Teaching First or Teaching Last: Does the Timing Matter in Simulation-Based Surgical Scenarios?

Benjamin Zendejas, MD,*† David A. Cook, MD,†‡§ and David R. Farley, MD*†

*Department of Surgery, †Mayo Clinic Multidisciplinary Simulation Center, ‡Department of Medicine, and the §Office of Education Research, College of Medicine, Mayo Clinic, Rochester, Minnesota

OBJECTIVE: The optimal timing of instruction in simulation-based scenarios remains unclear. We sought to determine how varying the timing of instruction, either before (teaching first) or after (teaching last) the simulation, affects knowledge outcomes of surgical trainees.

DESIGN: We conducted a pretest/posttest crossover study in which fourth-year medical students and general surgery residents (PGY 1–3) participated in 3 instructional modules, each repeated twice in consecutive weeks: endocrine surgery (sessions 1 and 2), trauma resuscitation (sessions 3 and 4), and team training (sessions 5 and 6). Each session comprised 3 cases, each involving a prescenario briefing, a simulated scenario, and a postscenario debriefing. The timing of instruction varied between sessions. For the teaching-first sequence (sessions 1, 4, and 6), participants received a lecture during each prescenario briefing. In the teaching-last sequence (sessions 2, 3, and 5), trainees received an identical lecture during the postscenario debriefings. We assessed attitudes and knowledge using a postsession survey and identical 10-question multiple-choice tests at the start (pretest) and end (posttest) of each session, respectively. The mean differences in knowledge scores between groups were analyzed with repeated-measures analysis of variance (ANOVA).

RESULTS: Forty-nine participants (11 medical students and 38 surgical residents) attended at least 1 session, providing 76 observations. Mean pretest scores were equivalent ($p > 0.05$). The change in scores from pretest to posttest varied between the 2 groups ($p = 0.002$). The mean posttest score was 8.24 (standard error [SE], 0.29) for the teaching-last group and 6.68 (SE, 0.27) for the teaching-first group (mean difference, 1.56; 95% confidence interval, 0.79–2.33). Both teaching-first and teaching-last group participants preferentially rated debriefings and scenarios, respectively, as the better learning experience.

CONCLUSIONS: Participants who received instruction after simulated scenarios achieved higher mean knowledge scores than those who received instruction before simulated scenarios. Cognitive overload, stress, or activation of prior knowledge could all be involved as causal mechanisms. (J Surg 67:432-438. © 2010 Association of Program Directors in Surgery. Published by Elsevier Inc. All rights reserved.)

KEY WORDS: medical education, patient simulation, educational models, instructional design

COMPETENCIES: Medical Knowledge, Practice-Based Learning and Improvement, and Interpersonal and Communication Skills

INTRODUCTION

Simulation-based medical education (SBME) is here to stay; it has experienced an exponential growth in the last few decades and has become a centerpiece in the education of health professionals. In surgery, the evidence supporting the benefits of simulation for surgical skills training continues to accumulate and provide arguments in favor of the transfer of skills learned in the simulated environment to the clinical setting. In addition, a greater emphasis on patient safety and quality of medical education has led regulatory bodies to take a strong stand in favor of SBME; this issue is exemplified clearly by the mandate of the Residency Review Committee (RRC) in Surgery stipulating that every general surgery residency program have access to a simulation or skills center and have a formal skills curriculum in place for accreditation. Despite this simulation momentum, much remains to be learned about how to design and deliver effective simulation-based training.

The role of simulated scenarios in SBME stems from extensive experience within the military context, in which task demands are similar to those confronted by medical professionals. The goal of scenario-based training is to develop realistic and
instructionally sound scenarios that include appropriate measurement and feedback strategies. 7 Even though the value of debriefing (as a feedback strategy) has not been proven entirely objectively, 8 it has shown to improve nontechnical skills after simulation-based training 9,10 and is generally recognized as “crucial to the learning process in SBME.” 8-12 Thus, most simulated scenarios are preceded by an introduction or briefing process and are followed by a review or debrief (brief-scenario-debrief format). Most programs also include an element of formal instruction. However, the optimal timing of instruction in this sequence remains unclear.

