

Implementing Distributed Feedback in a Tool that Supports Peer-to-Peer Simulation in Healthcare

Sandra Katz, Patricia Albacete, Pamela Jordan, Scott Silliman and Matthew Wrzesniewski

¹ University of Pittsburgh, Pittsburgh PA 15260, USA
katz@pitt.edu

Abstract. MedDbriefer is a web based ITS designed to enable healthcare students to do clinical scenarios anytime, anywhere. While one student “voice treats” a scenario’s patient(s) as the leader of a mock Emergency Medical Services (EMS) team, a peer records the team’s actions by using the system’s checklists, on a tablet. When the scenario ends, MedDbriefer analyzes the event log and generates a debriefing. MedDbriefer also provides a platform for research on simulation-based training. This paper describes how the system’s debriefing engine could be extended to deliver feedback *during* a scenario, as well as afterwards. MedDbriefer could then be used to compare the effectiveness of different ways of timing feedback delivery in computer-based simulation systems.

Keywords: Automated debriefing, Computer-based simulation systems, EMS training, Feedback, MedDbriefer, Microdebriefing, Peer-to-peer simulation.

1 Introduction

Computer-based simulation systems allow students to gain supplemental practice outside of their courses’ regular simulation-based training (SBT) labs and receive personalized feedback from the automated tutor. Although extensive research has shown that feedback is one of the most critical components of SBT, little is known about how to deliver effective feedback during live or automated simulation exercises [1-4].

In automated simulation systems designed for healthcare provider training, feedback is typically deferred until a debriefing, after a clinical scenario ends. Post-scenario debriefings typically present a step-by-step replay of students’ actions, as illustrated in tutoring systems such as the American Heart Association’s *Heartcode BLS* and *Heartcode ALS*, and Wolters Kluwer’s *VSim for Nursing* [5, 6]. Some human SBT instructors also avoid giving feedback during a scenario, except when students stray far off the path to a successful solution [7]. This approach to timing feedback delivery prevents the cognitive overload that can result from students having to shift attention between a scenario (or other simulation exercise) and the instructor’s feedback, preserves scenario realism, and provides opportunities for students to experience “productive struggle” and self-regulate [e.g., 3, 8]. However, several simulation scholars have advocated replacing post-scenario debriefings with within-scenario debriefings—that is, brief pauses to a scenario to address errors or prompt students to reflect on their

performance so far.¹ Compared with post-scenario debriefings, within-scenario debriefings position feedback closer to the context in which errors occurred, prevent the frustration that can stem from realizing that the scenario's patient is not improving, and reduce the chance that persistent errors become cognitively ingrained [e.g., 7, 8, 9].

Our observations of human instructor guided scenarios for training emergency medical services (EMS) providers revealed that instructors typically take a hybrid approach, parceling feedback between brief within-scenario debriefings and a post-scenario debriefing. In future work, we plan to compare alternative ways of timing the delivery of feedback, to determine which approach, if any, predicts better learning outcomes in a computer-based simulation system. MedDbriefer, a web based ITS that supports peer-to-peer simulation, will serve as the research platform for these studies [10-12].

After presenting an overview of MedDbriefer, this paper illustrates distributed feedback during human facilitated simulation. It then describes how MedDbriefer analyzes simulation logs to generate post-scenario debriefings, and how this approach could be leveraged to provide within-scenario and distributed debriefings.

2 MedDbriefer

Students who struggle to acquire clinical skills often benefit from supplemental simulation practice, outside of their regular SBT labs. Unfortunately, programs across healthcare specialties face a shortage of simulation instructors [e.g., 13]. To address this problem, many instructors encourage students to get together with peers and practice scenarios. However, left unguided, peer-to-peer simulation often deteriorates, emblemizing “the blind leading the blind” [14].

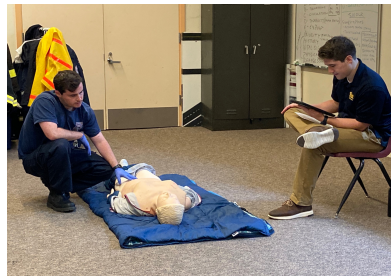


Fig. 1. MedDbriefer in use. Paramedic student (at left) treats a simulated patient while peer (at right) uses MedDbriefer's checklists to log actions.

