

A Mixed-methods Comparison of Participant and Observer Learner Roles in Simulation Education

Mark J. Bullard, MD, MS-HPed, Anthony J. Weekes, MD, Randolph J. Cordle, MD, Sean M. Fox, MD, Catherine M. Wares, MD, Alan C. Heffner, MD, Lisa D. Howley, PhD, and Deborah Navedo, PhD, CPNP

ABSTRACT

Background: Traditional simulation-based education prioritizes participation in simulated scenarios. The educational impact of observation in simulation-based education compared with participation remains uncertain. Our objective was to compare the performances of observers and participants in a standardized simulation scenario.

Methods: We assessed learning differences between simulation-based scenario participation and observation using a convergent, parallel, quasi-experimental, mixed-methods study of 15 participants and 15 observers ($N = 30$). Fifteen first-year residents from six medical specialties were evaluated during a simulated scenario (cardiac arrest due to critical hyperkalemia). Evaluation included predefined critical actions and performance assessments. In the first exposure to the simulation scenario, participants and observers underwent a shared postevent debriefing with predetermined learning objectives. Three months later, a follow-up assessment using the same case scenario evaluated all 30 learners as participants. Wilcoxon signed rank and Wilcoxon rank sum tests were used to compare participants and observers at 3-month follow-up. In addition, we used case study methodology to explore the nature of learning for participants and observers. Data were triangulated using direct observations, reflective field notes, and a focus group.

Results: Quantitative data analysis comparing the learners' first and second exposure to the investigation scenario demonstrated participants' time to calcium administration as the only statistically significant difference between participant and observer roles (316 seconds vs. 200 seconds, $p = 0.0004$). Qualitative analysis revealed that both participation and observation improved learning, debriefing was an important component to learning, and debriefing closed the learning gap between observers and participants.

Conclusions: Participants and observers had similar performances in simulation-based learning in an isolated scenario of cardiac arrest due to hyperkalemia. Findings support current limited literature that observation should not be underestimated as an important opportunity to enhance simulation-based education. When paired with

From the Department of Emergency Medicine (MJB, AJW, RJC, SMF, CMW, ACH) and the Department of Internal Medicine, Division of Critical Care (ACH), Carolinas Medical Center, Carolinas HealthCare System, Charlotte, NC; the Association of American Medical Colleges (LDH), Washington, DC; and the MGH Institute of Health Professions (DN), Boston, MA.

Received August 14, 2018; revision received October 15, 2018; accepted October 22, 2018.

Presented at the 18th International Meeting on Simulation in Healthcare, Los Angeles, CA, January 2018; and the 16th International Meeting on Simulation in Healthcare, San Diego, CA, January 2018.

Stipends (\$100/participant), focus group subject stipends (\$100/participant), transcription costs, and qualitative software costs were funded by the Carolinas Healthcare System Grant for Research and Education in Teaching (GREAT) from the Carolinas Healthcare Systems' Center for Faculty Excellence. The funding agreement ensured the authors' independence in designing the study, interpreting the data, writing, and publishing the report.

The authors have no potential conflicts to disclose.

Author contributions: study design—MJB, LDH, and DN; data acquisition—MJB, CMW, SMF, ACH, RJC, and AJW; data analysis and interpretation—MJB, CMW, SMF, ACH, and DN; drafting of manuscript writing—MJB, CMW, SMF, ACH, RJC, AJW, LDH, and DN; and critical manuscript editing—MJB, CMW, SMF, ACH, RJC, AJW, LDH, and DN. All authors read and approved the final manuscript.

Supervising Editor: Daniel P. Runde, MD, MME.

Address for correspondence and reprints: Mark J. Bullard, MD, MS-HPed; e-mail: mark.bullard@atriumhealth.org.

AEM EDUCATION AND TRAINING 2019;3:20–32

postevent debriefing, scenario observers and participants may reap similar educational benefits.

Medical simulation improves knowledge and skills acquisition while creating an educational method with a high level of learner satisfaction.¹ Simulation is a unique educational modality because it alleviates educational constraints of case availability, unifies learner exposure to critical thinking, hones procedural skills prior to performance on patients, allows for exposure to complex ethical and spiritual issues, and refines important interpersonal professional skill-sets.^{2,3}

Drawing from educational and psychosocial theories of Schön, Dewey, Kolb, and Russell,⁴⁻⁷ participation in simulation evokes learning due to deliberate practice and the experiential, activating nature of scenarios. As such, participation in simulation should be critical to learning. Study investigators, however, have anecdotally perceived the learning benefit of the observer role. Bandura's social learning theory further supports the notion that most experiential learning could be accomplished through observation.^{8,9} To date, there is limited evidence to support that participation within scenarios is necessary for maximum learning and performance improvement. The educational benefit of simulation observation is therefore worthy of investigation.⁸⁻¹² We hypothesized that the simple cognitive experience of considering clinical decisions, whether directly participating in a simulated scenario or not, benefits both observers and participants alike, particularly when paired with postevent debriefing.

Simulation education is both labor- and resource-intensive. Understanding the educational value of the learner role and other contributory factors, such as debriefing, may be instrumental to optimizing resource allocation in simulation-based education. If evidence reveals there is improved learning through scenario participation over observation, investments in equipment and dedicated simulation space may be required. Conversely, if evidence reveals that there is improved learning through scenario observation over participation, then a revised approach to the education of many learners with simulation could be employed. In addition, the value of quality postevent debriefing may demand further support for dedicated education of simulation educators. This pilot study aimed to 1) explore the relative impact of simulation-based learner role on learning by evaluating predefined critical actions (CAs) and performance metrics within a single simulated scenario and 2) identify other possible

factors affecting learning for simulation-based learner roles.

