# Emergent Themes for Instructional Design

## Alpha and Beta Testing During a Faculty Development Course



Jannet J. Lee-Jayaram, MD;
Benjamin W. Berg, MD;
Angela Sy, DrPH;
Kristine M. Hara, BAS

**Introduction:** Instructional design is an established discipline for designing educational activities for learners and is applied during the development of simulation-based healthcare educational activities. Although the iterative process is already being used with alpha/beta testing during development of the simulation, the process has not been described in detail. We sought to describe this process of design changes made during a novice faculty development course for simulation-based healthcare education where participants routinely design scenarios and conduct alpha/beta testing.
**Methods:** Using a mixed methods study, participant written narratives and checklist/rating scales were collected on changes made during both alpha and beta testing. Narratives were analyzed using the qualitative grounded theory approach to identify emergent themes. Checklist/rating scales were analyzed for changes made to previously identified areas and how critical these changes were to the success of the scenario.
**Results:** Several themes of frequent changes made during the alpha/beta testing process emerged from the participant narratives that included REALISM, PROTOCOLS, INTRODUCTIONS and ROLES. The quantitative analysis of potential areas for change correlated with the emergent themes.
**Conclusions:** Novice scenario designers make similar changes in thematic areas during the instructional design process. Faculty development courses for novice simulation healthcare educators should anticipate attention to these areas during discussions and consider specific didactics to avoid common design pitfalls of novice educators.
(*Sim Healthcare* 14:43–50, 2019)

**Key Words:** Alpha testing, beta testing, instructional design, faculty development, simulation based healthcare education, grounded theory, pilot testing.

Alpha and beta testing are software development strategies commonly used in a variety of industries to guide design modification before final public release of an optimized product. These transdisciplinary terms are also descriptors of elements in the healthcare education instructional design process for simulation-based scenarios and simulators.[1,2] Alpha testing is heuristic user testing of a product by development team members during early development phases to inform and guide initial software product trouble shooting. Alpha testing serves as a form of internal acceptance testing.[3] Beta testing of preproduction products is conducted by a representative target user cohort, not including members of the product development team. Beta testing is a form of external acceptance testing and provides feedback to the development team about a limited release product.[3] Both alpha and beta testing guide product design modification before final public release. Instructional design for simulation in healthcare commonly employs a similar iterative process, inconsistently identified as alpha or beta testing. Educational design research employs iterative cycles of development, implementation, and analysis to inform subsequent curriculum revision and refinement to optimally meet learning goals and objectives.[4]

Instructional design is an established discipline, based on sound educational theories and rubrics including the ADDIE (Analysis, Design, Develop, Implement, Evaluate) model for development[5] and the ARCS (Attention, Relevance, Confidence, Satisfaction) model of motivational design.[6] Other fundamental instructional design theories and strategies inform the design process and support the basis for simulation-based educational experiences. These include Gagne and Merrill's learning enterprises in which learners are engaged in a comprehensive activity where multiple objectives are integrated,[7] as well as Dick and Carey's systems approach model that focused on goals and objectives and addressed the interrelation between all the components of instruction.[8] However, specific instructional design for simulation-based healthcare education is less well developed than for other educational domains and lacks an evidence-based framework or well-articulated domain-specific applicable design theory. Despite this, general educational goal and objective oriented theory for curricular design, including adult learning principles, do apply.[9] Frameworks for instructional design of simulation-based healthcare education have been proposed and consist of scenario scripting tools, blueprints, and templates, without explicit internal or pilot-testing guidelines.[10,11] Published guidelines for simulation design

From the University of Hawaii (J.J.L., B.W.B., A.S., K.M.H); and SimTiki Simulation Center (J.J.L., B.W.B., K.M.H.), Telehealth Research Institute, John A. Burns School of Medicine, University of Hawaii at Manoa, Honolulu, HI.

Reprints: Jannet J. Lee-Jayaram, MD, 651 Ilalo St, MEB 212 John A Burns School of Medicine Honolulu, HI 96813 (e-mail: jannet98@hawaii.edu).

Supported in part by a novice investigator grant from the Society for Simulation in Healthcare.

The authors declare no conflict of interest.


DOI: 10.1097/SIH.0000000000000329

include recommendations for pilot testing simulation-based experiences with colleagues and members of the simulation team, as well as with an audience similar to the target learner group.[12,13] The iterative process is widely used but has not been systematically detailed or rigorously analyzed. For example, Society for Simulation in Healthcare Certified Healthcare Simulation Educator program identifies scenario development as a key educator skill, specifying that applicants "design simulation activity" and "conduct pilot activity for new simulations (eg, dress rehearsal, field test, run through)."[14] These models are consistent with the process used in software alpha and beta testing to guide product design modification but have not been explicitly described or named as such. We are unaware of any published empirical studies based on data gathered during the scenario design process for simulation-based healthcare education.

Our aim was to prospectively gather data during the process of scenario design exercises within a healthcare simulation faculty development course, specifically detailing design modifications occurring during alpha or beta testing. We anticipated that common novice scenario designers' modifications/issues could be codified into themes, that these themes could be incorporated back into the instructional design of the faculty development course, and that this could aid novice designers to avoid common pitfalls early in the iterative design process.

