An Electronic Competency-Based Evaluation Tool for Assessing Humanitarian Competencies in a Simulated Exercise

Andrea B. Evans, MDCM, MSc;[1,2] Jennifer M. Hulme, MD, MPH;[3,4] Peter Nugus, PhD, MA(Hons), Med;[5,6] Hilarie H. Cranmer, MD, MPH;[7] Melanie Coutu, MA;[8,9] Kirsten Johnson, MD, MPH[6,8,9,10]

1. Emergency Department, The Hospital for Sick Children, Toronto, Canada
2. Center for Global Child Health, The Hospital for Sick Children, Toronto, Canada
3. Emergency Department, University Health Network, University of Toronto, Toronto, Canada
4. Department of Family and Community Medicine, University of Toronto, Toronto, Canada
5. Centre for Medical Education, McGill University, Montreal, Canada
6. Department of Family Medicine, McGill University, Montreal, Canada
7. Global Health, Massachusetts General Hospital, Boston, Massachusetts USA
8. McGill Humanitarian Studies Initiative, McGill University, Montreal, Canada
9. Humanitarian U Inc., Montreal, Canada
10. Emergency Department, McGill University Health Centre, Montreal, Canada

Correspondence:
Kirsten Johnson, MD, MPH
McGill University
Department of Family Medicine
5858 Cote des Neiges, 3rd Floor, Suite 300
Montreal, Quebec, Canada H3S 1Z1
E-mail: Kirsten.johnson@mcgill.ca

Conflicts of interest: Kirsten Johnson is the CEO of Humanitarian U. The other authors declare no conflicts of interest.

Keywords: competency; evaluation; humanitarian response; humanitarian training; humanitarian workers; simulation exercise

Abbreviations:
CBHA: Consortium of British Humanitarian Agencies
MD: Medical Doctor
SimEx: simulation exercise

Abstract

Methods: The evaluation tool was first derived from the formerly Consortium of British Humanitarian Agencies' (CBHA; United Kingdom), now "Start Network's," Core Humanitarian Competency Framework and formatted in an electronic data capture tool that allowed for offline evaluation. During a 3-day humanitarian simulation event, participants in teams of eight to 10 were evaluated individually at multiple injects by trained evaluators. Participants were assessed on five competencies and a global rating scale. Participants evaluated both themselves and their team members using the same tool at the end of the simulation exercise (SimEx).

Results: All participants (63) were evaluated. A total of 1,008 individual evaluations were completed. There were 90 (9.0%) missing evaluations. All 63 participants also evaluated themselves and each of their teammates using the same tool. Self-evaluation scores were significantly lower than peer-evaluations, which were significantly lower than evaluators' assessments. Participants with a medical degree, and those with humanitarian work experience of one month or more, scored significantly higher on all competencies assessed by evaluators compared to other participants. Participants with prior humanitarian experience scored higher on competencies regarding operating safely and working effectively as a team member.

Conclusion: This study presents a novel electronic evaluation tool to assess individual performance in five of six globally recognized humanitarian competency domains in a 3-day humanitarian SimEx. The evaluation tool provides a standardized approach to the assessment of humanitarian competencies that cannot be evaluated through knowledge-based testing in a classroom setting. When combined with testing knowledge-based competencies, this presents an approach to a comprehensive competency-based assessment that provides an objective measurement of competency with respect to the competencies listed in the Framework. There is an opportunity to advance the use of this tool in future humanitarian training exercises and potentially in real time, in the field. This could impact the efficiency and effectiveness of humanitarian operations.

Evans AB, Hulme JM, Nugus P, Cranmer HH, Coutu M, Johnson K. An electronic competency-based evaluation tool for assessing humanitarian competencies in a simulated exercise. *Prehosp Disaster Med*. 2017;32(3):253-260.