METHODS

Study Design, Setting, and Population

The relative effects of learner role as scenario participant or observer within simulation-based education were studied within a well-established simulation-based critical care curriculum from January to August 2016 at a 900-bed hospital that serves as a tertiary referral center, trauma center, and branch campus of a medical school. The Intern Simulation Common Critical Care Curriculum (4Cs) provides focused instruction on core topics and procedural skills imperative for first-year residents caring for critically ill adult patients.¹³

Approximately 60 interns from internal medicine, general surgery, emergency medicine, family medicine, obstetrics/gynecology, and orthopedics are enrolled in the 4Cs annually. Within this curriculum, each intern participated in three separate 4-hour simulation-based sessions over a 6-month period using Laerdal Sim Man 3G (Laerdal Medical AS). Each session was attended by four residents. Within each session, each resident was involved in one of four discrete simulation scenarios as the participant role, with the other three learners observing the scenario via real-time video. There were 12 discrete clinical scenarios within the curriculum. The case scenario of cardiac arrest due to critical hyperkalemia was chosen for this investigation due to its definitive, evidence-based therapies and directly observable and quantifiable CAs.

The investigation scenario was designed for uniformity and used carefully scripted roles, scenario information, actor actions, and cues. Actors within this scenario included a nurse confederate, respiratory technician, and nursing assistant for ventilation and chest compressions as directed by the participant. The mannequin was outfitted with a dialysis catheter and presented as unresponsive, apneic, and pulseless, with a wide-complex rhythm displayed on the monitor. Laboratory values and historical patient information were provided to the scenario participant uniformly after each specific request. Chemistry panel results were provided to the scenario participant 5 minutes into the case, if not previously requested. This allowed participants the opportunity to demonstrate appropriate

clinical management of cardiac arrest due to hyperkalemia if the diagnosis remained elusive. The case concluded with consultation for definitive dialysis care or following 15 minutes of elapsed time.

Immediately following each scenario, the scenario participant and three scenario observers underwent a shared postevent debriefing. Debriefing lasted approximately 45 minutes and explored immediate reactions, case objectives, clinical management, pathophysiology, and guided feedback to improve future clinical performance. Debriefings were conducted by a dyad of four possible facilitators using the PEARLS model of debriefing¹⁴ and consisted of the same predetermined learning objectives for all trainees. The principal investigator (PI) was the lead debriefer, with one of three other faculty serving as an associate debriefer.¹⁵ Each faculty had previous experience facilitating simulation debriefings and had codebriefed in previous years' curricula. The PI led all codebriefings to assure all debriefings were uniform.

The 4Cs curriculum is mandatory for all interns caring for adult patients within our hospital; however, involvement in both the three-month follow-up assessment for this study and the focus group were voluntary. Stipends of \$100 were provided to learners who participated in the follow-up simulation assessment, as well as those who participated in the focus group. Follow-up assessment, video recording, and focus group participation were approved by our institutional review board and study subjects gave written informed consent.

Quantitative Data Sources and Analyses

As shown in Figure 1, 30 learners were recruited and enrolled in this study. One of the four learners within each curricular session was intentionally selected as the investigation scenario participant. Fifteen learners acted as the scenario participant during the curriculum (initial participants). This stratified intentional sample ensured that participant representation was proportional to the size of each residency discipline and prevented oversampling or under sampling of any discipline for the study (see Table 1). The other three learners within each 4Cs session were observers to the investigation scenario. The 15 initial investigation scenario participants were recruited for inclusion in the follow-up simulation assessment. There were no refusals. Fifteen scenario observers were additionally enrolled in the follow-up assessment in a similar intentional sampling to represent relative sizes of each residency discipline within our hospital.

The study involved two simulation assessments: baseline assessment occurred within the 4Cs, and follow-up assessment occurred 3 months after completion of the within-curriculum assessment. The scenario of cardiac arrest due to critical hyperkalemia was used for both assessments. During the follow-up assessment, the 15 initial participants repeated the investigation scenario ("follow-up participants"), and an intentional sample of 15 observers became scenario participants and performed the same investigation scenario ("follow-up observers"). Thus, there were three groups of videos for assessment at follow-up: 1) initial participants, 2) follow-up participants, and 3) follow-up observers. Initial participant performance data (group 1 in the previous sentence) were collected prior to postevent debriefing and were compared to follow-up performance metrics of both follow-up participants (group 2) and follow-up observers (group 3).

Each participant's clinical management was assessed by the number and frequency of completed CAs, time to CA, and performance assessment by two trained raters via retrospective video review using a performance assessment tool (Figure 2). The performance assessment tool evaluated the eight clinical domains included in oral board examination by the American Board of Emergency Medicine: data acquisition, problem solving, patient management, resource utilization, health care provided, interpersonal relations, comprehension of pathophysiology, and clinical competence.^{16,17} The performance assessment contained multiple dichotomous CAs decided a priori by expert consensus. CAs were anchored to each domain to assist raters with assessments. Domain ratings were scored on an 8-point ordinal scale. Independent clinical domain ratings and an averaged overall performance rating were attained for scenario participants. The performance assessment tool was chosen for this study as it has demonstrated high interexaminer agreement for CAs and performance ratings when used by calibrated raters in an oral board specialty examination.¹⁶⁻¹⁸ Additionally, this tool has demonstrated the ability to discriminate training level in the simulation setting,^{19,20} and correlates to objective test scores measuring relevant clinical knowledge.¹⁸ Furthermore, learner assessment with this tool has correlated with real-life performance in the intensive care unit setting.²¹

