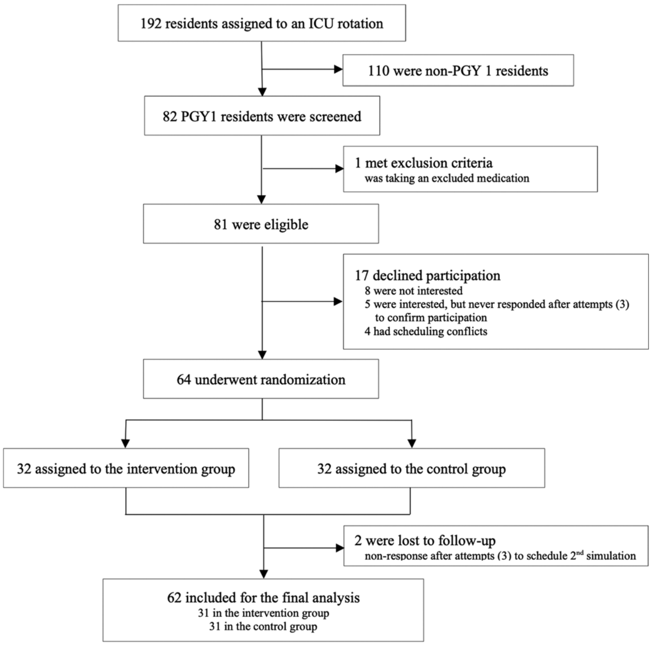

FIGURE 3. Enrollment, randomization, and follow-up.

(Fig. 4). There were no discernible within-subjects main effects for the remaining raters, $F(1, 58) = 0.654$, $P = 0.422$, or simulation timing, $F(1, 58) = 1.82$, $P = 0.185$, and no interactions effects, notably no interactions between simulation timing and group allocation, $F(1,58) = 0.556$, $P = 0.459$. Table 2 provides aggregate sample means of the overall scores.

Secondary Outcomes

OGRS Subscale Scores

The intervention group had significantly higher ratings on leadership, $F(1, 58) = 6.77$, $P = 0.012$; resource utilization, $F(1, 58) = 4.29$, $P = 0.043$; and communication subscales, $F(1, 58) = 4.35$, $P = 0.041$. There were no significant differences in problem solving, $F(1, 58) = 2.56$, $P = 0.115$, and situational awareness scores, $F(1, 58) = 2.36$, $P = 0.130$ (see text document, Supplemental Digital Content 5 [http://links.lww.com/SIH/B208], showing the subscale means by raters).

Stress Response Measurements

There were no between-group differences in stress response for the simulations (Fig. 5). Salivary cortisol measures were similar: Simulation T_0, $F(1,58) = 0.013$, $P = 0.908$; Simulation T_3: $F(1, 58) = 0.287$, $P = 0.594$. A within-subject effect was observed for the timing of measurements that suggested a bimodal peak of cortisol for both simulations. HR measurements showed similar results. There were no significant differences between groups for STAI state scores, $F(1, 57) = 0.234$, $P = 0.630$. Furthermore, no significant interactions were observed between study groups and the timing of measurements for both simulations: Simulation T_0, $F(3,171) = 2.564$, $P = 0.056$; Simulation T_3, $F(3,162) = 1.36$, $P = 0.257$.

Qualitative Results

A total of 126 interviews were analyzed. Three major themes were identified: (1) NTS education being contingent on TS