## METHODS

This study was approved by the University of Hawai'i Human Studies Program as exempt (CHS #21642).

We used a mixed methods research approach in which researchers collect, analyze, and integrate quantitative and qualitative data in a single study to address their research questions.[15] The quantitative data provided explanations for the relationships among study variables but more detailed understanding of what the statistical tests or effect sizes actually mean was also needed. The qualitative data and results helped build that understanding.[15]

A 2-day multimodal faculty development course for novice simulation healthcare educators from diverse healthcare disciplines, Fundamental Simulation Instructional Methods (FunSim), has been conducted since 2011 by the SimTiki Simulation Center, John A Burns School of Medicine, University of Hawaii, enrolling more than 400 healthcare educator participants. The participants come from varied healthcare backgrounds, including clinical nursing, nursing faculty, hospital educators, clinical physicians, medical school faculty, community healthcare providers, and others. FunSim is a workshop style program during, which two independently working small groups of 2–6 members collaboratively design healthcare simulation scenarios, using iterative heuristic analysis with formal alpha and beta testing. Imposed scenario design constraints include time limits and simulator type including standardized patient, mannequin, task-trainer, and hybrid-based scenarios. Workshop participants are actively mentored by course faculty, using a scenario design template to establish scenario objectives, methods, and an assessment strategy, for a 7-minute or less healthcare focused simulation scenario. Learner scenario design modification(s) are explicitly identified and incorporated into scenario designs after sequential alpha- and beta-testing sessions, which are followed immediately by a 20- to 30-minute facilitated debriefing of scenario design using a GAS (Gather-Analyze-Summarize) model debriefing structure,[16] focusing on identification of optimal scenario design modifications. Beta-test learners are volunteer course participants from learner teams who are unaware of details and development of the test scenario; as such, beta-test learners are not members of the intended learner cohort for the instructional design test scenario. Beta testing is therefore completed with a nonrepresentative target learner cohort—a simulated beta-test cohort. Simulated beta testing achieves one primary purpose of actual beta testing: to conduct scenario design evaluation with a scenario-naive learner cohort.

Refer to Figure 1 for the study flowchart. Data for this study were collected from a convenience sample of participants attending nine monthly FunSim courses from April to December of 2014. Each of the nine courses was attended by 4 to 12 participants. During each course, participants completed an exercise in two separate 2- to 6-member convenience groups that collaboratively designed a simulation scenario. After working in these small groups aided by an internal scenario design template to design a simple simulated healthcare scenario, participants then alpha tested their scenarios. The small-group participants collaboratively documented all changes made with alpha testing as a freeform written narrative. After a short didactic and activity on facilitation, the small groups beta tested their scenarios with the other small-group participants as beta-test subjects. Each beta test was immediately followed by an instructor-led reflection/discussion session regarding successful scenario design. The small-group participants collaboratively documented all changes they would make to improve their scenarios based on the beta test as a freeform narrative. The groups then completed checklist/rating scales for changes made with both alpha and beta testing. The two data collection instruments were: (1) a request for written freeform narrative report regarding scenario design modifications and (2) a checklist/rating scale request to report yes/no to any changes made, and if yes, the degree of importance to scenario design regarding changes in a-priori investigator identified themes (Appendix 1). Investigator-defined thematic categories for the checklist/rating scale were the following: *Orientation*, *Type of Simulator*, *Equipment/Supplies*, *Scenario Clinical Factors*, *Teaching Team Constructs*, and *Timing*. To decrease the risk of bias and to allow for emergence of previously unidentified items from the novice designer participants' perspectives, the written freeform narrative was completed for both alpha and beta testing before the checklist/rating scale. The six checklist/rating scale categories were empirically defined by the investigators, who were the course directors (B.W.B./J.L.) and the chief simulation specialist (K.M.H.) for the FunSim courses conducted in the 3 years before this research. Based on observations of participant scenario design alpha- and beta-testing exercises during those previous courses and using our internal scenario design template as a reference, we recollected patterns of changes by novice designers and classified them into defined categories.

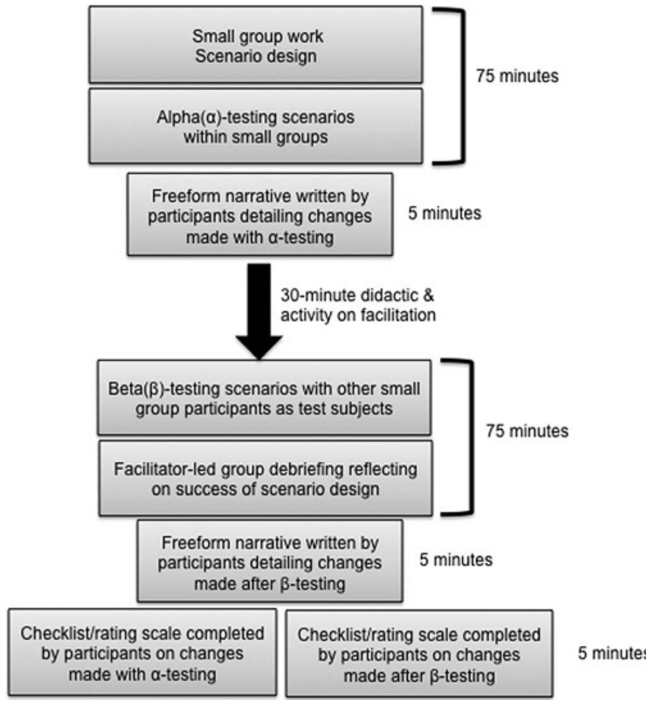


**FIGURE 1.** Study flowchart. Visual timeline for group work and collection of qualitative and quantitative data.

Lacking previously published results to inform a quantitative or qualitative estimate of thematic results, this was an exploratory descriptive group, with no applicable sample size estimate or power analysis.