Steps were taken to optimize rater agreement and reliability and reduce potential bias. Prior to the study, the two raters were calibrated by assessing three video

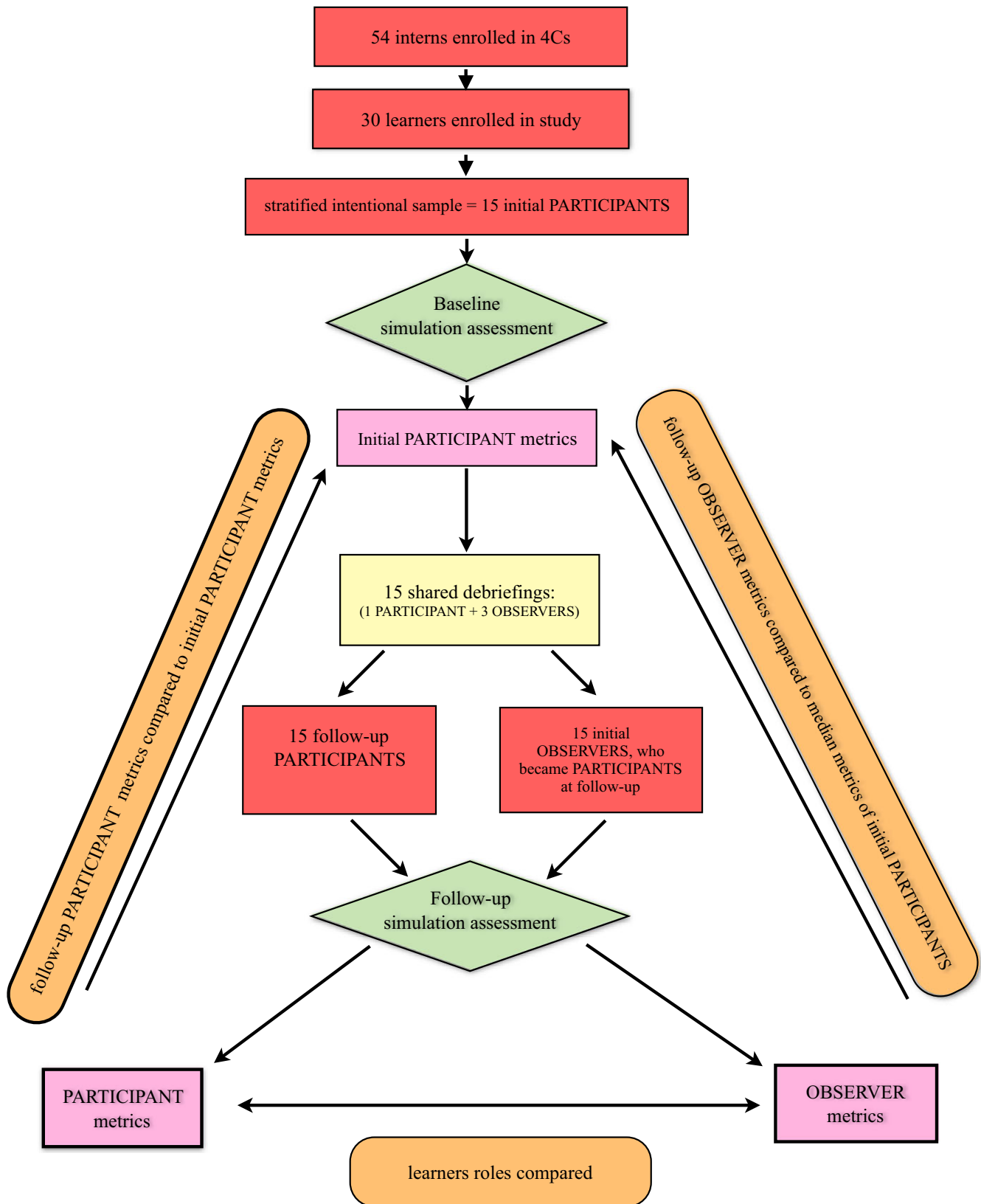


Figure 1. 4Cs quantitative study design flow diagram. 4Cs = Common Critical Care Curriculum.

recordings of the investigation scenario that represented excellent, moderate, and poor clinical performances. These calibration videos did not involve the

study subjects. After this calibration exercise, the PI and raters met to discuss any rating variation. Raters were not involved in the educational curriculum, and

Table 1
Characteristics of Study Subjects and 4Cs*

	Medical Specialty						Sex	
	EM	FM	IM	OB/Gyn	OS	GS	Male	Female
<i>Quantitative Study: Baseline and Follow-up Simulation Assessments (n = 30)</i>								
Initial participants	5	4	3	1	1	1	5	10
Initial observers	6	1	4	1	2	1	3	12
% of study subjects	36	17	23	7	10	7	27	73
Total learners in 4Cs	14	11	12	6	5	6	28	26
% of Learners in 4Cs	25	20	22	11	9	11	53	47
<i>Qualitative Study: Focus Group (n = 6)</i>								
Participants	1	1	1	1	1	1	4	2

4Cs = Common Critical Care Curriculum; EM = emergency medicine; FM = family medicine; GS = general surgery; IM = internal medicine; OB/Gyn = obstetrics and gynecology; OS = orthopedic surgery.

study subjects were identified by unique identifiers to decrease potential bias from possible clinical interactions raters may have had with study subjects. Raters were also blinded to whether scenario videos involved initial participants, follow-up participants, or follow-up observers.