Many theories of learning suggest that learning is promoted when relevant previous experience is activated. 13 Participating in a simulated scenario before receiving didactic information could activate prior knowledge and set a foundation for new knowledge. 14 Thus, the rationale for providing instruction or didactic information after the scenario is based on the belief that the scenario experience and the debrief session renders the learner to be more receptive to new information once he or she has identified a knowledge gap. 8 By contrast, other arguments and evidence suggest that providing didactic information before a scenario can also facilitate learning by reducing anxiety and/or decreasing cognitive load. For example, pretraining in computer-assisted instruction has been demonstrated to reduce cognitive load and improve learning. 15 Yet again, some evidence would suggest that the sequence of instruction makes no difference at all. For instance, in the field of web-based learning, Cook et al. 16 demonstrated that the timing of case-based problems, either before or after instruction, did not appear to alter learning outcomes.

We found no studies investigating this question in the simulation literature. To address this knowledge gap, we sought to determine how the timing of instruction, either before (teaching first) or after (teaching last) the simulated scenario, affects knowledge outcomes among surgical trainees.

Although we view 3 possible outcomes (improvement with teaching first, improvement with teaching last, or no difference), we feel that the theoretical framework behind the activation of prior knowledge through engaging the learner in a “mind primer” (simulated scenario) has a stronger foundation rooted in theories of cognitive psychology. Thus, we predicted that teaching after the simulated scenario would enhance learners’ knowledge outcomes.

**METHODS**

**Participants and Design**

With prior Institutional Review Board approval, we conducted a pretest/posttest crossover study in our multidisciplinary simulation center (level I American College of Surgeons accredited education institute) in the fall of 2009. Participants included categorical and preliminary general surgery residents (PGY 1-3), and fourth-year medical students who were currently in a surgical elective.

Participants completed simulation training using 2 different formats, with either didactic lecture before simulated scenario (teaching-first sequence) or lecture after the scenario (teaching-last sequence). The format was determined by the date of the training session (nonrandomized crossover design) prespecified within the weekly scheduled surgical educational curriculum at our institution.

**Interventions**

We developed 3 instructional modules to be delivered in the form of scenario-based simulation, in which scenarios are built around full-scale high-fidelity mannequins to simulate complex medical/surgical crises. 17 Each module was repeated as 2 identical 3-hour sessions (totaling 6 sessions) in consecutive weeks to accommodate residents’ schedules (Table 1). The topics of the modules were endocrine surgery (sessions 1 and 2), trauma resuscitation (sessions 3 and 4), and team training (sessions 5 and 6). Each session consisted of 3 cases related to the modules’ topic, each involving a prescenario briefing, a simulated scenario, and a postscenario debriefing. For each case, participants received a brief didactic lecture on relevant key concepts. However, we varied systematically the timing of the lectures to be delivered. For the teaching-first sequence (sessions 1, 4, and 6), participants received the lecture before each simulated scenario (ie, during the prebrief). In contrast, in the teaching-last sequence (sessions 2, 3, and 5) trainees received the identical

---

**TABLE 1.** Instructional Modules with Corresponding Cases and Sessions

<table>
<thead>
<tr>
<th>Instructional Modules*</th>
<th>Cases</th>
<th>Sessions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Endocrine surgery</td>
<td>1. Thyroidectomy for thyroid cancer</td>
<td>1 and 2</td>
</tr>
<tr>
<td></td>
<td>2. Adrenergic crisis during phaeochromocytoma resection</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3. Neck-incision closure and postoperative neck hematoma</td>
<td></td>
</tr>
<tr>
<td>Trauma resuscitation</td>
<td>1. Chest trauma with hemopneumothorax</td>
<td>3 and 4</td>
</tr>
<tr>
<td></td>
<td>2. Pelvic fracture with hypovolemic shock</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3. Head trauma with epidural hematoma</td>
<td></td>
</tr>
<tr>
<td>Team training</td>
<td>1. Intraoperative bleeding with communication breakdown</td>
<td>5 and 6</td>
</tr>
<tr>
<td></td>
<td>2. Aortic cross-clamping lacking situation awareness</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3. Postoperative asystole without team leader</td>
<td></td>
</tr>
</tbody>
</table>

*Each module was repeated twice in consecutive sessions to accommodate participants’ clinical schedule.
lecture after each simulated scenario (ie, during the debrief). All 3 cases in a given session used the same training sequence (Fig. 1). The session components (prebrief, scenario, and debrief) flowed from 1 to another without waiting time in between. The lectures were scripted to ensure consistency across sessions, with the same instructor for each pair of sessions.

