**Crowdsourcing Unmet Needs in Simulation-Based Education and Technology**

Cory Schaffhausen[1], Robert Sweet[1,2], David Hananel[1] Kathleen Johnson[2], Timothy Kowalewski[1]

[1] University of Minnesota (United States)
[2] American College of Surgeons (United States)

**Background**:

Medical manikin simulation technology has significant unrealized potential for a breadth of healthcare applications that need training solutions. Improvements to manikins must be driven by the needs of relevant stakeholders across the spectrum of healthcare applications. This study introduces a novel needs assessment method and compares it to a classic "focus group" needs assessment to define needs for manikin development.

**Methods**:

A needs assessment was distributed to 89 sites of the American College of Surgeons (ACS) Accredited Education Institutes (AEI) Consortium. The assessment was performed using a custom web application that displays visual (e.g. images) and textual (e.g. example needs) as context to aid entering open-ended need statements. Participants reviewed instructions and training, such as to refrain from describing desired solutions and to focus only on needs. Results were compared with data from focus group sessions. 80 focus group needs were identified from approximately 8 hours of audio representing 2 site visits at West Virginia University and University of Minnesota. Comparisons were automated using a semantic textual similarity (STS) algorithm to identify common and unique needs across methods.

**Results**:

The assessment was accessed by 21 individuals. 7 respondents proceeded through to completion. A total of 20 need statements were submitted. Comparing crowdsourcing statements to focus group statements demonstrated both overlapping (3 common) needs and also unique needs only submitted via crowdsourcing. Average minutes of participation per need statement ranged from 5 (crowdsourcing) to 30 (focus groups).

**Conclusions**:

Crowdsourcing methods can be effective in rapidly generating unmet needs and can identify common as well as unique unmet needs compared with more time- and resource- intensive focus groups.

## 1 Background

While manikin development has advanced significantly in previous decades, much existing technology remains in nascent stages [1]. The development of effective medical manikin simulation technology begins with identifying existing umnet user needs. The use of simulation manikins across many stakeholder groups creates a complex mix of user needs, and this qualitative information is time consuming to obtain. Traditional qualitative research methods such as in-depth interviews or focus groups can require individual participation times of 1 to 4 hours. This resource intensive process typically results in a limited number of participants (e.g. fewer than 30). Needs assessments relying on in-depth interviews include surgical training national stakeholder needs (n=22) [2]. Survey methods allow for higher numbers of participants (often 100 or more) including needs assessments for continuing professional development (n=71) [3] and undergraduate surgical training programs (n=123 graduates and n=55 surgeons) [4]. However, existing survey tools are not well suited to capturing qualitative data. Combined approached have been described for needs assessments of surgical residents as teachers (n=235 surveys, n=5 interviews, n=2 focus groups) [5]. Additional ethnographic methods such as personas and journey maps have been reported, as in development of electronic tools for cardiovascular health [6].

The present study evaluates a novel web-based needs assessment tool to improve the depth and completeness of qualitative responses. Rather than typical surveys with only textual information, the approach herein provides additional visual and textual information to provide context for thinking about simulation technologies. This information was intended as a stimulus because participants might be far removed from use environments while entering data (e.g. sitting at an office computer instead of a clinical setting). Previous crowdsourcing studies (n=100+ users) of consumer products demonstrate consistent increases (up to 50%) in need statement responses when viewing provided contextual images and text [7]. In addition, these large groups of diverse users consistently generate

extensive lists of unmet needs (500 or more statements) with rates of exact duplicates less than 1% [7] and rates of semantically equivalent statements less than 5% [8].

This study evaluates the feasibility of a novel web-based needs assessment method applied to simulation manikins. Crowdsourcing and focus group needs assessments were initiated in order to inform the creation of a requirements document to guide development of a new generation of advanced modular manikin technologies. Comparisons across methods help inform whether web-based methods might decrease resource-intensive participation times to collect detailed qualitative data and whether similar topics may be discussed by participants in groups of different methods.