Descriptive statistics, including means, standard deviations, medians, and interquartile ranges (IQR), were calculated for each measure. Participant data were paired within subjects. Since observers did not have initial within-curriculum performance evaluated, follow-up observer data were compared to initial median participant metrics (see Figure 1). Follow-up metrics of both learner roles were also compared. Wilcoxon signed rank and Wilcoxon rank sum tests were used for variables measured on a continuous scale, and Fisher's exact tests were performed for categorical variables. To compare scores between the two raters, a series of t-tests were conducted on each of the eight clinical domains. An intraclass correlation was determined using the raters' average of the eight domain assessments to assess inter-rater reliability and agreement. Analyses were performed with SAS Enterprise Guide Version 5.1 and StatsDirect Version 3.0.183.

Qualitative Data Sources and Analysis

A narrative-based case study methodology was implemented in the tradition of ethnographic approaches to explore perception differences between learner roles regarding their own simulation-based learning.²² This case study was bound by the 4Cs length and a postcurriculum focus group. Two qualitative sources were included.

First, detailed observations and reflective notes regarding learners and their simulation learning experiences were documented in study journals by faculty

(MB, SF, CW, AH). These journals yielded reflections and initial analysis during the study period. Recognizing reflexivity and bracketing as an important component to qualitative study, faculty were reminded to be explicit in their biases within their reflective journals, as well as in any interpretations.

Second, a voluntary postcurriculum focus group was conducted with a sample of six learners representing each of the six specialties and both learner roles. The focus group was led by the 4Cs director (PI), who has 10+ years of simulation debriefing experience, qualitative study expertise, and an MS-HPed degree. A structured guide with open-ended questions was used to guide the discussion of trainee perspectives on learner role and its impact on learning (Data Supplement S1, available as supporting information in the online version of this paper, which is available at <http://online.library.wiley.com/doi/10.1002/aet2.10310/full>). Since all learners fulfilled both participant and observer roles during the 4Cs, each had perspective on both experiences. Additional probing questions were used to clarify examples recalled from past experiences and other deeply held beliefs. The focus group also attempted to uncover factors that affected simulation-based learning and performance within the context of cardiac arrest due to critical hyperkalemia.

Qualitative data analysis was conducted as an applied thematic analysis²³ and was overseen by an experienced qualitative PhD researcher/educator (DN) who was not involved in the 4Cs. Two researchers conducted the analysis (MB and DN). Qualitative data from the focus group were audio recorded, professionally transcribed verbatim, and analyzed using both nVivo software and manual techniques. MB and DN reviewed the focus group data to identify emergent codes based on previously published theories on learner role within

simulation-based education. Data were coded independently, and these codes were discussed among the researchers until a consensus was reached. The final codes were clustered and grouped into emergent themes in an iterative process. Member checking was accomplished through sharing the resultant themes with learners via e-mail and asking them to provide

feedback regarding accuracy of the results to ensure their perceptions were appropriately represented.

RESULTS

Table 1 displays the specialty (medical discipline) and gender of all study subjects. The majority of subjects in the *quantitative study* were female (73%) and there were no significant differences between initial participants and initial observers, except for a slightly higher representation of family medicine among initial participants compared with observers (four and one, respectively). Table 1 shows that each specialty was represented in the *qualitative study*, and four of the six focus group participants were male.

Quantitative Results

Our performance assessment tool demonstrated excellent reliability. The intraclass correlation of the raters' average of the eight domain assessments was 0.92. While rater 1 assigned statistically significant greater scores than rater 2 on every domain, reliability remained high in that each rater agreed on the relative performance of the participants (Figure 3). Since rater performance assessments were found to have excellent reliability, the average of the two raters' domain

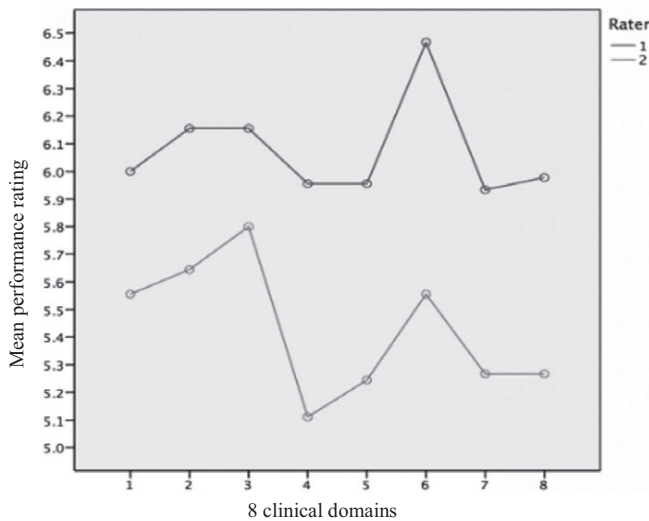


Figure 3. Graphical depiction of two rater performance ratings for each clinical domain.