### Qualitative Analytic Approach

Grounded theory was used to analyze qualitative data. Grounded theory is an inductive qualitative research method, seeking to gather data for systematic development of a theory derived from the data. This approach contrasts with traditional deductive research methods seeking evidence to verify or refute an explicit theory or hypothesis.[17] Grounded theory research begins without knowledge or presumption of issues, basic concepts, or frameworks that may be relevant to the research topic. Qualitative data collection identifies emergent themes, whereas a constant comparative method is applied. Themes with similarities are grouped into codes and codes are then reorganized, refined, and connected.[18,19]

A program evaluator (A.S.) with expertise in qualitative research methods, but without previous involvement in simulation education or in the design of the study, used an inductive approach to identify and code themes. This qualitative analyst had limited knowledge regarding simulation education learning concepts and therefore had no preconceived theories about expected emergent coding themes, enhancing the validity of the qualitative analysis. After orientation to the context of the course, the qualitative analysis expert reviewed the open-ended written narrative regarding feedback from each team and created a codebook of themes regarding scenario design changes based on alpha and beta testing. This lack of expertise in simulation design and education also meant that the instructor designated alpha- and beta-testing categories did not impact the coding process, further fostering its inductive characteristic. Identified scenario design changes were coded without restriction, into as many themes as relevant. Expert simulation designers (J.L./B.W.B.) completed the secondary analysis of raw coded themes and refined and described themes to align with language and concepts of healthcare simulation scenario design.

### Quantitative Analytic Approach

Data from the checklist/rating scale were analyzed using SPSS (IBM SPSS Statistics for Windows Version 22.0, 2013, IBM Corp). Differences in the frequency of modifications (no change/yes change) in the predetermined thematic categories in both alpha and beta tests were analyzed by $\chi^2$. In categories where changes were made, descriptive statistics were reported on the 5-point Likert scale rating of criticality to scenario success.

## RESULTS

### Demographics

Refer to Table 1 for participant demographics. Nine FunSim courses were studied with a total of 65 participants from mixed professional backgrounds. Fifty-one percent of the participants were from the nursing profession (n = 33) and 34% were physicians (n = 22). Physician specialties/subspecialties included pediatrics (n = 3), internal medicine (n = 3), pathology (n = 3), family medicine (n = 2), surgery (n = 2), intensive care (n = 2), neurology, otolaryngology, pediatric emergency medicine, ophthalmology, neonatology, and emergency medicine (remainder n = 1). There was a small cohort of other professions including athletic trainers (n = 5) and one each simulation specialist, paramedic, fire fighter, social worker, and educational programs manager. Forty-four (68%) of participants were female, with a mean age of 45 years. Mean years in practice were 17.5 years. The group was novice to simulation with a mean of 2.2 years of experience in simulation and 33 (51%) participants reporting 1 year or less experience. A total of 17 scenario design teams were composed of

**TABLE 1.** Participant Demographics

| Professions, No. Participants (%) | 65 |
|---|---|
| Nursing | 33 (51%) |
| Registered nurse | 27 |
| Nurse practitioner | 4 |
| Certified registered nurse anesthetist | 2 |
| Physicians | 22 (34%) |
| Pediatrics | 3 |
| Internal medicine | 3 |
| Pathology | 3 |
| Family medicine | 2 |
| Surgery | 2 |
| Intensive care | 2 |
| Neurology | 1 |
| Otolaryngology | 1 |
| Pediatric emergency medicine | 1 |
| Ophthalmology | 1 |
| Athletic trainers | 5 (7%) |
| Other | 5 (7%) |
| Simulation specialist | 1 |
| Paramedic | 1 |
| Social worker | 1 |
| Fire fighter | 1 |
| Manger | 1 |
| Sex, no. participants (%) | |
| Female | 44 (68%) |
| Male | 21 (32%) |
| Mean age in years (range) SD | 45 (22–72) SD = 10.9 |
| Mean years in clinical practice (range) SD | 17.5 (0–41) SD = 11.1 |
| Mean years in simulation (range) SD | 2.2 (0–10) SD = 2.3 |

Background of participant professions, sex, age, years in practice, and years in simulation.



two to six participants each. Seventeen participant-designed scenarios were available for analysis. No data, participants, workgroups, or scenarios from any course were excluded.

### Qualitative Analysis/Grounded Theory Emergent Themes

The number of scenario design modifications identified on written open-ended responses ranged from 0 to 7 per scenario. Fifty-five modifications were written following alpha testing, and 59 modifications were noted after beta testing. A mean of three modifications were made per alpha test and 3.8 per beta test.