TABLE 1. Characteristics of Study Participants

Characteristic	Control (n = 32)	Intervention (n = 32)	*P*
Age, mean ± SD	25.19 ± 1.79	25.59 ± 2.03	0.399†
Female sex, n (%)	20 (62.5)	17 (53.1)	0.448‡
Residency program, n (%)			0.139#
Medical	27 (84.4)	29 (90.6)	
Anesthesia	6 (18.8)	2 (6.3)	
Internal medicine	20 (62.5)	25 (78.1)	
Neurology	1 (3.1)	2 (6.3)	
Surgery	5 (15.6)	3 (9.3)	
ENT	2 (6.3)	0 (0)	
Digestive	1 (3.1)	0 (0)	
Orthopedic	1 (3.1)	0 (0)	
Plastic surgery	0 (0)	1 (3.1)	
Urology	0 (0)	2 (6.3)	
Vascular	1 (3.1)	0 (0)	
Degree, n (%)			1.000#
Undergraduate (pre-Bachelor's)	21 (65.6)	20 (62.5)	
Bachelor's	9 (28.1)	9 (28.1)	
Master's	1 (3.1)	2 (6.3)	
Doctorate	1 (3.1)	1 (3.1)	
Prior ICU experience, n (%)	0 (0)	0 (0)	n/a
Number of simulation experiences in last year, n (%)			0.371#
None	2 (6.3)	1 (3.1)	
1 to 5	24 (75)	29 (90.1)	
6 to 10	5 (15.6)	2 (6.3)	
>10	1 (3.1)	0 (0)	
Types of simulation, n (%)			
Theater based	23 (71.9)	27 (84.4)	0.226‡
Part task trainer	22 (68.8)	22 (68.8)	1.000‡
With actors	24 (75)	21 (65.6)	0.412‡
Hybrid	13 (40.6)	11 (34.4)	0.606‡
In situ	1 (3.6)	0	1.000‡
Prior CRM instruction, n (%)	5 (15.6)	4 (12.5)	1.000#
Number of CRM courses in last year, n (%)			1.000#
None	27 (84.4)	28 (87.5)	
1 to 5	5 (15.6)	4 (12.5)	
6 to 10	0 (0)	0 (0)	
>10	0 (0)	0 (0)	
Types of CRM courses, n (%)			
Simulation			
Theater based	3 (9.4)	4 (12.5)	1.000#
With actors	3 (9.4)	2 (6.3)	0.426#
In situ	0 (0)	0 (0)	n/a
Didactic	2 (6.3)	3 (9.4)	1.000#
Presimulation psychoactive substances, n (%)	1 (3.1)	2 (6.3)	1.000#
Coffee	1 (3.1)	2 (6.3)	1.000#
Cigarette	0 (0)	0 (0)	n/a
Drugs	0 (0)	0 (0)	n/a
Energy drinks	0 (0)	0 (0)	n/a

†Student *t* test; ‡chi-square test; #Fisher exact test.
ENT indicates ear, nose, and throat; n/a, not applicable.

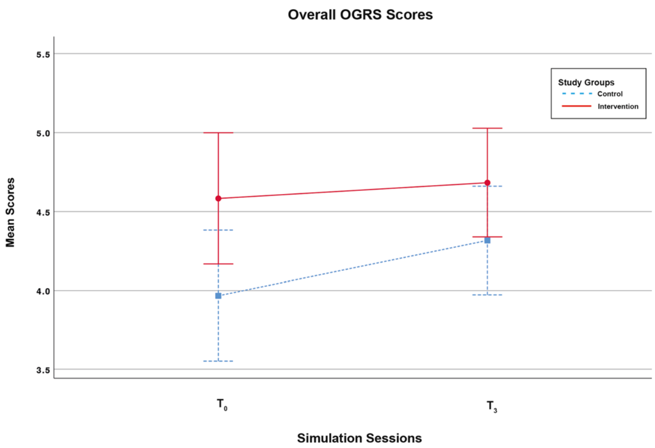

FIGURE 4. Overall OGRS scores for each simulation session, excluding Rater 1 assessment. T_0 = Initial simulation session, T_3 = 3-month simulation session. Error bars represent 95% CI.

competence, (2) asymmetrical effect of presimulation instruction of TS, and (3) pervasive intuition toward assessment and judgment. Each theme had several subthemes that offer a more comprehensive view of the overarching topic (Table 3).

NTS Education Is Contingent on TS Competence

Study participants found it easy to learn CRM but faced difficulties in applying these skills in practice.

(1) Participant 4 (control group): "(…) the knowledge [CRM principles] is hard to apply, but not to acquire because it's like in my head, but I'm not able to use it."

One reason for this difficulty is the high level of interconnectivity between different elements of the simulation cases.

(2) Participant 26 (control group): "We are so focused on all the details that our brain feels empty. I don't know if I really had the ability to think about CRM (…)."

Cognitive overload leaves little space for deliberate consideration of CRM principles.

(3) Participant 10 (control group): "[…] when the situation escalates, we get overwhelmed […] and I forget these notions [CRM] and I focus on the medical perspective."

Participants associate cognitive overload with their lack of experience. Greater cognitive resources are required by the taskwork of the case because participants have yet to integrate enough clinical patterns.

(4) Participant 33 (control group): "We are not experienced enough, the medical knowledge takes a lot of room, and we don't have any cognitive attention for crisis resource management."

Participants' TS affect their ability to practice NTS. Bridging the TS performance gap should therefore enhance their CRM learning experience. Although logical, this may not be suited to everyone.

Asymmetrical Effect of Presimulation Instruction of TS

Both groups expressed varying perceptions of the study intervention. Paradoxically, some participants praised the presimulation instruction of TS, whereas others believed it was detrimental to their learning. Positive comments were in line with the trial hypotheses of reduced cognitive load and increased control.

(5) Participant 48 (intervention group): "Uh, I felt, well, before the simulation, that is, I think I was still a little nervous. But, uh, thanks to the preparation for the activity, I felt much more, uh, confident and calm."

(6) Participant 43 (intervention group) "Hm … well, I found it [cognitive load] was lighter, uh, in the sense that I was prepared for the situation. I felt that the cognitive effort was lower."