When fully developed MedDbriefer will enable pairs or small groups of paramedic trainees to engage in clinical scenarios on their own—anytime, anywhere. While one student “voice treats” a virtual patient as the leader of a mock EMS team, a peer uses the system on a tablet to log the team leader's verbalized actions, by selecting these actions from the interface's checklists. (See Figures 1-2.) After the scenario ends, the system analyzes the event log and generates a debriefing. In a randomized trial that is nearing completion, we used MedDbriefer to compare two ways to structure post-scenario debriefings [12]. Table 1 illustrates the traditional ap-

proach: a chronological replay of students' scenario actions, with feedback.

MedDbriefer implements several affordances to foster beneficial interaction between the student who voice treats the simulated patient (the “EMS team leader”) and a peer who uses the tablet to log the team's actions (the “observer”). As shown in

¹ Within-scenario debriefings are alternatively called “microdebriefings,” “reflective pauses,” “stop-and-go debriefings,” among other terms.

Figure 2, MedDbriefer’s observer interface (OI) provides two main checklists: an assessment checklist (Figure 2, left) and intervention checklist (Figure 2, right). When the observer checks an assessment action, the system displays a finding to call out. The observer does not need to invent a finding that is realistic and accurate. For example, if the team leader states that he is checking breathing quality, the observer may be cued to call out “gurgling,” as highlighted in yellow in Figure 2. Similarly, the intervention checklist includes routine actions that EMS providers perform, such as ventilating a patient and securing a patient onto a spinal board. Interspersed throughout this menu are prompts for the observer to issue if the team leader fails to provide sufficient detail about how he would perform an intervention. For example, Figure 2 (right) displays a respiratory intervention, bag-valve mask ventilation, with questions the observer should ask if the team leader doesn’t specify the ventilation rate, oxygen flow rate, etc..

Providing feedback during a post-scenario debriefing is MedDbriefer’s most important affordance. Immediately after the observer clicks on the Done button, the system analyzes the event log to generate a debriefing (e.g., Table 1), as described presently (Section 4.1). Ultimately, MedDbriefer could be scaled to support peer-to-peer simulation in other areas of healthcare education, such as physician and nurse training.

BSI and Scene Size-up

Primary Survey

- INITIAL ASSESSMENT
- AIRWAY
- BREATHING**
- CIRCULATION
- TRANSPORT DECISION
- TRAUMA EXPOSE

History Taking

Secondary Survey

- HEAD
- NECK
- CHEST
- ABDOMEN
- PELVIS
- EXTREMITIES
- POSTERIOR THORAX, LUMBAR, AND BUTTOCKS

Ongoing Management & Reassessment

MedDbriefer SCENARIO: SC8CP SAVE INFO 18:27

Primary Survey/Breathing checklist

Stated what you were looking/listening/feeling for while checking breathing.

Checked breathing

- ☒ rate ☐ Very slow respirations
- ☒ rhythm ☐ Unremarkable
- ☒ quality (sounds, depth, chest rise/fall) ☐ Gurgling

Interventions

Bag-valve mask

At what rate will you ventilate the patient?

☒ 10-12 BPM ☐ 20 BPM ☐ Other Clear

Oxygen flow rate?

☒ 15 L/min or greater ☐ Other Clear

Target O2 saturation?

☒ 95-99% ☐ Other Clear

Continue

Patient Status

Vitals SAMPLE OPQRST LOC A B C

Vital	Current Value	Action
Pain	Not available	Requested
P	66, weak	Requested
Skin	Cyanosed, cool, wet	Requested
Spo2	No capture	Requested
Temp	95.3 degrees F (35.2 degrees C)	Requested
R	6, gurgling; LS rhonchi bilaterally	Requested
ETCO2	56 mm Hg	Requested
BP	78/50 (MAP 59)	Requested
Glucose	100 mg/dl (5.6 mmol/l)	Requested
GCS	3 (E-1, V-1, M-1)	Requested

Intervention Status

Intervention	Status
Bag-valve mask	after adjustments, breathing assisted Requested

Fig. 2. MedDbriefer's observer interface.