## 2 Methods

The comparison of needs assessment data used data from two sources: previously recorded audio from 2 focus group sessions and new data generated from an online crowdsourced needs assessment.

### 2.1 Focus Group Needs Assessment Data

Focus groups were performed at 2 locations, West Virginia University and University of Minnesota. Each focus group included a minimum of 5 participants and an experienced facilitator. Individuals within the focus group may have participated during the complete session or portions of the session depending on availability; therefore, participant counts varied. Participants represented interdisciplinary fields including clinicians, technicians, and educators and were primarily civilian simulation users. The session at the University of Minnesota included tables arranged around 2 functional manikins (iStan and SimMan) to facilitate interactive discussions, whereas the West Virginia University meeting did not include similar visual aids.

Broad discussion points included high-level requirements as well as detailed specific functions. Summary notes were taken during sessions. Sessions were recorded, and after completion, audio tracks were isolated from video recordings. A single analyst reviewed approximately 8 hours of audio as well as written notes taken during sessions. Any statement representing an unmet need for simulation technology was documented in list form.

**2.2 Online Crowdsourced Needs Assessment Data**

The crowdsourced needs assessment was performed using a custom web application developed with Zoho Creator software (Zoho, Inc, Pleasanton, CA). The web application could be accessed through an emailed link. This link was distributed to 89 sites of the American College of Surgeons (ACS) Accredited Education Institutes (AEI) Consortium. The assessment interface was designed to display visual (e.g. images) and textual (e.g. example needs) information as simulation-specific context as an aid to entering open-ended need statements. The application was accessible for a two-week duration.

Upon clicking the emailed link, participants would view an introduction to the study and then enter demographic information (e.g. civilian/military, provider/educator or simulation support, years of experience). Participants viewed instructions explicitly stating a preference for descriptions of unmet needs rather than descriptions of desired solutions. This was a standard instruction for all previous work to limit confusion over intellectual property and to approach the problem-solving phase of development independent of problem definition. The instructions also stated to list unmet needs using complete sentences. The instructions were reinforced with a set of appropriate and inappropriate example need statements, and an optional quiz was available to further check training comprehension.

The interface to enter unmet needs used side-by-side viewing windows where one included text boxes to type open-ended statements and the other provided space to simultaneously view information to help think of needs (e.g. images as a visual stimulus). Figure 1 represents a screen capture of the need entry interface with a display of images related to simulation training and technology.