Assignment of Interventions

As part of their weekly educational activities, trainees participated voluntarily and came for 1 session per topic as their schedules allowed. We did not assign individuals to specific sessions, but by considering residents’ schedules and repeating each intervention twice, we provided equal opportunity for trainees to participate in both training formats. Participants were masked to the study hypothesis (until attendance at the session) and to the sequence of instruction.

Measurements

We assessed knowledge of key concepts related to the session topic using identical 10-question multiple-choice tests at the start of the session (pretest) and after all 3 scenarios at the end of the given session (posttest). We planned the questions to represent evenly the content of the 3 didactic lectures in each session. Cronbach’s alpha was used to evaluate the internal consistency of each test. After the posttest, study participants were also asked to complete a brief attitude survey that required them to point out which of the 3 session components (prescenario briefings, scenarios, or debriefings) made them feel most nervous and which one they valued as the better learning experience.

Data Analysis

Descriptive statistics are provided as counts (percentage) and mean (standard error or standard deviation) as appropriate. Baseline (pretest) scores were compared using the Student t test. We compared the 2 interventions (instruction sequences) using an analysis of variance with repeated measures on participants and module topic. In this analysis, a significant intervention by time interaction would indicate a significant difference in score change between the 2 groups. To assess the possibility of confounding caused by recency of learning related to the proximity of the posttest to the final lecture (recency effect), we conducted a sensitivity analysis that excluded the 3 questions on each test that were related to the final lecture. The proportions of responses to attitude survey were compared among groups with a χ² test. The analyses were performed using SAS (version 9.1.3; SAS Institute, Inc., Cary, North Carolina) and JMP (version 8.0; SAS Institute, Inc.). All hypothesis testing was 2-sided and the alpha level was set at 0.05 for statistical significance.

RESULTS

In all, 49 participants (11 medical students and 38 surgical residents) attended at least 1 study session and all consented to participate. Demographic information is presented in Table 2. Of the 49 study participants, 4 (8%) completed all 3 sessions, 19 (39%) completed 2 sessions, and 26 (53%) completed only 1 session, providing a total of 76 observations (42 with the teaching-first sequence and 34 using the teaching-last sequence). Of the 23 trainees who

![FIGURE 1. Instructional sequences. Participants attended sessions based on their clinical availability; each session comprised 3 cases using the same instructional sequence (prespecified depending on the calendar day).](image)

<table>
<thead>
<tr>
<th>TABLE 2. Study Participant Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>n = 49 (%)</td>
</tr>
<tr>
<td>Gender</td>
</tr>
<tr>
<td>Females</td>
</tr>
<tr>
<td>males</td>
</tr>
<tr>
<td>Training level</td>
</tr>
<tr>
<td>PGY-3</td>
</tr>
<tr>
<td>PGY-2</td>
</tr>
<tr>
<td>PGY-1</td>
</tr>
<tr>
<td>PGY-1 preliminary</td>
</tr>
<tr>
<td>Medical students*</td>
</tr>
<tr>
<td>Age</td>
</tr>
<tr>
<td>Mean (SD)</td>
</tr>
<tr>
<td>Previous simulation experience†</td>
</tr>
</tbody>
</table>

*Fourth-year medical students in a surgical elective during the study.
†Had previously participated in a scenario-based simulation format.
participated in more than 1 session, 14 participated in the same training sequence and 9 participated at least once with the alternate training sequence. We compared the students who participated only in the training-first sequence (n = 23) with those who participated only in the training-last sequence (n = 18), using the demographic variables in Table 2, and found no significant difference (p > 0.05).

The mean pretest scores for a given topic, compared between training sequences, were equivalent among participants (p > 0.05). The change in knowledge scores from pretest to posttest varied significantly between the 2 groups (pInteraction = 0.002). The mean posttest score, averaged across the 3 topics, was 8.24 [standard error of the mean (SE), 0.29] for the teaching-last group and 6.68 (SE, 0.27) for the teaching-first group (mean difference, 1.56; 95% confidence interval [CI], 0.79-2.33). The effect size (Cohen’s d) of teaching-last instruction was 0.80.