3 An Example of Distributed Feedback

Table 2 shows excerpts of distributed feedback that took place during a live, human instructor guided paramedic training scenario. At the start of the scenario, the instructor (simulation facilitator, abbreviated “Fac” in Table 2) stated the dispatch information:

You are dispatched to a 26-year-old female having trouble breathing. She's at a bus terminal, sitting on the curb.

Table 1. Excerpt from a MedDbriefer post-scenario debriefing on a child near-drowning

ID	Action description	Feedback
15	 Performed bag-valve mask ventilation. Purpose: Oxygenation Oxygen flow rate: 15 liters/min or greater Target O2 saturation rate: 95-99% Ventilation rate: 10-12 BPM	<i>Incorrectly administered intervention.</i> You were incorrect about ventilation rate. This patient should have been ventilated at a rate of 20 breaths per minute, which is the recommended rate for a child aged 2-12. When performing BVM ventilation, you use 15LPM O2 to both oxygenate and ventilate the patient to a target SpO2 >90%.
16	 Checked skin condition (moisture). Found: <i>Wet</i>	
17	 Performed a gross blood sweep. Found: <i>No major external bleeding</i>	<i>Missing assessment step.</i> A gross blood sweep (negative in this patient) will help you identify any life-threatening bleeding you may have missed earlier when forming your general impression.
18	 Checked skin color. Found: <i>Cyanosed</i>	
19	 Which pulse? Pulse to check: carotid and radial	
20	 Checked pulse rate. Found: <i>Slow</i>	<i>Mistimed assessment step.</i> All unconscious patients should have a pulse check before starting the ABCs. However, paramedics often check breathing while doing a pulse check. If the patient is pulseless, you will start high-quality CPR immediately...

The facilitator paused the scenario several times to achieve an overarching goal: to get the team leader (abbreviated “TL” in Table 2) to do a more thorough patient assessment before settling on a diagnosis—specifically, to understand *why* the patient is in respiratory distress. Through within-scenario debriefings, often extended during the post-scenario debriefing, the instructor addressed a common problem in healthcare education: premature closure, aka “jumping to conclusions” [e.g., 15]. Especially when patients present with medical conditions, as opposed to physical injury, paramedics must gather enough evidence to determine which of several possible conditions is the most likely cause of ambiguous symptoms like “difficulty breathing,” so that they can manage the patient appropriately until they arrive at a definitive care facility.

Table 2. Example of distributed debriefing during a human facilitated scenario

Pause# (elapsed time in minutes)	Within-scenario Debriefings	Excerpts from the post-sce- nario debriefing that extend within-scenario debriefings
1 (~12:00)	1. Fac: OK; let's stop. What else do you know about her? 2. TL: She's a 26-year-old female. 3. Fac: OK. What else do you know? 4. TL: Nothing. 5. Fac: What about SAMPLE, OPQRST?	18. TL: I didn't think it was a cardiac problem once I put her on the monitor... 19. Fac: Don't females have them...especially if they're pregnant, diabetic, blah blah blah, right?...That's why you need to ask all those questions. She could be having a STEMI.
2 (~15:00)	6. Fac: So, let's pause again. What's the problem? Why are we here? 7. TL: Shortness of breath. I think she might have a PE. 8. Fac: But is it a breathing problem? Does she have a cough? Cold? Is the cold productive? Does she smoke? You know...think of all the breathing questions. When you have belly pain it's all the belly questions...	
3 (~19:00)	9. Fac: What's going on? 10. TL: I think she has a PE 11. Fac: Why? 12. TL: Pleuritic chest pain on one side. She's on birth control. 13. Fac: There's many people that have pleuritic chest pain. Maybe she lifted an air conditioner yesterday...What else do you know about her? 14. TL: Shortness of breath. Are you a smoker? 15. Patient: yes... 16. Fac: Where is she? 17. TL: Bus terminal...She's been sitting for a long time. That's a contributing factor for an expected PE.	20. Fac: Ok...She's 26, she's on birth control, she's been sedentary for a couple of days. She smokes. She does have the pleuritic chest pain. She's a little bit tachycardic, her oxygen saturations are low. Is that a much better story for a PE?... Like you hit PE super early on, but you didn't know why PE. You had like, it was a little piece of the puzzle it was a good guess...I think you did a really good job. You got to the point where you went from a guess to an educated informed, medical—I think they call it a “differential diagnosis” or something; a field impression.