However, for some, presimulation instruction of TS had deleterious effects, increasing anxiety and pressure to perform.

(7) Participant 20 (intervention group): "It's [presimulation TS instruction] somewhat stressful. I worry I'll forget. […] when you know, it's like "Okay so I better be good at it and succeed because I'm supposed to know it all."

Mixed feelings toward the intervention were also reported, with participants perceiving it as both an agonist and antagonist to CRM education.

(8) Participant 1 (intervention group): "It [presimulation TS instruction] had a neutral impact. While it did reduce my anxiety by helping me say, 'Ok good, I can expect this and that,' it also created the feeling I had to be good because I already knew what was going to happen."

This discrepancy in perceptions may be influenced by learners' achievement goal orientation.

(9) Participant 32 (control group): "I think that it always depends on what we value important. […] if we can put our ego aside, maybe we can have a bit less stress with simulations. […] I wasn't particularly nervous because I think I had more to gain than I had to lose."

This implies that SBE may have a cost, and this cost depends on learners' goals and may determine whether learners perceive an educational intervention as an opportunity or a threat.

Pervasive Intuition of Assessment and Judgment

Our data suggest that long-standing exposure to a culture of assessment may generate the impression that performance is at the forefront of training.

TABLE 2. Sample Means of the OGRS Overall and Subscale Scores

		Control	Intervention
	Simulation Session	Sample Mean (SE)	Sample Mean (SE)
OGRS			
Overall	T_0	3.96 (0.208)	4.58 (0.208)
	T_3	4.31 (0.172)	4.68 (0.172)
OGRS subscales			
Leadership	T_0	3.87 (0.184)	4.50 (0.184)
	T_3	3.67 (0.155)	4.70 (0.155)
Problem solving	T_0	4.35 (0.200)	4.81 (0.200)
	T_3	4.65 (0.182)	4.80 (0.182)
Situational awareness	T_0	3.88 (0.202)	4.50 (0.202)
	T_3	4.28 (0.168)	4.30 (0.168)
Resource utilization	T_0	4.06 (0.186)	4.56 (0.186)
	T_3	4.35 (0.172)	4.70 (0.172)
Communication	T_0	4.73 (0.165)	5.18 (0.165)
	T_3	4.85 (0.172)	5.16 (0.172)

SE indicates standard error; T_0, time = 0 months; T_3, time = 3 months.