Introduction

There is an urgent need to expand the professional humanitarian workforce. Gaps remain in standardization of training and tools for benchmarking competency and readiness for deployment. A growing number of humanitarian training programs are using simulation exercises (SimEx) in an effort to train and prepare humanitarians for work in the field. This study presents a novel, competency-based evaluation tool designed for rapid electronic offline use in a field-based SimEx.

Received: February 21, 2016
Revised: June 30, 2016
Accepted: July 10, 2016

Online publication: February 21, 2017

doi:10.1017/S1049023X1700005X

The United Nations Office for Humanitarian Affairs (UNOCHA; New York, USA and Geneva, Switzerland) called for US $20.1 billion in 2016 to provide life-saving humanitarian assistance to over 87.6 million people across 37 countries.[1] Crises are becoming more protracted and displacement levels are unprecedented with more than 60 million people fleeing their homes in 2015.[1] The number, frequency, and severity of humanitarian crises are only predicted to increase due to an increasing number of fragile states affected by conflict, urbanization, and climate-related events.[1,2] Experts predict that climate-related disasters alone will affect 375 million people annually.[3] Furthermore, more than 50% of the world's population will be living in urban settings by 2030 – most of these in slums that present an increase in vulnerability to disaster.[4]

Currently, the number of humanitarians responding to the global need is estimated at 450,000.[5] Many are untrained and ill equipped to provide an effective and efficient response.[6,7] There is an urgent need to expand the humanitarian workforce with competent leaders who have the proper training to provide them with the right tools, knowledge, skills, and behaviors.[8-11]

Humanitarian training includes agency-run, pre-departure trainings and a growing number of academic-affiliated humanitarian training programs.[8] An increasing number of humanitarian trainings include simulation-based practice and learning.[12-14] SimEx are an important adjunct to the traditional classroom-based trainings that should be essential for individual and team preparation prior to working in the field. Simulation is widely applied in medicine, military, and aviation.[15] Similarly, humanitarian workers face many situations that demand prior exposure and realistic training. SimEx training amplifies real experiences with guided ones, often "immersive" in nature, that evoke or replicate substantial aspects of the real world in a fully interactive fashion.[15] Simulation-based learning may be the most appropriate way to develop humanitarian professionals' knowledge, skills, and attitudes, while protecting them, their organizations, and beneficiaries from unnecessary risks.[15] Simulation-based training techniques, tools, and strategies can be applied in designing structured learning experiences, as well as be used as a measurement tool linked to targeted competencies and learning objectives.[15]

Humanitarian SimEx training has evolved to offer innovative scenarios and skills stations, also called "injects" for participants.[13,14,16] However, few SimEx trainings include methods and tools designed to assess the essential humanitarian competencies that participants must demonstrate in the SimEx and in the field.[17] Competency-based participant evaluation during a SimEx allows learners to benchmark their performance, identify areas for improvement, and has the potential to determine readiness.[18-24] This study presents a novel, competency-based evaluation tool designed for rapid electronic offline use in a field-based SimEx.

Methods

This observational study was conducted at Camp Interval, Saint Lucie des Laurentides, Quebec, Canada from May 15-17, 2015 as a mandatory part of the Canadian Disaster and Humanitarian Response Training Program. This Program is a collaborative effort between the Canadian Consortium for Humanitarian Training (CCHT; Montreal, Quebec, Canada), McGill University's Humanitarian Studies Initiative (HSI; Montreal, Quebec, Canada), and Humanitarian U (Montreal, Quebec, Canada). All 63 participants were included (Figure 1). The study was approved by McGill University's Research Ethics Office (IRB) and informed consent was obtained from each SimEx participant.