During the scenario excerpted in Table 2, the team leader suspected early on that the patient was suffering from a pulmonary embolism (PE) (e.g., Pause 2, turn 7), which is a clot that blocks blood flow to an artery in a lung. Although the student happened to be correct in this case, the facilitator was not happy with the *process* by which the student reached this diagnosis. The student had gathered other findings besides difficulty breathing that could indicate a PE such as chest pain, low blood pressure, and poor oxygen saturation (not shown in Table 2). However, a similar cluster of findings

could indicate a cardiac issue, including a heart attack. At each pause in the scenario, the instructor prodded students to gather more evidence, to rule out other possible conditions and strengthen the team leader’s proposed diagnosis (i.e., a PE).

Errors trigger within-scenario pauses. For example, Pause 1 in Table 2 is triggered by the team leader’s failure to take a patient history after administering oxygen and realizing that the patient still had difficulty breathing (i.e., SAMPLE and OPQRST history, referred to in turn 5).² Findings gathered from history taking would strengthen or weaken the likelihood of a PE. Continuing in the same vein, Pause 2 is triggered by TL’s failure to ask focused questions that would reveal more about the possible causes of the patient’s “breathing problem” (turn 8). Similarly, Pause 3 (turns 9-17) is triggered by the student’s failure to inquire about other telltale factors that could lead to a PE, such as being a smoker and/or sedentary for considerable time.

During the post-scenario debriefing, the facilitator clarified why he prodded the team leader to “ask all those questions” (turn 19): because the patient could be experiencing a “STEMI” (i.e., an “**ST-Elevation Myocardial Infarction**,” aka a “heart attack”), not a PE. As stated in turn 20, the facilitator tried to steer students from taking a “shot in the dark” (guess) towards deriving an “informed, medical...differential diagnosis.”

4 Towards Automated Distributed Feedback in MedDbriefer

MedDbriefer analyzes the entire event log after a scenario ends in three phases to generate a debriefing (e.g., Table 1). This section summarizes the log analysis process (Section 4.1; see [11] for more detail) and how it could be leveraged to implement within-scenario and distributed feedback (Section 4.2). In essence, instead of analyzing the entire log at the end of a scenario, the system would analyze the log dynamically, at selected “checkpoints,” pausing to provide feedback on students’ performance between checkpoints. Examples in this section refer to the near-drowning scenario associated with the interface screenshot shown in Figure 1 and the debriefing excerpt shown in Table 1. The scenario starts when the observer reads the dispatch information:

You and your partner are working in a suburban EMS service. You are dispatched to a single-family home for a child drowning. It is 19:30 hr. and the air temperature is 87°F. From the scene you are 11 minutes to a level 2 trauma center and 16 minutes by air to a level 1 pediatric trauma center.

4.1 How MedDbriefer Generates Post-scenario Debriefings

In addition to the scenario event log (EL), each analysis phase uses two other types of data: the assessment hierarchy (AH) and the management hierarchy (MH). The AH is a downward branching tree whose parent node is the goal of completing a full patient assessment and branches are assessment phases and subphases. Figure 1 (left) shows the top two levels of the assessment hierarchy. Lower levels are displayed when the

² SAMPLE and OPQRST are acronyms for history-taking questions; for example, S=Signs/Symptoms; A=known Allergies; M=Medications; O=Onset of pain; P=Palliation/Provocation (i.e., what makes the pain better/worse?).

observer selects a menu item. For example, Figure 1 shows the checklist that would appear when the observer selects *Breathing* in the *Primary Survey* menu.