Qualitative analysis extracted and coded a total of 75 participant entries. There were four emergent themes, three of which had subthemes. Of most to least frequently identified themes, REALISM had four subthemes. PROTOCOLS had five subthemes. INTRODUCTIONS had three subthemes. Finally, ROLES had no subthemes (refer to Table 2: Qualitative Data Emergent Themes and Frequencies).

The theme of REALISM comprised the greatest number of scenario design modifications overall (n = 33), with the preponderance of changes occurring with alpha testing (n = 21), and fewer with beta testing (n = 12). Realism modifications addressed increasing or improving realism with respect to the subthemes of Props, Patient, Timing, and Clinician. Specific examples of written responses included “Need to have phone available,” “Reprogram… Actual lung sounds,” and “Gradual change of VS” [verbatim participant response; VS assumed to represent vital signs].

The second greatest number of modifications was reported in the PROTOCOLS theme (n = 22), predominantly after beta testing (n = 15). PROTOCOLS theme modifications addressed improving Clinical issues during the simulation, modifying Sequence, providing learners additional Information, and issues with Student(s) and Timing. Examples of written responses included “Allow VS to be taken before seizure,” “Measure… only 1 student,” and “Changed the timing from 2 minutes to 1 minute.” The subtheme of Timing was the only subtheme that overlapped more than one main theme, appearing under both REALISM and PROTOCOLS.

The theme of INTRODUCTIONS, with 19 scenario design modifications, was initially coded as a subtheme under PROTOCOLS; however, it emerged with enough breadth and depth to generate its own subthemes and therefore was recorded as a main thematic category. Most modifications in introductions occurred with beta testing (n = 12). Modifications coded in the theme of introductions addressed changing or clarifying Objectives, Specific individual changes, or just a General mention of change without specific details. Examples included “Give student broadened objectives so intervention isn't obvious,” “Orient the student to know exactly what the seizure/motion/sounds are,” and “Modified our intro.”

The least frequently modified scenario thematic area was ROLES (n = 11), which did not have any emergent subthemes. Scenario changes after alpha (n = 6) and beta (n = 5) testing were evenly distributed. Changes within this theme dealt with improving, increasing, or clarifying roles during simulation. Specific examples included “no confederate,” “change in personnel,” and “clarify roles more.”

### Quantitative Analysis: A Priori Categories

Quantitative analysis of the checklist/rating scale evaluated participant perceptions regarding scenario design modifications related to a priori defined categories by the instructors. There were two possible checklist outcome events and five possible rating scale outcome events. Statistically significant differences between making (checklist answer yes) and not making a relevant scenario modification (checklist answer no) during

**TABLE 2.** Qualitative Data Emergent Themes and Frequencies

| Theme | Subtheme | Total No. Modifications | No. Alpha Modifications | No. Beta Modifications |
|---|---|---|---|---|
| Realism | | 33 | 21 | 12 |
| | Props | 9 | 8 | 1 |
| | Patient | 10 | 6 | 4 |
| | Timing | 8 | 3 | 5 |
| | Clinician | 6 | 4 | 2 |
| Protocols | | 22 | 7 | 15 |
| | Clinical | 13 | 5 | 8 |
| | Sequence | 11 | 5 | 6 |
| | Information | 9 | 3 | 6 |
| | Students | 8 | 2 | 6 |
| | Timing | 5 | 2 | 3 |
| Introductions | | 19 | 7 | 12 |
| | Objectives | 8 | 4 | 4 |
| | Specific | 8 | 2 | 6 |
| | General | 3 | 1 | 2 |
| Roles | | 11 | 6 | 5 |

Summarizes qualitative data and categorizes emergent themes and subthemes.

combined alpha and beta testing were analyzed for each theme. For themes *Type of Simulator* (yes = 5/no = 29) and *Timing* (yes = 10/no = 24), a lesser proportion of modifications were reported than were nonmodifications ($P < 0.001$). More than 85% of teams reported no modification to the *Type of Simulator* and more than 70% reported no modifications in *Timing*. There were no statistically significant differences regarding the proportion of yes/no responses regarding thematic modifications for the other categories for combined alpha and beta testing (refer to Fig. 2: Number of Changes in Checklist Categories).

Participant results indicating the extent that the testing change was critical to the success of the scenario were described for the six instructor-determined categories. The rating scale means (ie, 1–5, completely disagree–completely agree) on the degree of importance of the design modification to the success of the scenario were determined in each category for which teams indicated that they had made a modification (yes answers noted previously). At least one response was entered for each thematic category (refer to Table 3: Rating Scale Mean Scores for Changes Critical to Scenario Success). Rank ordering of importance for combined alpha and beta testing results from most critical to least critical was *Clinical Factors, Timing, Orientation, Teaching Team Constructs*, and *Type of Simulator*. The ranges of the scores had a larger range for the categories of *Orientation, Type of Simulator,* and *Equipment/Supplies* than *Scenario Clinical Factors, Teaching Team Constructs,* and *Timing*, indicating a tendency toward agreement that the latter three were more important changes than the former three.