Table 2
Performance Domain Assessments

Assessments	Participant						Observer			Wilcoxon Signed Rank (Initial Participant vs. Follow-up Participant)	Wilcoxon Rank Sum (Initial Participant vs. Follow-up Observer)	Wilcoxon Rank Sum (Follow-up Participant vs. Follow-up Observer)
	Initial (Average 1 and 2)			Follow-Up (Average Raters 1 and 2)			Follow-Up (Average Raters 1 and 2)					
	No.	Median	IQR	No.	Median	IQR	No.	Median	IQR			
Overall domain rating	15	6	5.0–6.0	15	6	5.5–6.5	15	5.75	5.25–6.75	0.17	0.48	0.77
Comprehension of pathophysiology	15	5.5	5.0–6.5	15	6	5.5–6.5	15	6	4.0–6.5	0.39	0.65	0.84
Clinical competence	15	6	5.5–6.0	15	6	5.0–6.5	15	5.5	5.0–6.5	0.82	0.98	0.93
Data acquisition	15	6	4.5–6.0	15	6	6.0–6.5	15	6	5.0–6.5	0.03	0.40	0.40
Health care provided	15	5.5	4.0–6.0	15	6	5.5–6.5	15	5.5	5.0–6.5	0.17	0.34	0.58
Interpersonal relations	15	5.5	5.0–6.5	15	6.5	5.5–6.5	15	6.5	5.5–6.5	0.26	0.26	0.89
Patient management	15	6	5.5–6.5	15	6	5.5–7.0	15	6	5.5–7.0	0.79	0.76	0.68
Problem solving	15	6	5.0–6.5	15	6	5.5–7.0	15	6	5.5–6.5	0.05	0.47	0.39
Resource utilization	15	5.5	4.5–6.0	15	5.5	5.0–6.0	15	6	5.0–6.5	0.56	0.16	0.55

IQR = interquartile range.

performance ratings was used to compare initial participant and 3-month follow-up performance data.

Averaged overall participant performance rating demonstrated no statistical improvement. Likewise, there was no improvement in averaged overall ratings for observers. Participants, however, were rated higher upon follow-up assessment in two independent clinical domains reaching statistical significance: data acquisition [6 (IQR = 4.5–6.0) – 6 (IQR = 6.0–6.5)] and problem solving [6 (IQR = 5.0–6.5) – 6 (IQR = 5.5–7.0)]. These improvements were driven by the difference in IQR although the medians remained the same. When comparing participants and observers at the follow-up assessment, no significant difference was found in averaged overall rating or any independent clinical performance domain (Table 2).

Tables 3 and 4 display participant and observer metrics for time to CA completion, number of CAs completed, and individual CA completion frequency. Median time to CA completion improved for six of six CAs measured for participants, and four of six for observers over the assessment interval. Except for participants' time to calcium administration (316 to 200 seconds; $p = 0.0004$), improvements in time to CA completion did not rise to the level of statistical significance. Analyses comparing follow-up participants to follow-up observers for time to CA completion demonstrated a statistically significant difference between groups for calcium administration (200 and 238 seconds, respectively; $p = 0.02$). Both learner roles demonstrated no statistical improvement in the median number of CAs completed. Initial participants completed a median number of 5 CAs (IQR = 4–6). The total number of CAs completed in the follow-up assessment for participants and observers were 6 (IQR = 4–6) and 5 (IQR = 5–6), respectively. The findings were similar when analyzed by each individual CA completion frequency.

To identify confounding exposure to critical hyperkalemia during the study period, learners were surveyed regarding exposure to lectures and patient encounters at the initiation and completion of the investigation period (Tables 5 and 6). There was no statistically significant change in education on hyperkalemia or patient encounters external to the 4Cs for scenario participants. Observers, however, were exposed to a statistically significant number of external hyperkalemia lectures ($p = 0.0024$). Observers were exposed to nearly two more lectures on hyperkalemia compared with participants ($p = 0.0426$) over the 9-

month investigation period. Learner roles did not differ in hyperkalemia patient encounters.

Qualitative Results

Qualitative data uncovered six themes pertaining to factors that affect learning in the simulation-learning environment: 1) *Participation improved learning*: scenario participation created an activated learning environment and facilitated increased learning. One learner stated: "... being the active participant ... [makes me] always learn ... because the pressure is on, and makes it stick ... because you are responsible for the decisions. When you're in the room it feels real, and [only] you are responsible for the outcome of the scenario." 2) *Observation improved learning*: observing scenarios engaged learners and facilitated learning, but perhaps not to the same degree as participation. One learner remarked: "[One] could argue that the person actually in the scenario is more engaged ... but I never actually felt that way. [As an observer] you are [mentally] in [the scenario]; trying to problem solve with [the scenario participant] ..." Another learner noted that even "... when you are the observer, you are still participating on a macro level. Because the case is going on in real-time and [the learners] are talking about the case, you are [mentally] going through the case, like you were [actually] in there." 3) *Debriefing was key*: debriefing was necessary for maximal learning in both learner roles. "I think whether you were the observer or the participant, everyone was engaged and debriefing sessions took the scenario to another level." 4) *Debriefing closed the learning gap*: debriefing elevated learning benefits of observation to that of scenario participation. "I think you could argue that the debrief is probably the most important part for the people who are the observers. Even though you are not [participating] within the scenario, the group discussions are where you come full circle and reflect on the case." 5) *Interdisciplinary richness was valued*: interdisciplinary perspective was important to learning and enabled critical thinking. One learner stated: "It's hearing other people's perspectives and how they would have approached that scenario." Another remarked: "I learned ... from talking and problem-solving with other people and learning from [their] experience." 6) *Safe learning environment was critical*: a constructive, safe learning environment was crucial for simulation-based learning. Best stated by one of the learners: "A non-threatening environment allowed learners to open up and reflect."