Like the assessment hierarchy, the management hierarchy is a downward branching tree whose parent node is the goal of managing the clinical problems identified during patient assessment, children are separate problems (e.g., severe bleeding, hypovolemic shock), and grandchildren are interventions necessary to address these problems, including appropriate alternatives. For example, in the near-drowning scenario, the main management goals are to control the patient's obstructed airway and compromised breathing. Managing the child's airway requires suctioning and, eventually, intubation.

Analysis Phase 1: Interpreting the Event Log. During the first phase of analysis, observed events in the EL are interpreted by comparing them to two models: the expected patient assessment actions specified in the AH and solutions to clinical problems specified by domain experts, represented in the MH. In addition, the system scores any responses to the observer's requests for additional details (e.g., Table 1, ID 15). Interventions (the leaf nodes) in the MH, are designated as either "required" or "optional" and there may be more than one acceptable alternative for required interventions. Interventions that are not part of any solution are designated as "not indicated."

There is usually a simple one-to-one mapping between assessment actions in the EL and the AH, and between interventions in the EL and the lowest levels and leaf nodes in the MH. The analysis system picks the solution path that best fits the EL. By associating events in the EL with items in the AH and the MH, the system gains knowledge about the possible role of each event, such as what to expect some time before or after a particular event and the purpose of that event. This information facilitates recognition of assessment sections and management goals that may not have been completed during one contiguous time frame—for example, the student interrupted an assessment section to start a different one and returned to the interrupted section later. It is also used as part of checking temporal constraints in the second analysis phase and organizing the final debriefing presentation in the third phase, as described presently.

The AI in Analysis Phase 1 is this matching process—a search to find the solution path that best explains the events logged. Similar approaches have been used in other intelligent tutoring systems [16, 17]—that is, generate solutions and do plan recognition by matching observations of what the student did to possible solutions [18].

Analysis Phase 2: Applying temporal constraints. Some interventions must be performed in an expected order to be effective, whereas timing is less critical for other interventions. For example, in the near-drowning scenario, it is important to suction the child's airway to clear it before oxygenating and ventilating him. We represent temporal constraints as decision rules and use these rules to assess the ordering of actions recorded in the EL. For example, Table 1 (ID 20) illustrates a temporal constraint violation. The student checked the unconscious patient's pulse late. Temporal representations and constraints, and constraints in general, are part of problem solving and plan recognition and thus are important in reasoning [e.g., 19, 20].

Analysis Phase 3: Identifying and marking missing actions. In the final phase of analysis, missing actions are identified and inserted in the section of the annotated EL in which they best fit and are assigned a status of "missing." The suggested orderings

implied by the AH and MH are utilized so that missing actions are inserted in the annotated log where they are inferred to be most appropriate. The insertion heuristic first tries to locate other events related to the same assessment phase or management goal and inserts the missing one relative to the ordering specified in the AH or MH. If a management goal is missing entirely from the student's solution, the missing intervention is inserted at the end of the assessment section in which the MH indicated it should appear. For example, if the student doesn't check the patient's pulse at all, as opposed to checking it late, "checks pulse" would be inserted in the *Primary Survey/Circulation* section of the debriefing narrative and tagged as a "missing assessment step," with a red X (e.g., Table 1, line 17). Missing interventions are likewise identified and inserted into the annotated debriefing based on the solutions specified for their management goal in the MH and relative to where they best fit in the student's solution (the EL).

4.2 Adapting the Log Analysis System to Vary the Timing of Feedback

Most complex tasks can be divided into subtasks, each with a goal, possible subgoals, and actions to achieve these (sub)goals. For example, the assessment hierarchy shown in Figure 2 (left) represents the goal structure for managing a trauma patient during prehospital emergency care. Its subtasks include performing a *Primary Survey*, *History Taking*, *Secondary Survey*, etc., referred to herein as "assessment phases." Phases can be subdivided into subphases and their associated actions—for example, *Check Breathing* is a subphase of the *Primary Survey*, whose leaf node actions include checking the patient's breathing rate, rhythm, and quality (Figure 2, center).