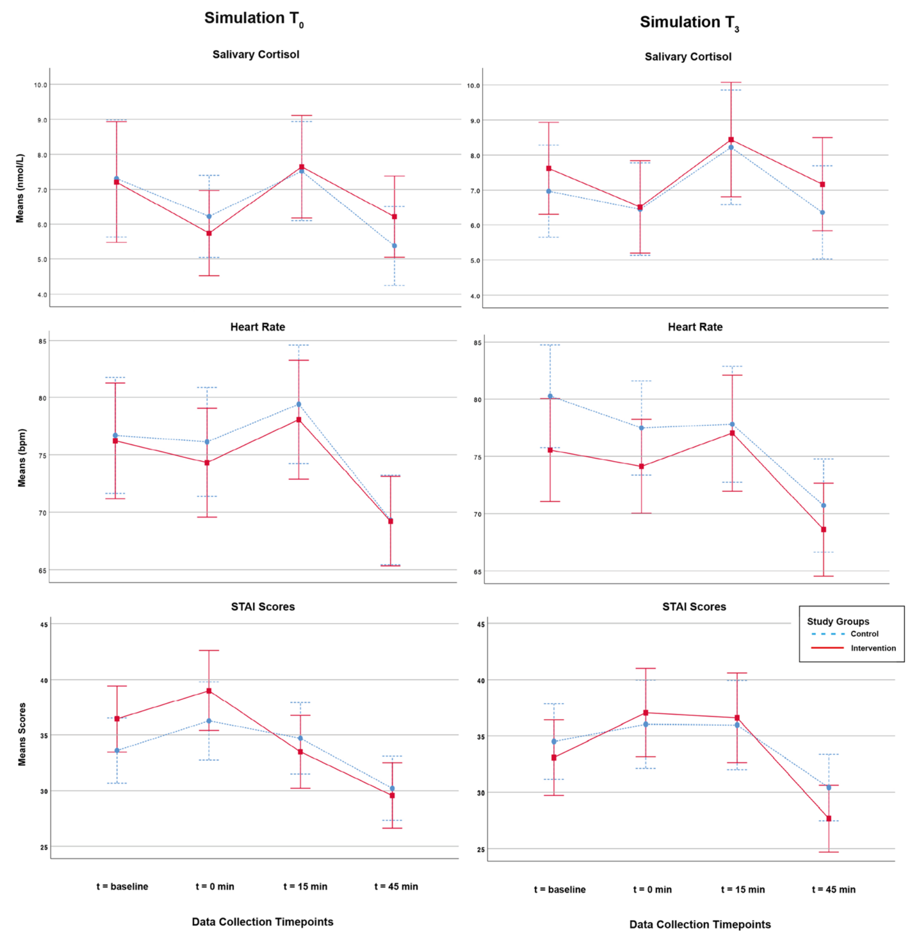

FIGURE 5. Stress response assessed at 4 timepoints. t = baseline collected when residents were greeted at the simulation laboratory; t = 0 minute and t = 15 minutes immediately before and after the simulation activity, respectively; and t = 45 minutes after the debriefing session. Error bars represent 95% CI.

(10) Participant 44 (control group): "We constantly feel our performance being assessed. It's part of who we are, what we live every day, what we are taught, and how we feel from clerkship up to now."

Some learners have an unshakeable feeling of impending judgment. Although they eventually rationalize that there are no dubious intentions, their initial instinct is to prepare for a threat.

(11) Participant 62 (intervention group): "You tell yourself, 'Of course they're going to judge me.' But in reality, they [facilitators] couldn't care less."

This intuition stems from a perceived threat, leading individuals to take the path of least resistance by demonstrating proficiency in more familiar skills. As a result, there may be an unconscious emphasis on TS, which can overshadow CRM learning.

(12) Participant 32 (control group) "(…) most of my attention is still directed more toward the medical side, because maybe I'm not yet comfortable enough or quick enough in my decisions to really be able to maximize my teamwork skills."

Students' focus on TS, combined with the ingrained intuition that assessment is omnipresent, may distort how they react to the study intervention. For some individuals, it increases the standards they believe they must uphold.

(13) Participant 2 (intervention group): "I knew that I was going to have to manage an atrial fibrillation and an anaphylaxis. I had the impression that the expectations were higher because I had prepared myself prior to the simulation."

Ultimately, this focus affects how they respond to the study intervention and may result in the paradoxical effect mentioned earlier.

DISCUSSION

We found that combining the preinstruction of TS with SBE to aid CRM instruction resulted in increased overall

TABLE 3. Thematic Analysis Results: Main and Subthemes

NTS Education Is Contingent on TS Competence	Asymmetrical Effect of Presimulation Instruction of TS	Pervasive Intuition Toward Assessment and Judgment
CRM is easy to learn but hard to apply	Unexpected paradoxical effects of presimulation teaching of TS	Culture of assessment and standards to live up to
High item interaction increases cognitive load during simulation	Achievement goal orientation as a perceptual modulator	Learner's instinct and cognitive biases
Clinical experience modulates cognitive load		Subconscious emphasis on TS

CRM proficiency. In particular, it had a positive impact on leadership, resource utilization, and communication skills. Although this increase in performance is modest, our finding of a statistically significant between-subject difference in overall OGRS scores favoring the intervention group, without any within-subject or interaction effects, indicates that the intervention group consistently outperformed the control group in CRM skill levels.

The intervention appears to have provided participants with a "head start" in developing CRM skills, an advantage that was sustained over time. However, our trial was not designed to assess participants' baseline skill levels, making it difficult to definitively determine whether the intervention directly resulted in an "educational boost." That said, the alternative explanation—that participants in the intervention group were inherently better at CRM than those in the control group—seems unlikely. Such a scenario would imply selection bias, which is improbable given the randomized design of the study and the balanced baseline characteristics of participants in both the control and intervention groups.

Some researchers have expressed concerns that NTS scales may inadvertently capture elements of TS,[12] leading to construct-irrelevant variance.[51] In our study, the intervention caused an uneven distribution of TS among participants. Consequently, subscale items of the OGRS assessing both NTS and TS could have hypothetically inflated scores for the first simulation and contributed to the observed divergence between groups. However, the only OGRS subscale that incorporated concepts that potentially overlapped TS, the problem-solving subscale, did not reveal any between-group differences. This makes it unlikely that construct-irrelevant variance played a meaningful role. In addition, the lack of interaction effect between group assignment and simulation sessions further supports this conclusion, suggesting that skill improvements from the first simulation alone do not account for the observed between-group differences.

We are therefore left to consider that the preinstruction of TS intervention may indeed have provided a sustained early advantage in CRM skill integration, an effect that persisted even 3 months later. Although Figure 3 shows that the intervention group outperformed the control group in CRM competencies, the steeper increase in OGRS scores for the control group across simulations raises an important question about the durability of this advantage: Does the initial benefit of preinstruction of TS persist over time, or do the 2 groups ultimately converge in performance?