Table 3
Time to CA* and Total Number of CAs Completed

	Participant						Observer						Wilcoxon			
	Initial			Follow-up			Follow-up			Signed Rank (Initial Participant vs. Follow-up Participant)			Rank Sum (Initial Participant vs. Follow-up Observer)		Wilcoxon Rank Sum (Follow-up Participant vs. Follow-up Observer)	
Time to CA	No.	Median	IQR	No.	Median	IQR	No.	Median	IQR	No.	Median	IQR	Percent	Wilcoxon Rank Sum	Wilcoxon Rank Sum	Wilcoxon Rank Sum
Epi or Vaso	15	439	208-900	15	228	110-900	15	235	144-900	15	235	144-900	0.64	0.51	0.84	0.84
Ca+ Gluc or chloride	15	316	184-426	15	200	142-240	15	238	210-522	15	238	210-522	0.0004	0.96	0.02	0.02
Compressions	15	59	32-900	15	40	35-60	15	54	37-78	15	54	37-78	0.06	0.39	0.34	0.34
Dialysis	15	518	500-900	15	307	270-735	15	648	335-900	15	648	335-900	0.14	0.78	0.22	0.22
Ventilation	15	146	53-259	15	90	46-812	15	102	75-124	15	102	75-124	0.95	0.45	0.94	0.94
Pulse	15	30	20-140	15	27	20-31	15	30	18-38	15	30	18-38	0.09	0.48	0.59	0.59
No. of CAs	15	5	4.0-6.0	15	6	4.0-6.0	15	5	5.0-6.0	15	5	5.0-6.0	0.37	0.33	>0.99	>0.99

CA = critical action; Ca+ Gluc = calcium gluconate; Epi = epinephrine; Vaso = vasopressin.
*If critical action was not completed, then time was set to 900 seconds.

Table 4
Frequency (Freq) of Completed CAs

CA Freq	Participant						Observer						p-value	
	Initial	Freq	Percent	Follow-up	Freq	Percent	Initial vs. Follow-up Observer	Follow-up	Freq	Percent	Follow-up Participant vs. Follow-up Observer	Follow-up Observer	Follow-up Observer	
Pulse	Yes	12	80.00	Yes	15	100.00	0.60	Yes	14	93.33	1.00	1.00		
	No	3	20.00	No	0	0.00		No	1	6.67				
Compression	Yes	11	73.33	Yes	14	93.33	0.33	Yes	14	93.33	1.00	1.00		
	No	4	26.67	No	1	6.67		No	1	6.67				
Ventilation	Yes	13	86.67	Yes	12	80.00	1.00	Yes	14	93.33	0.60	0.60		
	No	2	13.33	No	3	20.00		No	1	6.67				
Epi or Vaso	Yes	9	60.00	Yes	8	53.33	1.00	Yes	10	66.67	0.71	0.71		
	No	6	40.00	No	7	46.67		No	5	33.33				
Ca+ Gluc or chloride	Yes	15	100.00	Yes	15	100.00	1.00	Yes	14	93.33	1.00	1.00		
	No	0	0.00	No	0	0.00		No	1	6.67				
Dialysis	Yes	12	80.00	Yes	14	93.33	1.00	Yes	12	80.00	0.60	0.60		
	No	3	20.00	No	1	6.67		No	3	20.00				

CA = critical action; Ca+ Gluc = calcium gluconate; Epi = epinephrine; IQR = interquartile range; Vaso = vasopressin.

Table 5
Hyperkalemia (K+) Lectures and Patient Encounters

	Participant						Observer					
	Precurricular			Postcurricular			Precurricular			PostCurricular		
	No.	Median	IQR	No.	Median	IQR	No.	Median	IQR	No.	Median	IQR
K + lectures	15	1	0–2	15	2	0–3	15	2	1–2	15	4	1–7
K + patients	15	1	0–2	15	0	0–2	15	1	0–2	15	1	0–2

IQR = interquartile range.

Table 6
Hyperkalemia (K+) Lectures and Patient Encounters

	Participant: Pre vs. Post (Wilcoxon Signed Rank Sum)	Observer: Pre vs. Post (Wilcoxon Signed Rank Sum)	Pre-4Cs Participant vs. Observer (Wilcoxon Rank Sum)	Post-4Cs: Participant vs. Observer (Wilcoxon Rank Sum)
K + Lectures	0.1602	0.0024	0.2986	0.0426
K + Patients	0.5625	1.0000	0.7613	0.6882

4Cs = Common Critical Care Curriculum.

DISCUSSION

Adult education theory suggests that learners who are more activated during an experience should demonstrate improved learning and clinical performance as opposed to less activated learners. As such, it has been assumed that scenario participation has an important role for learning and performance improvement when compared to observation. However, simulation-based observation may prove as beneficial as scenario participation to learning *if* the observer is appropriately “activated.”

Limited studies have demonstrated the benefit of simulation observation;^{11,12,24} however, factors that affect observational learning have yet to be defined. To date, most simulation studies have had observers use structured observation tools, such as CA checklists, to activate them during the experience.^{8,11,24} Kaplan et al.,¹¹ for example, demonstrated no learning differences between participants and observers; however, observer activation was augmented by the use of CA checklists to critique participant performance while watching the scenario unfold. Bong et al.¹² promoted observer activation by having the observer believe that he or she might be called upon to assist within the scenario while observing directly in the patient encounter room. Conversely, our study demonstrated benefit in the “nondirected” observer role without tools, specific guidance, or objectives to maintain observer activation.⁸ This finding suggests that the use of specific observational tools or “directed observation” may not be required for observer activation and simulation observation to be beneficial.

When evaluating the success of an educational intervention, the recognized confounder of skill decay must be considered. Skill and knowledge can quickly decay after simulation-based education,^{25–28} with one study reporting decay as early as 2 weeks.²⁷ Educational theory suggests the benefit of simulation stems from repetitive practice.⁶ Yet even with repetition, skills and knowledge decay, Braun et al.²⁶ demonstrated that learner performance declined at 2 months even with repetitive practice until “mastery-level performance” was met. Cardiopulmonary resuscitation skills and knowledge using simulation education in nursing students demonstrated decreased retention at 3 months despite having over 4 hours of training.²⁵ Our education on critical hyperkalemia consisted of only 45 minutes of postevent debriefing.