The analyses remained unchanged after adjustment for training level or gender. Sensitivity analyses excluding test questions related to the last case for each topic revealed essentially identical results (pInteraction = 0.001). The Cronbach’s alpha results for the posttests on endocrine surgery, trauma resuscitation, and team training were 0.72, 0.74, and 0.79, respectively.

In the teaching-first sequence, most study participants (66%) rated the actual scenarios as the component of each case where they felt the most nervous, followed by debriefings (19%) and prescenario briefings (14%). The component in which they learned the most followed a similar pattern, with most (64%) participants identifying scenarios as most valuable, followed by debriefings (22%) and prescenario briefings (14%). For the teaching-last sequence, the scenarios were also rated as the component in which study participants felt the most nervous by most participants (91%), followed by prescenario briefings (6%) and debriefings (3%). In contrast, most participants (53%) in the teaching-last sequence rated debriefings (which included the lectures), instead of scenarios, as the component that they considered the better learning experience, followed by scenarios (44%) and briefings (3%). The difference in proportion listing debriefing as the better learning experience was significantly higher for teaching-last than for teaching-first (53% vs 22%, p = 0.004). However, in the teaching-last sequence, significantly more participants felt the most nervous during the scenarios (91% vs 66%, p = 0.01) (Fig. 2).

**DISCUSSION**

This study examined how the timing of instruction, either before (teaching first) or after (teaching last) the simulation-based scenario, influences the knowledge outcomes of surgical trainees who are exposed to scenario-based simulation training. In accordance with our prediction, the results of our study show that trainees who participated in the teaching-last sequence achieved significantly higher mean knowledge posttest scores when compared with those who underwent the teaching-first sequence of instruction. These findings remained essentially
unchanged after adjusting for participant gender and training level, and after excluding the test questions related to the last lecture of the session to minimize the effect of recency of learning. According to Cohen’s classification of the effect size (small: 0.2–0.5, medium: 0.5–0.8, and large: >0.8),18 the educational impact of the teaching-last sequence of instruction (effect size, 0.8) was “large,” which translated to a 16% improvement in knowledge posttest scores, or 1–2 additional correct questions (in a 10-question test).

The sequence of instruction also had an impact on participants’ perceptions toward the simulation. Significantly fewer participants in the teaching-first sequence reported the actual scenarios as being the most “anxiety-provoking” component as opposed to those in the teaching-last sequence, for which most participants reported scenarios as the component with which they felt the most nervous. Contrary to participants in the teaching-last sequence, where most valued postscenario debriefings (with instruction) as the component that provided the better learning experience, most participants in the teaching-first sequence valued scenarios as the better learning experience, followed by debriefings and prescenario briefings (with instruction). These findings suggest that participants’ perception of a better learning experience was not entirely dependent on when they received the didactic lecture.

The Results of This Study Can Be Viewed Through the Lens of Three Different Conceptual Frameworks

Stress. Acute stress can be classified as eustress or “healthy stress” if an individual has personal and environmental resources that meet the perceived demands of a situation, or distress if the demands outweigh the resources of the individual.19 Although some studies have shown impaired performance under acute stress,20,21 others have shown improved performance.22 Thus, the perfect balance between “healthy” and “harmful” stress for optimal learning remains to be elucidated. Nonetheless, we do know that an individual’s response to a stressful event is highly dependent on that individual’s perception of the demands,25 and this perception leads to the activation of the hypothalamic–pituitary axis and ensuing cortisol rush, which in turn activate specific cognitive brain areas (amygdala, hippocampus, and prefrontal cortex).24 In our study, most teaching-first participants rated the actual scenarios as the element that made them most nervous (a surrogate for stress) and considered these events to be the better learning experience. By contrast, nearly all teaching-last participants rated the scenario as most stressful, and the debrief (with instruction) as the better learning experience. Because learning also differed between these groups, these findings suggest an interaction among stress, timing of instruction, and learning outcomes that merits subsequent investigation. Accordingly, Inzana et al.25 in the military context observed that presenting preparatory information during scenario-based training reduced subject anxiety and improved performance accuracy on Naval decision-making tasks.