Interpreting this more pronounced improvement is challenging. One possible explanation is that this reflects a time-limited effect where participants who did not receive the initial "educational boost" from the preparatory teaching may have simply had more room for early growth, eventually stabilizing and following a parallel learning trajectory with the intervention group. Alternatively, the greater increase in CRM proficiency observed between the first and second simulations may suggest that the control group is catching up and could eventually surpass the intervention group. This would imply that the intervention may have inadvertently dampened the educational growth of its participants. In this context, it would be valuable to explore the intervention's impact on learners' adaptive skills, such as uncertainty tolerance (ie, could the familiarity with a specific scenario limit learners' ability to manage the unpredictability of a different case?).[52] Differentiating between these scenarios would require an in-depth analysis of the learning curves[53] associated with both groups, an approach that demands data beyond the scope of our current study.

Although these uncertainties highlight the need for further research into long-term sustainability of our intervention, our trial did investigate potential mechanisms contributing to the short-term benefits of offloading TS learning for the benefit of CRM instruction. To this end, we incorporated both quantitative and qualitative secondary objectives aimed at exploring the potential factors driving the observed effects. One hypothesized mechanism was a reduction in cognitive load through decreased anxiety (Fig. 5). However, we found no significant differences between groups in stress responses, as measured by salivary cortisol, the STAI, or HR. These findings suggest that anxiety reduction did not play a role in explaining the observed differences between the intervention and control groups.

Nonetheless, the data suggest that a modulation of cognitive load occurred. Our thematic analysis revealed that limited experience contributed to cognitive overload, particularly among control group participants. In contrast, many participants in the intervention group reported reduced cognitive load, which enhanced their ability to apply CRM principles. This finding supports previous studies that explained increased CRM proficiency through cognitive load reduction associated with higher TS competencies.[9] Expertise is, in fact, linked to higher automatic processing and clinical pattern recognition capacities.[15,54] Hence, instructional designs that provide preparatory procedural information to promote the formation of automated schemas may help reduce cognitive load and improve performance.[15]

Preliminary evidence from studies assessing the impact of flipped learning (FL) of TS combined with SBE for TS acquisition corroborates this claim. Although heterogeneous in their designs, these trials demonstrated increased skills acquisition and high satisfaction associated with the combination of FL and SBE.[55–64] Although our methodology differed in that we measured NTS, we similarly offloaded TS before the in-person simulation and, as a result, expected comparable levels of learner satisfaction.

However, our results revealed a distinct divergence from this pattern. We observed considerable interindividual variability in learners' perceptions toward the intervention. Several participants cited benefits such as decreased anxiety and increased control over the activity, in line with previous work.[26] Concurrently, however, a considerable number of participants reported the exact opposite with higher performance anxiety. Some appraised the activity through an evaluative lens and held a negative perception of the study intervention, even considering it detrimental at times. Although this is inconsistent with the FL literature, the negative response to the study intervention does align with the social intuitionist approach to moral judgment.[65] In this model, a situation unintentionally elicits a moral intuition that leads an individual to generate judgment of an event. For some participants, the study intervention seems to have stimulated an intuition of assessment and impending judgmental threat emanating from the simulation and led them to feel that performance expectations were higher. This feeling may have offset the hypothesized benefit of the trial by increasing cognitive load with added performance pressure. If so, this may have conceptually resulted in an underestimation of the intervention's efficacy, because the inclusion of learners less responsive to the treatment within the intervention group could have attenuated the observed effect.

The reason behind some participants' limited sensitivity to the intervention is elusive but could reside in how they self-regulate and assess how they achieve their learning goals. According to achievement goal theory, learners exhibit different behavioral patterns in the face of learning activities, depending on their orientation toward achievement—mastery or performance.[30,31] Depending on their goal orientation, learners will therefore react differently to learning activities and changes to instructional designs. Our thematic analysis suggests that participants' goal orientation may influence their appraisal of the study intervention, which may explain why some participants felt that achievement standards were higher and ultimately perceived increased performance pressure. Further studies are required to corroborate these hypotheses and clarify the impact of achievement goal orientations on SBE.

Limitations

Some limitations should be considered when interpreting our results. First, this was a single-site study, which may limit the generalizability of our findings. Second, although a block randomization was initially planned, a simple 1:1 randomization was inadvertently used because of a clerical error by the research staff. Third, some variance in the results came from discrepancies in raters' performance assessments. To address this, we excluded the ratings of a rater who misunderstood scoring directives and adopted a statistical plan that was not defined a priori, both of which can limit the strength of our claims. Fourth, the lack of preintervention data on baseline CRM skills limits our ability to draw definitive conclusions about the intervention's impact on participants' learning trajectories. Lastly, our study was not designed to assess the impact of preinstruction of TS on students from minority backgrounds.[66] However, we recognize that socioeconomic disparities may limit access to the technology required for this instructional approach, and this factor should be considered when interpreting our findings.