This hierarchical goal structure can be used to specify when an automated simulation system should initiate within-scenario debriefings. Transitions between scenario phases offer natural assessment "checkpoints"—for example, when the student shifts focus away from the *Primary Survey*. As illustrated in the examples discussed previously (Tables 1-2), students often leave a phase prematurely, before conducting all necessary assessment actions and interventions. MedDbriefer's log analysis system can detect phase transitions because each assessment action is associated with only one phase or subphase. For example, if the student had been performing actions that belong to the *Primary Survey*'s branch of the AH and then performs actions that belong to the *Secondary Survey*'s branch (e.g., checking the patient's abdomen for distension), the system can infer that the student considers the *Primary Survey* completed (although it may not be), analyze the event log segment that includes *Primary Survey* actions, and initiate a within-scenario debriefing if it detects errors.

MedDbriefer's log analysis system could be modified to enable it to dynamically apply similar rules and procedures to the log segment that corresponds to a phase that the student has shifted focus away from (perhaps prematurely) as the system currently uses to retrospectively analyze the entire log, as described in Section 4.1. Specifically, the system would compare the student's plan for the corresponding (sub)phase with an expert plan for that (sub)phase, searching for missing actions, temporal constraint violations, and incorrectly performed interventions (e.g., see Table 1). User testing will enable us to determine what adjustments to the current log analysis procedure, additional decision rules, etc. are needed to generate within-scenario debriefings.

In future randomized trials, we plan to compare different ways of timing feedback delivery in MedDbriefer, keeping feedback content constant across conditions. At this writing, we are examining our corpus of human facilitated EMS scenarios to gain insight into the types of errors that trigger within-scenario debriefings and how experienced simulation facilitators parcel feedback between these discussions and post-scenario debriefings. To our knowledge, these questions have not yet been investigated. Initial corpus analysis suggests that facilitators tend to keep within-scenario pauses brief, focused on keeping students on track towards a successful solution (e.g., Table 2). For example, instead of addressing every assessment action that a team leader skipped, human facilitators tend to focus on those that would have yielded significant findings in the current scenario (e.g., checking the patient's pulse would have indicated an abnormal heart rate). Similarly, facilitators focus feedback on missing critical, life-saving interventions, not less critical interventions (e.g., dressing minor wounds). We plan to emulate this “focused feedback” behavior in MedDbriefer. Ultimately, a more adaptive system will assess students' performance and gradually address less critical errors after the student consistently and correctly performs critical interventions.

To keep within-scenario debriefings as brief as possible and encourage student self-regulation, we plan to have the system issue feedback in increasingly directive doses. For example, if the log analysis system detects that a student left the *Primary Survey* prematurely—such as advancing to the *Secondary Survey* without first managing the patient's difficulty breathing—the observer would be cued to click on the Feedback button. This would invoke audio-recorded feedback, to preserve the peer observer's role as a simulation assistant, not an instructor. Initial feedback would offer a high-level reminder of what the *Primary Survey* entails, for example: “Make sure that you have fully assessed the patient's airway, breathing, and circulation and managed any problems before you move onto the focused, Secondary Survey.” If the student then checks breathing but does not oxygenate and ventilate the patient, the next level of feedback would prompt the student to interpret his findings, as the facilitator does in the example shown in Table 2 (lines 1 and 3), for example: “What do you know about this patient?” If the student still does not manage the patient's breathing, the system would interpret significant findings and tell the student what to do but not how to do it, for example: “You checked the patient's breathing and found it to be slow, with gurgling. You should manage his breathing.” Finally, if the student still does not ventilate and oxygenate the patient, feedback would specify what the student should do, for example: “The patient's bradypnea [slow breath rate] and gurgling should have prompted you to administer oxygen and perform bag-valve mask ventilation, or another type of ventilatory support. Perform BVM ventilation with high flow oxygen at this time.”