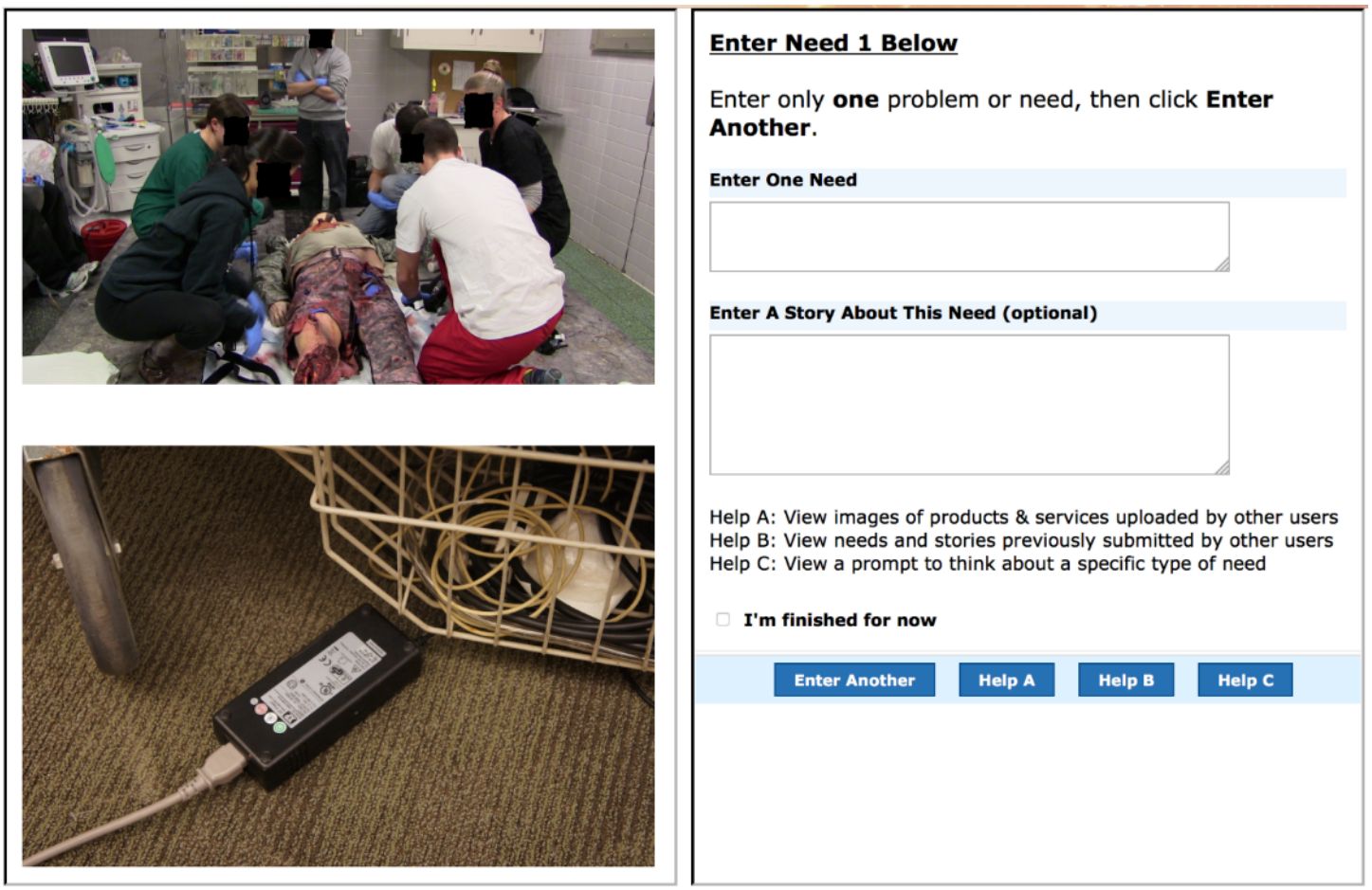


Figure 1: Screen capture of need entry interface and scrollable list of images

Participants were instructed to use a single sentence need statement as a summary and to provide more detailed background and context as an optional story about the need. The story would more fully explain relevant details (who, what, where, etc.). During the needs assessment, the participant could select any type of information displayed to help think of needs and could repeat the selection any number of times.

### 2.3 Images and Text Displays for Needs Assessment

Three types of information were available to display: 1) Images showing simulation products or technology or relevant training scenarios; 2) Previously generated need statements describing what other people consider unmet needs; and 3) Short narrative prompts that might be analogous to questions asked during qualitative research (e.g. interviews or focus groups).

**Displayed Images:**

The pool of images included photographs taken during the University of Minnesota focus group as well as a broader selection of simulation scenarios selected from a photograph archive at the University of Minnesota in SimPORTAL / CREST. A total of 140 images were included in the pool. Visible names and faces were de-identified. Each time a participant clicked the help button for images the display window refreshed with a scrollable list of 10 randomly selected (and non-repeated) images. Figure 2 displays thumbnail views of a sample of this image pool.

Figure 2: Random selection of images included in image pool for web application

**Example Need Statements:**

The pool of example need statements includes previously collected need statements and allowed participants to view additional examples of appropriate statements and to potentially build off of statements from other users in a manner similar to divergent thinking exercises (e.g. group brainstorms). Previous studies demonstrate minimal duplication of new need statements compared to displayed need statements [8] when using this method. Each time a participant clicked the help button for need statements the display window refreshed with a list of 10 randomly selected (and non-repeated) need statements and full-length stories (if included). An example of a previously identified need statement included in the pool is “The strength of the pulse doesn't vary enough.”