Implications

Although this study showed evidence supporting the effectiveness of offloading TS instruction for the acquisition of NTS during SBE, it remains unclear whether the intervention is equally beneficial for all participants. Our qualitative data suggest interindividual variability in responses, raising the question of whether subgroup differences existed in CRM performance within the intervention group. Hence, it calls for further research to investigate the impact of trainee characteristics, such as goal orientation, on learning behaviors during simulations to assess if preinstruction of TS is suitable for all participants.

CONCLUSION

In PGY1 residents, presimulation instruction of TS to support SBE of NTS resulted in higher proficiency in CRM skills. Some of our data suggest that offloading TS instruction before simulation training may be beneficial because of a reduction in participants' cognitive load. However, how participants experienced the study intervention was highly variable. In some instances, it induced a paradoxical effect and increased performance pressure. This should serve as a catalyst to explore and investigate learner characteristics, such as achievement goal orientations, that may influence responses to presimulation instruction of TS. In turn, this could help align instructional design to learner specificities.

REFERENCES

1. Institute of Medicine, Kohn LT, Corrigan JM, Donaldson MS, eds. In: *To Err Is Human: Building a Safer Health System [Internet]*. Washington, DC: The National Academies Press; 2000: [cited 2021 Dec 1]. 312 p. Available from: https://www.nap.edu/catalog/9728/to-err-is-human-building-a-safer-health-system.
2. Gaba DM. Training and nontechnical skills: the politics of terminology. *Simul Healthc* 2011; 6(1):8–10.
3. O'Connor P, O'Dea A. When I say … non-technical skills. *Med Educ* 2024;58(12): 1439–1441.
4. Buljac-Samardzic M, Doekhie KD, van Wijngaarden JDH. Interventions to improve team effectiveness within health care: a systematic review of the past decade. *Hum Resour Health* 2020;18(1):2.
5. Passiment M, Sacks H, Huang G. *Medical Simulation in Medical Education: Results of an AAMC Survey [Internet]*. Association of American Medical Colleges; 2011: Available from: https://www.aamc.org/system/files/c/2/259760-medicalsimulationinmedicaleducationanaamcsurvey.pdf. Accessed July 12, 2021.
6. Qayumi K, Pachev G, Zheng B, et al. Status of simulation in health care education: an international survey. *Adv Med Educ Pract* 2014;28(5):457–467.
7. Abildgren L, Lebahn-Hadidi M, Mogensen CB, et al. The effectiveness of improving healthcare teams' human factor skills using simulation-based training: a systematic review. *Adv Simul Lond Engl* 2022;7(1):12.
8. Boet S, Bould MD, Fung L, et al. Transfer of learning and patient outcome in simulated crisis resource management: a systematic review. *Can J Anesth*. 2014;61(6):571–582.
9. Gordon M, Darbyshire D, Baker P. Non-technical skills training to enhance patient safety: a systematic review. *Med Educ* 2012;46(11):1042–1054.
10. Cook DA, Brydges R, Hamstra SJ, et al. Comparative effectiveness of technology-enhanced simulation versus other instructional methods: a systematic review and meta-analysis. *Simul Healthc* 2012;7(5):308–320.
11. Cook DA, Hamstra SJ, Brydges R, et al. Comparative effectiveness of instructional design features in simulation-based education: systematic review and meta-analysis. *Med Teach* 2013;35(1):e867–e898.
12. Riem N, Boet S, Bould MD, Tavares W, Naik VN. Do technical skills correlate with non-technical skills in crisis resource management: a simulation study. *Br J Anaesth* 2012;109(5): 723–728.
13. Krage R, Zwaan L, Tjon Soei Len L, et al. Relationship between non-technical skills and technical performance during cardiopulmonary resuscitation: does stress have an influence? *Emerg Med J* 2017;34(11):728–733.
14. Van Merriënboer JJ, Sweller J. Cognitive load theory in health professional education: design principles and strategies: cognitive load theory. *Med Educ* 2010;44(1):85–93.
15. Young JQ, Van Merrienboer J, Durning S, Ten Cate O. Cognitive load theory: implications for medical education: AMEE Guide No. 86. *Med Teach* 2014;36(5):371–384.
16. Gutierrez G, Lunsky IO, Van Heer S, Szulewski A, Chaplin T. Cognitive load theory in action: e-learning modules improve performance in simulation-based education. *A pilot study CJEM* 2023;25(11):893–901.
17. Price JW, Price JR, Pratt DD, Collins JB, McDonald J. High-fidelity simulation in anesthesiology training: a survey of Canadian anesthesiology residents' simulator experience. *Can J Anesth* 2010;57(2):134–142.
18. Savoldelli GL, Naik VN, Hamstra SJ, Morgan PJ. Barriers to use of simulation-based education. *Can J Anesth*. 2005;52(9):944–950.
19. Kurrek MM, Fish KJ. Anaesthesia crisis resource management training: an intimidating concept, a rewarding experience. *Can J Anaesth* 1996;43(5):430–434.
20. Paskins Z, Peile E. Final year medical students' views on simulation-based teaching: a comparison with the best evidence medical education systematic review. *Med Teach* 2010; 32(7):569–577.
21. Mills B, Carter O, Rudd C, Claxton L, O'Brien R. An experimental investigation into the extent social evaluation anxiety impairs performance in simulation-based learning environments amongst final-year undergraduate nursing students. *Nurse Educ Today* 2016; 1(45):9–15.
22. LeBlanc VR. The effects of acute stress on performance: implications for health professions education. *Acad Med* 2009;84(Suppl):S25–S33.
23. LeBlanc VR, McConnell MM, Monteiro SD. Predictable chaos: a review of the effects of emotions on attention, memory and decision making. *Adv Health Sci Educ* 2015;20(1): 265–282.
24. Pekrun R, Frenzel AC, Goetz T, Perry RP. The control-value theory of achievement emotions: an integrative approach to emotions in education. In: *Emotion in Education [Internet]*. Elsevier; 2007 [cited 2021 Dec 10]:13–36 Available from: https://linkinghub.elsevier.com/retrieve/pii/B9780123725455500034.