**TABLE 3.** Rating Scale Mean Scores in Categories Critical to Scenario Success

| | Orientation | Simulator | Equipment | Clinical | Team | Timing |
|---|---|---|---|---|---|---|
| Alpha | | | | | | |
| Mean (range) | 4.5 (3–5) | 3.0 | 4.4 (3–5) | 4.4 (3–5) | 4.5 (4–5) | 4.3 (3–5) |
| n | 11 | 1 | 5 | 10 | 11 | 3 |
| SD | 0.82 | NA | 0.89 | 0.70 | 0.52 | 1.15 |
| Beta | | | | | | |
| Mean (range) | 4.2 (1–5) | 3.0 (1–5) | 3.9 (2–5) | 4.7 (4–5) | 4.1 (2–5) | 4.4 (4–5) |
| n | 11 | 4 | 9 | 6 | 11 | 7 |
| SD | 1.17 | 1.63 | 1.17 | 0.52 | 0.94 | 0.53 |
| Both | | | | | | |
| Mean (range) | 4.4 (1–5) | 3.0 (1–5) | 4.1 (2–5) | 4.5 (3–5) | 4.3 (2–5) | 4.4 (3–5) |
| n | 22 | 5 | 14 | 16 | 22 | 10 |
| SD | 1.00 | 1.41 | 1.07 | 0.63 | 0.78 | 0.70 |

Likert scale "This change was critical to the success of the scenario" 1–5 = completely disagree-completely agree.
Quantitative descriptive data for mean rating scale scores.
NA, not available.

## DISCUSSION

Overall, our qualitative results revealed themes that would be familiar to those experienced in instructional design for simulation in healthcare as important elements during the design process. Combined with our quantitative results, they begin to reveal the experience of novice designers when they first start learning how to design simulation scenarios.

Most scenario design modifications were under the theme of REALISM during alpha testing and were efforts to increase the purported fidelity of the scenario. The Healthcare Simulation Dictionary defines realism as the "ability to impart suspension of disbelief" and refers the reader to see also definitions of fidelity.[20] Fidelity is defined as the degree or ability to replicate or reproduce real-world elements.[20] This theme thus emerges as a major focus area for novice scenario designers, even as the relative importance of realism and definitions of simulation fidelity are under scrutiny in the field of simulation in healthcare.[21] For simulation in healthcare, it remains uncertain what level psychological stress is required for effective learning and how to correlate the degree of realism to knowledge attainment.[22] It is unclear whether this high valuation of realism or fidelity by the novice designers was unwarranted or if their attention to this theme in alpha testing precluded a need for realism modifications during beta testing. There were no modifications proposed to decrease realism nor evaluation that the effort to increase realism added only extra labor without any contribution to achieving the objectives of the scenario. Given the ambiguous value of realism with respect to learning outcomes, perhaps novice designers should be cautioned about the tendency to elevate the value of fidelity during alpha testing and advised to refocus efforts on better established evidence-based elements of instructional design modification at this stage of design.

Most scenario design PROTOCOLS theme modifications were uncovered and reported during beta testing. PROTOCOLS subthemes focused on issues impacting facilitation or maintenance of the simulation scenario framework. These design issues

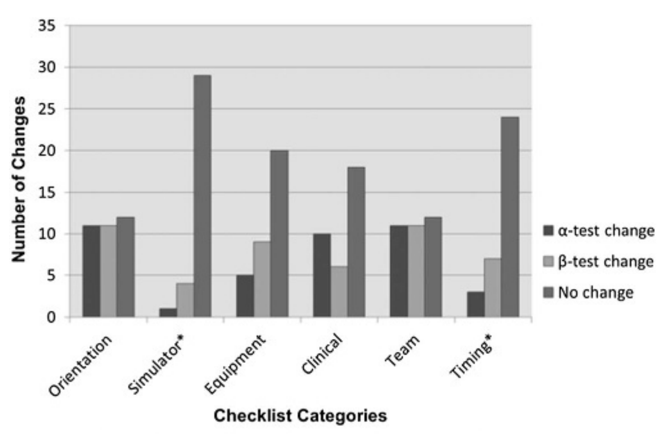


**FIGURE 2.** Number of Changes in Checklist Categories. Qualitative data comparing any change versus no change, checklist items significantly different in categories type of simulator and timing.

logically come to light during beta testing because scenario designers are challenged to recognize scenario flow variation in the absence of learner perceptions and behaviors. Despite design efforts to anticipate learner in-simulation behaviors and perceptions, unanticipated learner reactions identified during beta testing comprised a meaningful proportion of necessary design modification.

The subtheme of Timing overlapped under the two main themes of REALISM and PROTOCOLS. Participants concentrated on timing issues during both alpha and beta testing, perhaps having difficulty with anticipating the flow of the entire simulation scenario framework. Novice scenario designers may have difficulty with scripting the entire scenario and published guides suggest that it may help them envision the whole scenario as a flowchart to anticipate various directions and eventualities.[13] Scripting not only refers to lines spoken by members of the teaching team but also to the written plan and timing for a simulation event.[20] While emerging as a subtheme under two main themes in the qualitative analysis, *Timing* was actually one of the checklist categories that the participants were more likely to report they made no changes. This category, which seems to track to the critical design area of scenario scripting, may not be considered as noteworthy by novice scenario designers and introductory faculty courses may consider dedicating some time to this element of scenario design.