Along with the issue of skills decay, there is also current uncertainty in appropriate timing of education assessment intervals to evaluate learning within educational interventions. Assessment periods with short intervals may not reflect the effect of knowledge and skills decay. Thus, short duration between testing periods may bias previous findings comparing scenario participation and observation.^{11,12,24} Stegmann et al.,²⁴ for example, had interval learning assessments only days after the educational intervention. Longer assessment intervals may be biased due to failure to account for external education or influences, such as repetitive practice. Bloch and Bloch¹⁰ demonstrated the benefits of observational simulation learning among pediatric emergency medicine fellows over a 7-month assessment period. However, learning in their

study was influenced by repetition of tested skills within curriculum, as well as learning within the larger pediatric emergency medicine fellowship, which included education on similar material. Additionally, Bong et al.¹² demonstrated the equivalence of observational learning for nontechnical skills when compared with scenario participation; however, their study included repetitive education and a short, 3-week interval assessment period. Learners in our study demonstrated improvement trends over a lengthy (3-month) testing interval without the benefit of repetitive practice, with minimal external education on hyperkalemia, no significant hyperkalemic patient encounters and despite expected knowledge/skill decay. Additionally, our learners were multidisciplinary, with many specialties having little or no external daily experience with the investigation subject matter.

Similar to previous work by Gordon et al.,^{20,21} our study further demonstrates the value of using a learner performance assessment tool within a simulation-based environment that was designed for and previously used in specialty oral board examinations.^{16–19} The excellent performance of our tool supports further use of similar learning and performance assessment tools in acute care simulation scenarios.

Our study explored the impact of learner role on simulation-based education. Quantitatively, this study demonstrated no significant advantage for either learner role for the scenario of cardiac arrest due to critical hyperkalemia. Congruent with quantitative findings, emergent themes from the focus group reinforced this theory, further demonstrating the value of nondirected simulation-based observation when paired with postevent debriefing. The results of this study may offer practical solutions to combat challenges faced by today's simulation educators such as space, time, equipment, large numbers of learners, and availability of trained simulation educators. Participants demonstrated statistically significant decreased time to calcium administration when compared with observers. Given that this CA is the most clinically important for initial stabilization of the patient in cardiac arrest due to hyperkalemia, scenario participation would be highlighted as critical to learning. Conversely, as suggested by most of our data analysis, observation is of equal benefit as scenario participation and provides a reasonable approach to educating a large number of learners. Finally, it was apparent that high-quality, faculty-facilitated postevent debriefing elevated the learning benefits of observation to that of participation. As such, it could be surmised that

simulation technology simply provides a social forum for activated didactics, and learning (particularly for scenario observers) is most affected by debriefing, which requires proper education of simulation educators.

LIMITATIONS

This study had several limitations. First, the quantitative study was limited by the need to integrate it into a preexisting curriculum and the number of 4Cs learners. While this may have underpowered the study, we enrolled all available subjects for our study population. It is possible this sample size was not large enough to demonstrate differences between learner roles; however, our results did suggest improvement trends in both participants and observers. Using our quantitative results as exploratory data, this pilot may open the door for increased resources and ability to further understand the impact of learner role in simulation-based education.

Second, the simulation-based education on critical hyperkalemia was limited to a single 45-minute shared postevent debriefing. This limited time frame, without the benefit of repetitive practice may have affected study findings. Additionally, the 3-month assessment interval may also have contributed to study findings given previous discussion on knowledge decay and uncertain time-frame for appropriate assessment interval.

Third, unlike participants, observers did not have an initial assessment to pair performance metrics at follow-up assessment. Thus, observers were compared to the performance metrics of the initial within-curriculum participants (Figure 1). This strategy assumed our intentional sampling of 15 initial participants appropriately captured a representative performance for all 54 learners in the 4Cs curriculum. Thus, the initial participant metrics may or may not be a true representation of the entire learner group. Randomization might be suggested to improve the impact of the study; however, this sampling was implemented to assure representation and perspective across disciplines and avoid over- or underrepresentation of any one specialty. Randomization (if resulting in over- or underrepresentation of any specific discipline) might bias the results due to specific disciplines, such as emergency medicine, being exposed to critical hyperkalemia more than others. Additionally, adult learners have varying motivation to learn the investigation subject matter relative to its applicability to their respective discipline. Finally, the prescheduled curriculum made randomization infeasible. From a

qualitative standpoint, it is worth mentioning that while the stratified sampling of the focus group contributed to the inclusion of all perspectives, the applied thematic analysis methodology²³ does not assure thematic saturation.

Finally, the content of the simulation (cardiac arrest due to critical hyperkalemia) and the training level of the study subjects may affect the generalizability of our results to other learner groups and clinical scenarios. While learners were from different disciplines, all study subjects were trainees 6 months post-medical school graduation and thus had similar baseline knowledge. Plus, hyperkalemia represents a clinical pathophysiology that affects patients on all services. Observers did report two more didactics on hyperkalemia than participants over the course of the 9-month investigation (6-month curriculum + 3-month interval assessment). This, too, may have affected results; however, given the previous discussion on knowledge decay, it likely had little effect.

CONCLUSION

This pilot study provided a preliminary evaluation and comparison of simulation-based scenario participants and observers who experience a shared postevent debriefing. With high-quality debriefing, there was no demonstrable advantage of one learner role in simulation-based education in an isolated scenario of cardiac arrest due to hyperkalemia. Observation should not be underestimated as an important opportunity to enhance simulation-based education. Future work should look to further evaluate the learning benefits of the simulation observer role and specific factors affecting its benefit.