In a randomized trial to compare the effectiveness of within-scenario debriefing, post-scenario debriefing, and distributed debriefing, the within-scenario condition would present increasingly directive feedback, as illustrated in the previous paragraph. Some within-scenario debriefings might elaborate on this basic feedback. Referring to the previous example, a possible elaboration would be: “Always check BVM compliance. If relevant vital signs do not improve or ventilations do not seem to be working, reassess the patient's head position (nose tilted upright), body position (thorax elevated) and ensure a proper BVM seal.” The distributed debriefing condition would provide

the same increasingly directive feedback as within-scenario debriefings but defer any elaboration until the post-scenario debriefing. The post-scenario debriefing condition would summarize the information that unfolds through the within-scenario sequence and include any elaborated feedback, entirely after the scenario ends.

5 Conclusion

The debate concerning the timing of feedback delivery during simulation-based training echoes the well-documented, unresolved problem in instructional science known as the assistance dilemma: "...how, when, and if support should be given to learners during training" [8, p. 442]. Most research on feedback timing during simulation focuses on technical skills training (i.e., procedural, motor, and psychomotor skills) [e.g., 3, 4, 8, 21]. Studies typically compare the effectiveness of *concurrent feedback* versus *terminal feedback* for initial skill acquisition and retention. Concurrent feedback takes place while a task or action is in progress, while terminal feedback takes place after a task/action, with variable delay. Due to mixed results, these studies have yielded few guidelines on timing feedback delivery for technical skills training [4]. Nonetheless, they are important for having uncovered factors that can moderate the effectiveness of feedback delivery, such as skill level, type of feedback (e.g., outcome vs. process), training context, and outcome measures (e.g., user acceptance vs. performance gains).

In contrast to the abundant research on feedback for technical skills training, few studies have investigated feedback timing in scenario-based training [9]. Scenarios challenge students to apply technical and non-technical skills (e.g., clinical decision making and team coordination). Echoing the distinction between concurrent and terminal feedback, two studies compared within-scenario with post-scenario debriefings during human guided clinical simulations [7, 9]. Their primary aim was to test the common belief that students would find within-scenario feedback disruptive. Both studies found no between-condition differences for user satisfaction, as measured by survey. Unfortunately, due to methodological limitations that the study authors acknowledge (e.g., lack of an objective measure of learning in [7]; lack of a baseline measure of performance in [9]), these studies do not indicate whether within-scenario or post-scenario debriefings are more effective for improving students' clinical knowledge and scenario performance, regarding technical and non-technical skills.

To our knowledge, no research to date has examined the effectiveness of distributing feedback across within-scenario and post-scenario debriefings, which we observed to be the norm during human facilitated EMS training scenarios (e.g., Table 2). Future research is needed to compare alternative ways of timing feedback delivery during scenario-based training in various domains—that is, within-scenario, post-scenario, and/or distributed debriefings. These studies investigate gains in students' knowledge, technical and non-technical skills, as well as user satisfaction. We expect that findings will be moderated by the same factors that moderate the effectiveness of feedback timing for technical skills training. This paper and our prior work [e.g., 4, 10] illustrate how computer-based simulation systems like MedDbriefer can serve as a research platform to investigate feedback delivery and other aspects of simulation-based training.

Acknowledgements. This research is supported by grant 2016018 from the National Science Foundation. The ideas and opinions expressed are those of the authors and do not necessarily represent the views of the NSF. We thank Audrey Aronis, John Gallagher, Priya Gupta, Karen Kornblum, Emily Miller, Lily Nong, Collin O'Connor, Erin O'Meara, Thomas Platt, Stuart Prunty, Samuel Seitz, Emma Sennott, Keith Singleton, Zachary Smith, Marideth Tokarsky, Jingyi Xiong, and Tiffany Yang for their contributions.