The theme of INTRODUCTIONS, which was identified by the qualitative analyst expert, seemed to encompass what is more commonly referred to as orientation, briefing, or prebriefing. These terms are often not clearly distinguished from each other and refer to information or session preceding the simulation to a framework and guidelines for learner participation.[20] Explicit orientation to objectives and setting of learner expectations were identified as key areas for scenario design modification. Our findings reflect the difficulty novice designers face gauging the amount and type of information to provide during an orientation that sets the stage for a successful experiential discovery learning experience.[23]

Information provided during orientation is critical to establishing learner psychological safety. It has been demonstrated that psychological safety, or belief that the environment is safe for interpersonal risk-taking, is a precursor to optimal learning behaviors.[24] Based on a literature review, expert experience, and structure from a rating scale, Rudolph et al[25] suggested a format with four content elements for how to establish psychological safety to allow the learner to effectively engage in the entire simulation exercise.[25] Their suggested elements of clarifying expectations and attending to logistic details match closely with scenario design modifications classified in the INTRODUCTIONS theme in our study.

Their suggested element of establishing a fiction contract was not identified within the theme of INTRODUCTIONS; however, the issue of REALISM, a construct that is integral to establishing a fiction contract or dispelling disbelief, emerged as a separate theme. One definition of a fiction contract in the Healthcare Simulation Dictionary describes it as the learners' willingness to set aside disbelief and engage in the simulation as being real.[20] Novice designers and design guidelines should incorporate the establishment of a fiction contract during orientation to streamline and decrease the time and effort required to modify scenarios in order to maximize realism. Our results from the rating scale demonstrated a tendency for novice designers to value *Scenario Clinical Factors* over *Orientation*, which corresponds to the theme of REALISM emerging as the most mentioned. A factor that may have contributed to this high valuation of clinical components could include the long number of years in clinical practice of the participants, although this would be difficult to ascertain as the group responses were completed collaboratively. This high consideration of orientation during the design process counterbalancing the effort toward faithful replication of a clinical situation deserves emphasis during introductory faculty development courses on simulation-based healthcare education and we have started to incorporate this concept into FunSim.

Finally, their fourth suggested element of committing to learner psychological safety likely did not emerge as a discrete factor given the research setting of faculty development course, where beta testers were not representative of a true learner cohort during scenario testing.

The checklist/rating scale was created for the purposes of this study, incorporating design elements based on instructor recollections of changes made during previous iterations of the course. The qualitative research expert who developed the codebook while doing the qualitative grounded theory analysis had not reviewed of the checklist. The themes and subthemes that emerged from the qualitative analysis link directly to each of the checklist/rating scale categories (refer to Table 4: Correlation of Checklist Items With Emergent Themes). These correlations suggest that longitudinal observations of scenario design expert instructors are well founded, based on rigorous qualitatively identified themes and subthemes. Likewise, the correlations suggest that coding of scenario design process themes and subthemes by a qualitative methods expert without prior knowledge of or experience in simulation is accurate.

The expert-designated checklist item *Type of Simulator* was not explicitly identified in this study as an emergent theme or subtheme during scenario modification, likely due to the construct of the study within a faculty development exercise, in which simulator choice was limited to those available at the teaching center. It was considered a critical element of scenario design and mapped by the expert faculty researchers to the theme of REALISM and subtheme of Props. We have observed frequent modifications of simulator type (eg, high technology simulator changed to standardized patient) in our multiyear experience teaching this course and in our actual educational experience as mentors for faculty who design simulation scenarios.

**TABLE 4.** Correlation of Checklist Categories With Emergent Themes

| Expert-Designated Design Category | Correlated Emergent Themes | Correlated Emergent Subthemes |
|---|---|---|
| Learner orientation | Introductions | Objectives/Specific/General |
| Type of simulator | Realism | Props |
| Equipment/supplies | Realism | Props |
| Scenario clinical factors | Protocols, Realism | Clinical/Patient |
| Teaching team constructs | Introductions, Roles | Specific |
| Timing | Protocols, Realism | Timing |

Qualitative data and quantitative data categories correlate, validating emergent themes.

One of the limitations of this study was the study design that mandated that designer(s) and instructor(s) for the course were investigators/authors. Instructors were responsible for coaching during alpha testing and guiding reflection and discussion after beta testing. Instructors may have unknowingly introduced bias influencing the participants' scenario modifications. This instructional construct has been used for multiyear iterations of the course before the study period years and therefore is an accurate measure of this particular faculty development course. Validation of the emergent and expert defined themes for application in scenario development exercises in other faculty development programs with different facilitators, using methodology similar to this study, would guide generalization and incorporation of these themes into scenario design templates and scenario design curricula.

The small size of the study with 17 scenario development exercises and responses from teams with members from a variety of backgrounds represent additional limitations. Our results do not represent individual novice designers and may not apply to scenario designers of all professional domains. However, in educational practice, many scenarios are designed by a group of content experts and our findings thus reflect practice-based scenario design processes. The fact that the participants' experience as educators was not explicitly ascertained also impacts the generalizability of the results. A final limitation is that checklist/rating scale for alpha testing was completed after beta testing was completed, potentially subject to recall bias, omission of items, and misallocation of elements to one or the other scenario testing session. Open-ended narrative for alpha testing was completed immediately after alpha testing and may have mitigated potential recall bias. Data collection sequencing of timing was purposeful, aiming to limit the influence of instructor-conceived themes in the freeform narrative for beta testing. This design, however, resulted in a time lag of 1 to 2 hours between the end of the alpha test and completion of the alpha test checklist/rating scale.