**Short Narrative Prompts:**

A third form of displayed information was a short narrative prompt in the form of a question. The pool of 18 questions described a wide range of themes such as past experiences of frustration, uncertainty, or observations of others. The questions allowed participants to think about past experiences with a focus on specific types of unmet needs. For example, one prompt focusing on needing help asked, “Think of a time when you were working with a manikin, and you needed to stop and ask for help (or you wished you could). This could have been asking a friend or coworker, looking online, or making a call. What was the exact problem you needed help with?” Each time a participant clicked the help button for prompts the display window refreshed with one randomly selected (and non-repeated) prompt.

### 2.4 Comparisons Across Methods

Two comparisons were performed across needs assessment methods. First, estimated participation times were calculated to determine relative rates of data collection. Second, an automated semantic textual similarity (STS) algorithm was used to identify common and unique needs across methods. The STS algorithm was previously identified as a state-of-the-art algorithm for comparisons of

sentence-length text [9,10] and proved useful for analyzing high volumes of need statements [8]. All pairwise combinations of individual statements were generated. The STS algorithm provides a numerical score from 0 (not similar) to 1 (equivalent meaning). Low scores of crowdsourced need statements compared to all focus group need statements indicated a need identified via crowdsourcing that was not identified via focus groups. Any score for a comparison across methods above a cutoff of 0.6 indicated some similarities in meaning where topics for both methods overlap.

## 3 Results

The web application was accessed by 21 individuals. 7 respondents proceeded to complete the assessment and submit data. A conservative online time estimate includes a time allowance of those who started the assessment and did not complete it. While exact durations in this scenario are not known, an estimate of 3 minutes each was used as sufficient time to read introductory pages and detailed instructions. Times for completed crowdsourcing assessments were recorded (min: 2 minutes, max: 17 minutes, mean: 7.5 minutes). A total of 20 need statements were submitted with a range of 1-7 per person. Table 1 summarizes a comparison of estimated participation times. Estimated participation times for focus groups are approximate due to variations in group size throughout the sessions. The required participation time per need statement using online crowdsourcing is approximately 15% that of focus group sessions.

Table 1: Participation times and data collection rates for focus groups and crowdsourcing methods

| Method | Total Need Statements | Participation time | Minutes / Need Statement |
|---|---|---|---|
| Focus Groups | 80 | 5 users * 4 hours * 2 sessions = 2400 minutes | 30 minutes |
| Crowdsourcing | 20 | 14 Incomplete sessions (3 minutes each) + 7 complete sessions (7.5 minutes each) = 94.5 minutes | 5 minutes |

Searching for redundancy within the crowdsourcing statements resulted in two similar, but not equivalent, statement pairs. Comparing crowdsourcing statements to focus group statements demonstrated both a set of overlapping (3 common) needs and also unique needs submitted via crowdsourcing and not identified during focus groups. Table 2 includes all submitted crowdsourcing need statements (full length stories omitted) where similar focus group statements were not identified. Table 3 includes crowdsourcing need statements identified as overlapping and the corresponding focus groups need statements