Cognitive load theory. Novel information must be processed within working memory to construct schemas or mental models in long-term memory. Unfortunately, working memory has a finite capacity (7 plus/minus 2 chunks of data),26 which is affected by the intrinsic complexity of the learning task and by the way in which it is presented.27 In our study, we believe that participants who were exposed to a complex learning task (scenario-based simulation), which imposed a high intrinsic load on their cognitive system, were overwhelmed in their capacity to retain the new information received previously to the scenario. However, this effect lessened when this information was given afterward, when the debriefing and lecture had helped decrease the cognitive load imposed by the scenario. Bingener et al.28 found similar detrimental effects of an intervention with high intrinsic cognitive load in a study showing that novice students who were exposed to an error-recognition curriculum (high intrinsic load) designed to improve laparoscopic surgical skills performed worse than those who were not given this curriculum and only allowed to practice.

Activation of prior knowledge. Because long-term memory has no known limit, the activation of prior knowledge in the form of retrieved schemas or mental models from long-term memory can help “free up” space in working memory and allow for better integration of the new information with prior knowledge.29 Our study results suggest that participants who received the lecture after the scenario had activated their prior knowledge, had more attentional capacity, and were primed to retain better the new information provided from the postscenario lecture. In accordance to our findings, Alvermann et al.30 found that students who had activated prior knowledge outperformed students who had not in test recall of high cognitive load experiences, but they found no difference in low cognitive load situations. This finding corroborates the notion that activation of prior knowledge is an effective strategy to improve learning in contexts with high intrinsic cognitive load.27

LIMITATIONS

This study has some limitations that might affect the applicability of its results. As with any nonrandomized intervention, we could have had uneven distribution of known and unknown confounders between groups. We attempted to minimize this limitation by using a study design in which participants could cross over to the other intervention during subsequent sessions, but only 23 participants completed more than 1 session, of which only 9 participated in different training sequences. However, the demographic measures were similar among those who completed only 1 training sequence, and all analyses were adjusted for pretest scores, gender, and training level. Despite a relatively small sample size (which limited subgroup analysis), the power of the study was adequate to detect significant differences. Although it is unlikely that such differences are caused by chance, the possibility for bias remains. Bias in the sample could exist if trainees who did not attend study sessions were systematically different from those who did; we attempted to minimize
this effect by including all participants who attended regardless of training level. Bias in the measurement of outcomes could also be present, as the tests used for the purpose of pretesting and posttesting were not validated prior to their use. However, questions were constructed from information directly from the lecture scripts and were applied nondifferentially to both groups; in addition, all tests had adequate internal consistency. Study investigators were not blinded to the sequence of instruction of study participants, but lectures and sessions were scripted to ensure consistency across groups. Outcome assessment was also not conducted in a blinded fashion. Nonetheless, the answers to the multiple-choice questions were objective and prespecified, and scoring followed a predefined algorithm with overview by the senior investigator. This study was limited to cognitive learning and mostly to novice learners (PGY-1s, 55%; medical students, 22%); therefore, we do not know whether our results apply to other competencies (ie, skills) or senior learners. Finally, the study involved a single training program, and the composition of our program may not be representative of that of other programs.

Implications to Current Practice and Future Research

The findings of this study highlight the need for a thorough understanding of how instructional design principles interplay in scenario-based simulation. Providing instruction after the scenario potentially can decrease participants’ cognitive load in addition to the activation of prior knowledge that results from scenario exposure, which will enhance the integration of new information and, most likely, will result in improved knowledge outcomes. The short-term improvement observed in this study is not a guarantee of a long-term effect, and this result should be studied further. As well, our study findings might not apply to other types of instructional modes or even to other competencies, such as surgical skills, where pretraining has been shown to improve technical skills. Thus, how the timing of instruction in scenario-based simulation affects other competencies will need to be elucidated. In addition, how these results apply to other competencies (ie, skills) or senior learners. Finally, the study involved a single training program, and the composition of our program may not be representative of that of other programs.

CONCLUSIONS

Participants who received instruction in a lecture-based format after simulated scenarios achieved higher mean knowledge scores than those who received the same instruction before the simulated scenario. Cognitive load, stress, or activation of prior knowledge could all be involved as causal mechanisms behind these findings. Additional studies are needed to expand our understanding of how to deliver simulation-based educational interventions optimally.

ACKNOWLEDGMENTS

The authors would like to thank the general surgery residents and medical students who participated in this study, as well as Mark D. Sawyer and Jackie Arnold, for their help as instructors, and the rest of the Mayo Clinic Multidisciplinary Center Staff who made this study possible.

REFERENCES