## CONCLUSIONS

Alpha and beta testing are common processes during instructional design of simulation scenarios in healthcare education. Though not previously and routinely referred to as such, we propose that this terminology could be applied to the iterative process that instructors use to design, test, and modify simulation scenarios for healthcare education. With this study, we observed and documented novice designer behaviors that shed light on how educators learn to design simulation scenarios through analysis of alpha and beta testing in a faculty development program. Key emergent themes suggest that those who are conducting introductory courses on instructional design for simulation should guide participants to examine areas of fidelity, orientation, and scenario scripting during the early stages of scenario design. Generalizable best practices for alpha and beta testing should be based on development and assessment of protocols incorporating instructional design principles. We encourage ongoing educational research focused on development of guidelines for alpha and beta testing. Guidelines should be based on assessment of design outcomes resulting in early-stage scenarios with high-quality features including well-designed objectives, effective orientation, facilitator guidelines, and rigorously measurable performance standards.

## REFERENCES


1. Al-Qadhi SA, Pirie JR, Constas N, et al. An innovative pediatric chest tube insertion task trainer simulation: a technical report and pilot study. *Simul Healthc* 2014;9(5):319–324.
2. Jackson A, Rajbhandari S. Developing a grapefruit model for assessment and training of diabetic foot ulcer debridement. *Simul Healthc* 2014;9(5):331–336.
3. ISTQB, *Standard Glossary of Terms Used in Software Testing Version 3.01*, 2015, International Software Testing Qualifications Board Glossary Working Group.
4. Edelson DC. Design research: what we learn when we engage in design. *J Learn Sci* 2002;11(1):105–121.
5. Morrison G, Ross S, Kemp J. *Designing Effective Instruction*. 6th ed. New York: John Wiley & Sons, Inc.; 2010.
6. Keller J. Development and use of the ARCS model of instructional design. *J Instr Dev* 1987;10:2.
7. Gagne R, Merrill M. Integrative goals for instructional design. *Educ Technol Res Develop* 1990;38(1):23–30.
8. Dick W, Carey L. *The Systematic Design of Instruction*. 3rd ed. New York: HarperCollins Publishers; 1990.
9. Anderson JM, Aylor ME, Leonard DT. Instructional design dogma: creating planned learning experiences in simulation. *J Crit Care* 2008;23:595–602.
10. Huffman J, McNeil G. Essentials of Scenario Building for Simulation Based Education. In: Grant V, Cheng A, eds. *Comprehensive Healthcare Simulation: Pediatrics*. Switzerland: Springer International Publishing; 2016.
11. Benishek LE, Lazzara EH, Gaught WL, et al. The Template of Events for Applied and Critical Healthcare Simulation (TEACH Sim): a tool for systematic simulation scenario design. *Simul Healthc* 2015;10(1):21–30.
12. Lioce L, et al. Standards of best practice: simulation standard IX: simulation design. *Clin Simul Nurs* 2015;11:9–315.
13. Alinier G. Developing high-fidelity health care simulation scenarios: a guide for educators and professionals. *Simul Gaming* 2011;42(1):9–26.
14. *CHSE Examination Blueprint*. Society for Simulation in Healthcare. 2011 Available at: http://www.ssih.org/Portals/48/Certification/Examination_Blueprint.pdf. Accessed August 25, 2017.
15. Creswell JW, Plano-Clark VL. *Designing and Conducting Mixed Methods Research*. 2nd ed. Thousand Oaks: Sage Publications; 2013.
16. Phrampus P, O'Donnell J. Debriefing Using a Structured and Supported Approach. In: *The Comprehensive Textbook of Healthcare Simulation*. Levine A, et al Editors. 2013, Springer New York.
17. Walker D, Myrick F. Grounded theory: an exploration of process and procedure. *Qual Health Res* 2006;16:547–559.
18. Khandkar SH. *Open Coding*. Available at: http://pages.cpsc.ucalgary.ca/~saul/wiki/uploads/CPSC681/open-coding.pdf. Accessed October 15, 2015.
19. Glaser B, Strauss AL. *The Discovery of Grounded Theory: strategies for qualitative research*. New Brunswick (USA) and London (UK): Aldine Transaction; 1967.
20. *Healthcare Simulation Dictionary*. Lopreiato JO, et al., Editors. 2016. Available at:http://www.ssih.org/dictionary. Accessed December 27, 2017.
21. Tun JK, et al. Redefining simulation fidelity for healthcare education. *Simul Gaming* 2015;46(2):159–174.
22. Schaumberg A. The matter of 'fidelity': keep it simple or complex? *Best Pract Res Clin Anaesthesiol* 2015;29:21–25.
23. Bruner JS. The act of discovery. *Harvard Educ Rev* 1961;31:21–32.
24. Edmondson A. Psychological safety and learning behavior in work teams. *Adm Sci Q* 1999;44(2):350–383.
25. Rudolph JW, Raemer DB, Simon R. Establishing a safe container for learning in simulation: the role of the presimulation briefing. *Simul Healthc* 2014;9:339–349.