Table 2: Need statements only identified via online crowdsourcing

| | **Complete Need Statement Text** |
|---|---|
| 1 | We would like to be able to place organs in Sim man's abdominal cavity so we do laparoscopy sim within an inter-professional education simulation e.g. with anesthesia and nursing staff in an OR |
| 2 | None of the available small intestines have a mesentery |
| 3 | Trauma man: the window for chest tube insertion is too low in the axilla. We teach the students only to place the tubes at nipple line or higher, and only a very small proportion of the window is above the nipple line. |
| 4 | Trauma man: the overlying window is bigger than the underlying window for chest tube insertions, so students inadvertently cut through the overlying skin in the boarder around the underlying window, and damage the manikin |
| 5 | Trauma man Intercostal vessels are not anatomically correct- they get cut and fluid spills out when students insert chest tubes |
| 6 | There needs to be a window in the overlying skin where chest tubes are inserted so that we don't need to replace a full (and expensive) skin after chest tube insertion. |
| 7 | It would be good to have female options for Sim Man in terms of a programmed female voice + chest appearance so that the manikin has breasts |
| 8 | A manikin needs to be affordable. |
| 9 | The materials of the manikin need to be like a human. |
| 10 | More realistic skin |
| 11 | To fix the constant failure of the more advanced Manikins connection issues, more than not, by not having the companies Tech Support blame the issues on user error. |
| 12 | I wish mouth could close but jaw could be hinged and opened if needed. |
| 13 | I wish manikin had rotating wrist, elbow, and knee joints. |
| 14 | I wish lung sounds were not so mechanical or affected by background mechanical noise. |
| 15 | I wish manikin voice was not so difficult to hear - pre-recorded or operator generated. If you turn up the volume - distortion occurs. If scenario is being live streamed or recorded - microphone distortion is increased. |
| 16 | I need to be able to have training to aid me in programing the simulators easier. Especially if they require certain physiological traits to program. |
| 17 | virtual reality trainers (like Bronch mentor) that communicate with LMS |

Table 3: Similar need statements identified in focus groups and online crowdsourcing

| | Crowdsourcing Need Statement | Focus Group Need Statement |
|---|---|---|
| 1 | I wish radial pulse spot was more anatomically correct. | Manikin only has a pulse on the right radial wrist. |
| 2 | The manikin needs to elicit a human connection with the trainee. | Current manikins lack a human connection. |
| 3 | The manikin's response needs to be lifelike. | I want an immediate response or reaction to an input to the manikin. |

## 4 Discussion

The results indicate a web application can be a feasible tool to collect qualitative data for a simulation technology needs assessment with results comparable to those of focus groups. The participants stated a wide range of topics in open-ended responses and generated need statements with greater efficiency compared to focus groups. A larger sample size is warranted to further evaluate group sizes that may generate a comparable number of need statements. Based on present duplication rates, a total of approximately 30 participants completing the assessment could exceed the need statement count extracted from the focus group sessions. Assuming an average of 7.5 minutes each participant, the total of 225 minutes compares favorably to 2400 minutes of focus group participation.

A number of improvements to the web application are warranted given the high number of participants choosing not to complete the assessment. Participant feedback indicated the time to review instructions and examples was overly cumbersome. These portions should be minimized and future trials can evaluate the impact on clarity and completeness of responses when users are given less guidance. A feasibility trial using alternative and simplified methods applicable to medical conference settings has shown promise. Here, participants did not view stimulus information online. Instead participants were engaged in viewing and discussing current clinical challenges during conference sessions and descriptions of unmet clinical needs could be submitted in real time via text

message [11]. While this method has advantages, it is limited to schedules and technical support available for target conferences.

While this study describes the development of a promising new method for gathering information on needs, a number of limitations to the study are evident. When transcribing need statements from focus group recordings, only a single analyst reviewed the data. Additional analysts may increase the counts of need statements generated from the same recordings. The automated algorithm comparison is susceptible to false positives and false negatives, as previously described [8]; however, manual methods become less feasible as user group sizes increases. In-depth interviews were not included in the current study. While this method might decrease participation times (e.g. 20 interviews requires 1200 minutes), the analyst time to review 20 hours of recordings (rather than 8) is much larger. The present study does not attempt to prioritize or rank need statements; however previous work has demonstrated one method of prioritizing large sets of need statements [12] and suggested a benefit of including needs from short duration activities [13].

While the present study relates to medical simulation manikins, the same methods may be appropriate for a wide range of research areas, comparable in breadth to traditional focus group and interview methods. The current climate for increasing use of human centered design in health services and delivery as well as rapid development of new medical device technologies represent key areas demanding data on user needs.