References

1. Cook, D.A., et al., Technology-enhanced simulation for health professions education: a systematic review and meta-analysis. *Jama*, 2011. **306**(9): p. 978-988.
2. Cheng, A., et al., Debriefing: the state of the art and science in healthcare simulation. *Healthcare Simulation Education: Evidence, Theory and Practice*, 2017: p. 158-164.
3. Hatala, R., et al., Feedback for simulation-based procedural skills training: a meta-analysis and critical narrative synthesis. *Advances in Health Sciences Education*, 2014. **19**(2): p. 251-272.
4. Wijewickrema, S., et al., Feedback Techniques in Computer-Based Simulation Training: A Survey. *arXiv preprint arXiv:1705.04683*, 2017.
5. Oermann, M.H., et al., Advantages and barriers to use of HeartCode BLS with voice advisory manikins for teaching nursing students. *International journal of nursing education scholarship*, 2010. **7**(1).
6. Oermann, M.H., et al., HeartCode™ BLS with voice assisted manikin for teaching nursing students: Preliminary results. *Nursing Education Perspectives*, 2010. **31**(5): p. 303-308.
7. Van Heukelom, J.N., T. Begaz, and R. Treat, Comparison of postsimulation debriefing versus in-simulation debriefing in medical simulation. *Simulation in Healthcare*, 2010. **5**(2): p. 91-97.
8. Tullis, J.G., R.L. Goldstone, and A.J. Hanson, Scheduling scaffolding: The extent and arrangement of assistance during training impacts test performance. *Journal of motor behavior*, 2015. **47**(5): p. 442-452.
9. Schober, P., et al., Effects of post-scenario debriefing versus stop-and-go debriefing in medical simulation training on skill acquisition and learning experience: a randomized controlled trial. *BMC medical education*, 2019. **19**(1): p. 1-7.
10. Katz, S., et al. MedDbriefer: A debriefing research platform and tool to support peer-led simulation-based training in healthcare. in *Proceedings of the 17th International Conference of the Learning Sciences—ICLS*. 2023.
11. Katz, S., et al. Debriefings on Prehospital Care Scenarios in MedDbriefer—A Tool to Support Peer Learning. in *Novel & Intelligent Digital Systems Conferences*. 2023. Springer.
12. Katz, S., et al. Comparing Alternative Approaches to Debriefing in a Tool to Support Peer-Led Simulation-Based Training. in *International Conference on Intelligent Tutoring Systems*. 2022. Springer.
13. McKenna, K.D., et al., Simulation use in paramedic education research (SUPER): a descriptive study. *Prehospital Emergency Care*, 2015. **19**(3): p. 432-440.
14. Haraldseid, C. and K. Aase, Variability among groups of nursing students' utilization of a technological learning tool for clinical skills training: An observational study. *Journal of Nursing Education and Practice*, 2017. **7**(7): p. 66-76.
15. Rencic, J., et al., Clinical reasoning education at US medical schools: results from a national survey of internal medicine clerkship directors. *Journal of general internal medicine*, 2017. **32**: p. 1242-1246.

12 S. Katz, P. Albacete, P. Jordan, S. Silliman, M. Wrzesniewski

16. Koedinger, K.R. and A. Corbett, Cognitive Tutors: Technology bringing learning science to the classroom. 1st ed. The Cambridge Handbook of the Learning Sciences, ed. R.K. Sawyer. 2008, New York, NY, USA: Cambridge University Press. 61-77.
17. Chu, Y.-S., et al., Implementation of a model-tracing-based learning diagnosis system to promote elementary students' learning in mathematics. *Journal of Educational Technology & Society*, 2014. **17**(2): p. 347-357.
18. Carberry, S., Techniques for plan recognition. *User modeling and user-adapted interaction*, 2001. **11**: p. 31-48.
19. Allen, J.F., Towards a general theory of action and time. *Artificial intelligence*, 1984. **23**(2): p. 123-154.
20. Köckemann, U. and L. Karlsson. Configuration planning with temporal constraints. in *Proceedings of the AAAI Conference on Artificial Intelligence*. 2017.
21. Cheng, A., et al., Debriefing for technology-enhanced simulation: a systematic review and meta-analysis. *Medical Education*, 2014. **48**(7): p. 657-666.