## APPENDIX 1.

**Alpha/Beta Testing**
**Changes with ALPHA test**

**Course: iSIM FunSim**

**Team:____________**

**Date:_____________**

**Location:__________**

**Scenario Information**
- **Teaching or testing:________________**
- **Learner level:_____________________**
- **Clinical topic:_____________________**
- **Assessment method:________________**

**Please list all the changes you made to your scenario after you *started* alpha testing and the reason why you made the changes. These include any changes made after first run of scenario up until the start of the first beta test.**

**Alpha/Beta Testing**
**Checklist changes with ALPHA test**

**Course: iSIM FunSim**
**Team:____________**
**Date:_____________**
**Location:__________**
**Scenario Information**
- **Teaching or testing:________________**
- **Learner level:_____________________**
- **Clinical topic:_____________________**
- **Assessment method:________________**

**Were any changes made after starting alpha testing in the following categories?**

| | | | | |
|---|---|---|---|---|
| **LEARNER ORIENTATION** | | | **YES** | **NO** |
| This change was critical to the success of the scenario | | | | |
| 1 Completely Disagree | 2 Somewhat Disagree | 3 Neither Agree nor Disagree | 4 Somewhat Agree | 5 Completely Agree |
| **TYPE OF SIMULATOR** | | | **YES** | **NO** |
| This change was critical to the success of the scenario | | | | |
| 1 Completely Disagree | 2 Somewhat Disagree | 3 Neither Agree nor Disagree | 4 Somewhat Agree | 5 Completely Agree |
| **EQUIPMENT/SUPPLIES** | | | **YES** | **NO** |
| This change was critical to the success of the scenario | | | | |
| 1 Completely Disagree | 2 Somewhat Disagree | 3 Neither Agree nor Disagree | 4 Somewhat Agree | 5 Completely Agree |
| **SCENARIO CLINICAL FACTORS** | | | **YES** | **NO** |
| This change was critical to the success of the scenario | | | | |
| 1 Completely Disagree | 2 Somewhat Disagree | 3 Neither Agree nor Disagree | 4 Somewhat Agree | 5 Completely Agree |
| **TEACHING TEAM CONSTRUCTS (confederates, facilitation etc)** | | | **YES** | **NO** |
| This change was critical to the success of the scenario | | | | |
| 1 Completely Disagree | 2 Somewhat Disagree | 3 Neither Agree nor Disagree | 4 Somewhat Agree | 5 Completely Agree |
| **TIMING** | | | **YES** | **NO** |
| This change was critical to the success of the scenario | | | | |
| 1 Completely Disagree | 2 Somewhat Disagree | 3 Neither Agree nor Disagree | 4 Somewhat Agree | 5 Completely Agree |

**Alpha/Beta Testing**
**Changes after BETA test**

**Course: iSIM FunSim**

**Team:____________**

**Date:_____________**

**Location:__________**

**Scenario Information**
- **Teaching or testing:________________**
- **Learner level:_____________________**
- **Clinical topic:_____________________**
- **Assessment method:________________**

**Please list all the changes you made to your scenario after your first beta test and the reason you made these changes.**

**Alpha/Beta Testing**
**Checklist changes after BETA test**

**Course: iSIM FunSim**
**Team:____________**
**Date:_____________**
**Location:__________**
**Scenario Information**
- **Teaching or testing:________________**
- **Learner level:_____________________**
- **Clinical topic:_____________________**
- **Assessment method:________________**

**Were any changes made after beta testing in the following categories?**

| | | | | |
|---|---|---|---|---|
| **LEARNER ORIENTATION** | | | **YES** | **NO** |
| This change was critical to the success of the scenario | | | | |
| 1 Completely Disagree | 2 Somewhat Disagree | 3 Neither Agree nor Disagree | 4 Somewhat Agree | 5 Completely Agree |
| **TYPE OF SIMULATOR** | | | **YES** | **NO** |
| This change was critical to the success of the scenario | | | | |
| 1 Completely Disagree | 2 Somewhat Disagree | 3 Neither Agree nor Disagree | 4 Somewhat Agree | 5 Completely Agree |
| **EQUIPMENT/SUPPLIES** | | | **YES** | **NO** |
| This change was critical to the success of the scenario | | | | |
| 1 Completely Disagree | 2 Somewhat Disagree | 3 Neither Agree nor Disagree | 4 Somewhat Agree | 5 Completely Agree |
| **SCENARIO CLINICAL FACTORS** | | | **YES** | **NO** |
| This change was critical to the success of the scenario | | | | |
| 1 Completely Disagree | 2 Somewhat Disagree | 3 Neither Agree nor Disagree | 4 Somewhat Agree | 5 Completely Agree |
| **TEACHING TEAM CONSTRUCTS (confederates, facilitation etc)** | | | **YES** | **NO** |
| This change was critical to the success of the scenario | | | | |
| 1 Completely Disagree | 2 Somewhat Disagree | 3 Neither Agree nor Disagree | 4 Somewhat Agree | 5 Completely Agree |
| **TIMING** | | | **YES** | **NO** |
| This change was critical to the success of the scenario | | | | |
| 1 Completely Disagree | 2 Somewhat Disagree | 3 Neither Agree nor Disagree | 4 Somewhat Agree | 5 Completely Agree |