## 5 Conclusion

Crowdsourcing methods can be effective in rapidly generating unmet needs and can identify common as well as unique unmet needs compared with focus groups with fewer resources. Individual time commitments for online participants are substantially lower than focus group participants as is the combined participation time per need statement collected via online needs assessment.

**Acknowledgement**

DOD funding source: This research is work done as part of the MedSim Combat Casualty Training Consortium (CCTC) funded by Telemedicine & Advanced Technology Research Center (TATRC) under contract W81XWH-11-2-0185. This research and development project was conducted by and was made possible by a contract vehicle which was awarded and administered by the U.S. Army Medical Research & Materiel Command and the Medical Simulation and Information Sciences Joint Program Committee, at Fort Detrick, MD under award number: W81XWH‐11‐02‐0185. The views, opinions and/or findings contained in this publication are those of the authors and do not necessarily reflect the views of the Department of Defense and should not be construed as an official DoD/Army position, policy or decision unless so designated by other documentation. No official endorsement should be made.

**References**


[1] Cooper JB, Taqueti VR. A brief history of the development of mannequin simulators for clinical education and training. Qual Saf Health Care. 2004;13(Suppl 1):i11-i18.

[2] Kim S, Dunkin BJ, Paige JT, Eggerstedt JM, Nicholas C, Vassilliou MC, Spight DH, Pliego JF, Rush RM, Lau JN, Carpenter RO. What is the future of training in surgery? Needs assessment of national stakeholders. Surgery. 2014 Sep 30;156(3):707-17.

[3] Wallace T, Birch DW. A needs-assessment study for continuing professional development in advanced minimally invasive surgery. The American journal of surgery. 2007 May 31;193(5):593-6.

[4] Birch DW, Mavis B. A needs assessment study of undergraduate surgical education. Canadian journal of surgery. 2006 Oct 1;49(5):335.

[5] Rotenberg BW, Woodhouse RA, Gilbart M, Hutchison CR. A needs assessment of surgical residents as teachers. Canadian Journal of Surgery. 2000 Aug;43(4):295.

[6] Lis Neubeck, Genevieve Coorey, David Peiris, John Mulley, Emma Heeley, Fred Hersch, Julie Redfern, Development of an integrated e-health tool for people with, or at high risk of, cardiovascular disease: The Consumer Navigation of Electronic Cardiovascular Tools (CONNECT) web application, International Journal of Medical Informatics http://dx.doi.org/10.1016/j.ijmedinf.2016.01.009; In Press.

[7] Schaffhausen CR, Kowalewski TM. Large-Scale Needfinding: Methods of Increasing User-Generated Needs From Large Populations. Journal of Mechanical Design. 2015 Jul 1;137(7):071403.

[8] Schaffhausen CR, Kowalewski TM. Large scale needs-based open innovation via automated semantic textual similarity analysis. In ASME 2015 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference. 2015; DETC2015-47358: V007T06A045.

[9] Agirre E, Cer D, Diab M, Gonzalez-Agirre A, Guo W. SEM 2013 shared task: Semantic textual similarity, including a pilot on typed-similarity. In *SEM 2013: The Second Joint Conference on Lexical and Computational Semantics. Association for Computational Linguistics 2013.

[10] Han L, Kashyap A, Finin T, Mayfield J, Weese J. UMBC EBIQUITY-CORE: Semantic textual similarity systems. In Proceedings of the Second Joint Conference on Lexical and Computational Semantics 2013 Jun 13 (Vol. 1, pp. 44-52).

[11] Schaffhausen CR, Kowalewski TM. Crowdsourcing Unmet Clinical Needs in Minimally Invasive Surgery. Journal of Medical Devices. 2016; In Press.

[12] Schaffhausen CR, Kowalewski TM. Assessing Quality of User-Submitted Need Statements From Large-Scale Needfinding: Effects of Expertise and Group Size. Journal of Mechanical Design. 2015 Dec 1;137(12):121102.

[13] Schaffhausen CR, Kowalewski TM. Assessing quality of unmet user needs: Effects of need statement characteristics. Design Studies. 2016; 44:1-27.