25. Artino AR, Holmboe ES, Durning SJ. Control-value theory: using achievement emotions to improve understanding of motivation, learning, and performance in medical education: AMEE Guide No. 64. *Med Teach* 2012;34(3):e148–e160.
26. Zendejas B, Cook DA, Farley DR. Teaching first or teaching last: does the timing matter in simulation-based surgical scenarios? *J Surg Educ* 2010;67(6):432–438.
27. Inzana CM, Driskell JE, Salas E, Johnston JH. Effects of preparatory information on enhancing performance under stress. *J Appl Psychol* 1996;81(4):429–435.
28. Cannon-Bowers JA, Rhodenizer L, Salas E, Bowers CA. A framework for understanding pre-practice conditions and their impact on learning. *Pers Psychol* 1998;51(2):291–320.
29. Cavalcanti RB, Detsky AS. The education and training of future physicians: why coaches can't be judges. *JAMA* 2011;306(9):993–994.
30. Meece JL, Anderman EM, Anderman LH. Classroom goal structure, student motivation, and academic achievement. *Annu Rev Psychol* 2006;57(1):487–503.
31. Kaplan A, Maehr ML. The contributions and prospects of goal orientation theory. *Educ Psychol Rev* 2007;19(2):141–184.
32. Creswell JW, Plano Clark VL. *Designing and conducting mixed methods research*. 3rd ed. Los Angeles: SAGE; 2018:492.
33. Teddlie C, Yu F. Mixed methods sampling: a typology with examples. *J Mix Methods Res* 2007;1(1):77–100.
34. Palinkas LA, Horwitz SM, Green CA, Wisdom JP, Duan N, Hoagwood K. Purposeful sampling for qualitative data collection and analysis in mixed method implementation research. *Adm Policy Ment Health* 2015;42(5):533–544.
35. Kim J, Neilipovitz D, Cardinal P, Chiu M, Clinch J. A pilot study using high-fidelity simulation to formally evaluate performance in the resuscitation of critically ill patients: the University of Ottawa Critical Care Medicine, High-Fidelity Simulation, and Crisis Resource Management I Study. *Crit Care Med* 2006;34(8):2167–2174.
36. Rudolph JW, Simon R, Dufresne RL, Raemer DB. There's no such thing as “nonjudgmental” debriefing: a theory and method for debriefing with good judgment. *Simul Healthc* 2006; 1(1):49–55.
37. Eppich W, Cheng A. Promoting excellence and reflective learning in simulation (PEARLS): development and rationale for a blended approach to health care simulation debriefing. *Simul Healthc* 2015;10(2):106–115.
38. LeBlanc VR, Regehr C, Tavares W, Scott AK, MacDonald R, King K. The impact of stress on paramedic performance during simulated critical events. *Prehospital Disaster Med* 2012; 27(4):369–374.
39. Harvey A, Bandiera G, Nathens AB, LeBlanc VR. Impact of stress on resident performance in simulated trauma scenarios. *J Trauma Acute Care Surg* 2012;72(2):497–503.
40. Spielberger CD. *State-Trait Anxiety Inventory: Bibliography*. 2nd ed. Palo Alto: CA. Consulting Psychologists Press; 1989.
41. Downing SM. Validity: on the meaningful interpretation of assessment data. *Med Educ* 2003; 37(9):830–837.
42. Barbadoro P, Brunzini A, Dolcini J, et al. Stress responses in high-fidelity simulation and standard simulation training among medical students. *BMC Med Educ* 2023;23(1):116.
43. Dickerson SS, Kemeny ME. Acute stressors and cortisol responses: a theoretical integration and synthesis of laboratory research. *Psychol Bull* 2004;130(3):355–391.
44. Kudielka BM, Schommer NC, Hellhammer DH, Kirschbaum C. Acute HPA axis responses, heart rate, and mood changes to psychosocial stress (TSST) in humans at different times of day. *Psychoneuroendocrinology* 2004;29(8):983–992.
45. Prabhu A, Smith W, Yurko Y, Acker C, Stefanidis D. Increased stress levels may explain the incomplete transfer of simulator-acquired skill to the operating room. *Surgery* 2010;147(5):640–645.