The authors acknowledge Megan Templin, Dickson Advanced Analytics, Carolinas Healthcare System, who completed quantitative data analysis; and Kelly Goonan, MPH, CPHQ, Scientific Writer, Department of Emergency Medicine, Carolinas Healthcare System, who edited manuscript.

References

1. Cook DA. How much evidence does it take? A cumulative meta-analysis of outcomes of simulation-based education. *Med Educ* 2014;48:750–60.
2. Stroud L, Wong BM, Hollenberg E, Levinson W. Teaching medical error disclosure to physicians-in-training: a scoping review. *Acad Med* 2013;88:884–92.
3. Park I, Gupta A, Mandani K, Haubner L, Peckler B. Breaking bad news education for emergency medicine residents: a novel training module using simulation with the SPIKES protocol. *J Emerg Trauma Shock* 2010;3:385–8.
4. Schön D. *The Reflective Practitioner: How Professionals Think in Action*. London: Temple Smith, 1983.
5. Dewey J. *Experience and Education*. New York: Kappa Delta Pi, 1938.
6. Kolb DA, Fry R. Toward an applied theory of experiential learning. In: Cooper C, editor. *Theories of Group Process*. London: John Wiley, 1975.
7. Russell JA. Core affect and the psychological construction of emotion. *Psychol Rev* 2003;110:145–75.
8. O'Regan S, Molloy E, Watterson L, Nestel D. Observer roles that optimise learning in healthcare simulation education: a systematic review. *Adv Simul* 2016;1:4.
9. Bandura A. *Social Learning Theory*. New York: General Learning Press, 1971.
10. Bloch S, Bloch A. Simulation training based on observation with minimal participation improves paediatric emergency medicine knowledge, skills and confidence. *Emerg Med J* 2015;32:195–202.
11. Kaplan BG, Abraham C, Gary R. Effects of participation vs. observation of a simulation experience on testing outcomes: implications for logistical planning for a school of nursing. *Int J Nurs Educ Scholarsh* 2012;9:1–15.
12. Bong CL, Lee S, Ng AS, Allen JC, Lim EH, 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 (Lond)* 2017;2:7.
13. Bullard MJ, Leuck J, Howley LD. Unifying interdisciplinary education: designing and implementing an intern simulation educational curriculum to increase confidence in critical care from PGY1 to PGY2. *BMC Res Notes* 2017;10:563.
14. 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;10:106–15.
15. Cheng A, Palaganas J, Rudolph J, Eppich W, Robinson T, Grant V. Co-debriefing for simulation-based education: a primer for facilitators. *Simul Healthc* 2015;10:69–75.
16. Platts-Mills T, Lewin M, Ma S. The oral certification examination. *Ann Emerg Med*. 2006;47:278–82.
17. American Board of Emergency Medicine. Oral Exam Results and Scoring. Available at: [https://www.abem.org/public/emergency-medicine\(em\)-initial-certification/oral-examination/scoring/performance-criteria](https://www.abem.org/public/emergency-medicine(em)-initial-certification/oral-examination/scoring/performance-criteria). Accessed Jun 1, 2018.
18. Bianchi L, Gallagher EJ, Korte R, Ham HP. Interexaminer agreement on the American Board of Emergency Medicine oral certification examination. *Ann Emerg Med* 2003;41:859–64.

19. Maatsch J. Assessment of clinical competence on the emergency medicine specialty certification examination: the validity of examiner ratings of simulated clinical encounters. *Ann Emerg Med* 1981;10:504–7.
20. Gordon J, Tancredi D, Binder W, Wilkerson W, Shaffer D. Assessment of a clinical performance evaluation tool for use in a simulator-based testing environment: a pilot study. *Acad Med* 2003;78:S45–7.
21. Gordon JA, Alexander EK, Lockley SW, et al. Does simulator-based clinical performance correlate with actual hospital behavior? The effect of extended work hours on patient care provided by medical interns. *Acad Med* 2010;85:1583–88.
22. Creswell JW. *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*, 4th ed. Los Angeles: SAGE Publications, 2014.
23. Guest G, MacQueen K, Namey E. *Applied Thematic Analysis*, 1st ed. Los Angeles: SAGE Publications, 2011.
24. Stegmann K, Pilz F, Siebeck M, Fischer F. Vicarious learning during simulations: is it more effective than hands on learning? *Med Educ* 2012;46:1001–8.
25. Aqel A, Ahmad M. High-fidelity simulation effects on CPR knowledge, skills, acquisition, and retention in nursing students. *Worldviews Evid Based Nurs* 2014;11:394–400.
26. Braun L, Sawyer T, Smith K, et al. Retention of pediatric resuscitation performance after a simulation-based mastery learning session: a multicenter randomized trial. *Pediatr Crit Care Med* 2015;16:131–8.
27. Ellis S, Varley M, Howell S, et al. Acquisition and retention of laparoscopic skills is different comparing conventional laparoscopic and single-incision laparoscopic surgery: a single-centre, prospective randomized study. *Surg Endosc* 2016;30:3386–90.
28. Gallagher A, Jordan-Black J, O’Sullivan G. Prospective, randomized assessment of acquisition, maintenance, and loss of laparoscopic skills. *Ann Surg* 2012;256:387–93.

Supporting Information

The following supporting information is available in the online version of this paper available at <http://onlinelibrary.wiley.com/doi/10.1002/aet2.10310/full>

Data Supplement S1. Guiding questions for focus group.