46. Norman G. Likert scales, levels of measurement and the “laws” of statistics. *Adv Health Sci Educ* 2010;15(5):625–632.
47. Sullivan GM, Artino AR. Analyzing and interpreting data from Likert-type scales. *J Grad Med Educ* 2013;5(4):541–542.
48. Sandelowski M. Whatever happened to qualitative description? *Res Nurs Health* 2000;23(4): 334–340.
49. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol* 2006;3(2): 77–101.
50. Kiger ME, Varpio L. Thematic analysis of qualitative data: AMEE Guide No. 131. *Med Teach* 2020;42(8):846–854.
51. Downing SM, Haladyna TM. Validity threats: overcoming interference with proposed interpretations of assessment data. *Med Educ* 2004;38(3):327–333.
52. Stephens GC, Lazarus MD. Twelve tips for developing healthcare learners' uncertainty tolerance. *Med Teach* 2024;46(8):1035–1043.
53. Howard NM, Cook DA, Hatala R, Pusic MV. Learning curves in health professions education simulation research: a systematic review. *Simul Healthc* 2021;16(2):128–135.
54. van Merriënboer JJG. Perspectives on problem solving and instruction. *Comput Educ* 2013; 64:153–160.
55. Liebert CA, Mazer L, Bereknyei Merrell S, Lin DT, Lau JN. Student perceptions of a simulation-based flipped classroom for the surgery clerkship: a mixed-methods study. *Surgery* 2016;160(3):591–598.
56. Liu KJ, Tkachenko E, Waldman A, et al. A video-based, flipped classroom, simulation curriculum for dermatologic surgery: a prospective, multi-institution study. *J Am Acad Dermatol* 2019;81(6):1271–1276.
57. Reed T, Pirotte M, McHugh M, et al. Simulation-based mastery learning improves medical student performance and retention of core clinical skills. *Simul Healthc* 2016;11(3): 173–180.
58. Chiu HY, Kang YN, Wang WL, et al. The effectiveness of a simulation-based flipped classroom in the acquisition of laparoscopic suturing skills in medical students—a pilot study. *J Surg Educ* 2018;75(2):326–332.
59. Uther P, Van Munster KA, Briggs N, O'Neill S, Kennedy S. Introducing early-phase medical students to clinical paediatrics using simulation and a flipped-classroom. *J Paediatr Child Health* 2019;55(9):1107–1112.
60. Bortnick E. The Effects of a Combined Flipped Classroom and Simulation-Based Education Session of Male Infant Genitourinary Exam on Pediatric Resident and Attending Comfort, Confidence, Performance, and Knowledge [Internet] [M.S.]. 2023 [cited 2024 Oct 7]. Available from: https://www.proquest.com/docview/2835797207/abstract/EBFB614C2C384BA1PQ/1
61. Hassan EA, Elsaman SEA. The effect of simulation-based flipped classroom on acquisition of cardiopulmonary resuscitation skills: a simulation-based randomized trial. *Nurs Crit Care* 2023;28(3):344–352.
62. Fischer KM, Sekula MN, Hannigan JD, et al. Effects of simulation-based flipped classroom in developing head and neck examination skills in dental students. *J Dent Educ* 2024;88(6): 848–855.
63. Ohlenburg H, Arnemann PH, Hessler M, Görlich D, Zarbock A, Friederichs H. Flipped classroom: improved team performance during resuscitation training through interactive pre-course content—a cluster-randomised controlled study. *BMC Med Educ* 2024;24(1):459.
64. Zheng Z, Liao J, Zeng L, et al. High-Fidelity patient simulation incorporated into a flipped classroom improves students' long-term knowledge retention of acute organophosphorus pesticide poisoning. *Simul Healthc* 2022;17(1):e68–e74.
65. Haidt J. The emotional dog and its rational tail: a social intuitionist approach to moral judgment. *Psychol Rev* 2001;108(4):814–834.
66. Bockrath R, Osman C, Trainor J, et al. Education scholarship assessment reconsidered: expansion of Glassick's criteria to incorporate health equity. *Acad Med* 2024;99(5):487–492.