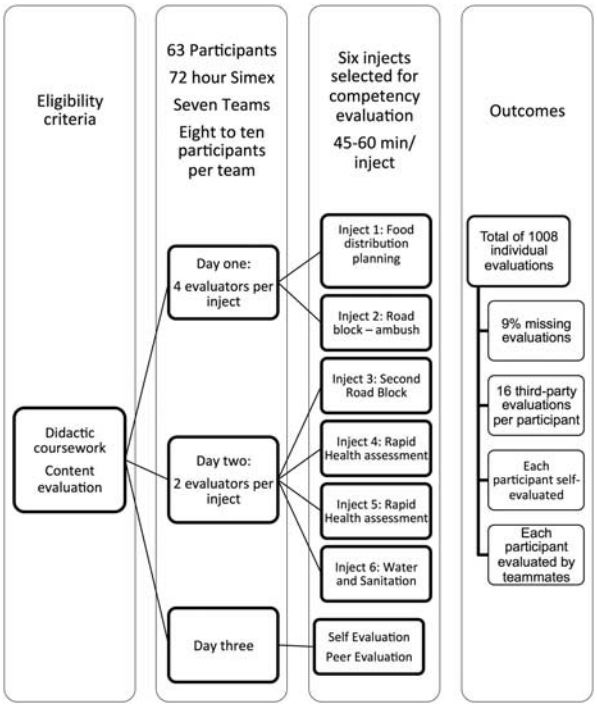

Figure 1. Humanitarian SimEx Schematic.
Abbreviation: SimEx, simulation exercise.

SimEx

The 72-hour-long SimEx took place in a rustic summer camp, 1.5 hours north of Montreal, Quebec, Canada. Forty-eight experienced faculty and volunteers who came from partner universities, humanitarian non-governmental organizations, and the private sector facilitated the SimEx. The minimum required experience for volunteers was at least one deployment to the field for humanitarian aid, global health, or disaster relief. Faculty had extensive experience abroad and were considered experts in the field of humanitarian aid and disaster relief. As per Bradt et al, professionals, or "experts," are defined as those meeting the criteria of experience in clinical medicine, public health, and/or disaster management, and is more commonly defined as months to years of full-time, hands-on field service in disaster or humanitarian settings.[25] The SimEx was based on a scripted, fictional humanitarian emergency in a conflict-affected country called *Simlandia* that is struck by an earthquake and then devastated by a tsunami.

The SimEx began 24 hours after the onset of a fictional tsunami on Day 1 of the disaster, when participants arrive in country after crossing customs and immigration. Day 2 of the SimEx is one week into the emergency response, and Day 3 is two weeks into the emergency response. Participants worked in eight teams of 10 to conduct needs assessments, mapping, communication, draft reports, give press conferences, attend cluster meetings, and work through different injects (skills stations). There were 11 injects in total, each lasting 45-60 minutes. These injects were developed with learning activities that linked to core humanitarian competencies.[17] Six injects were used as evaluation points, and a small evaluation team comprising two to six evaluators observed each participant as they completed tasks and demonstrated competencies specific to that inject (Figure 1). For example, specific behaviors demonstrating the competency "Managing oneself in a stressful environment"

https://doi.org/10.1017/S1049023X1700005X Published online by Cambridge University Press

were assessed at the Roadblock inject. Injects that were used for evaluation of competency domains included: Food Distribution, Roadblock, Vaccination Campaign, Ambush, Rapid Health Assessment, and Water and Sanitation. Overall, participants were evaluated individually 16 times over the course of the first two days. Participants conducted a self-evaluation and evaluated each of their team members on the third and final day of the SimEx.

Participants

Participants completed pre-requisite coursework prior to the SimEx. Coursework was done in one of three ways: (1) as part of Laval University's (Quebec City, Quebec, Canada) Masters in Humanitarian Studies Program;[26] (2) a 55-hour, in-class program at McGill University;[27] or (3) a 2-week program that involved one week online[28] and one week in-class.[16]

All students submitted a written application to participate in the SimEx, paid the SimEx fee, and signed a waiver to participate. All participants had a working proficiency in English, completed the course requirements described above, and agreed to participate fully in the 3-day event.

Evaluation Tool

The competency-based evaluation tool used competencies from the formerly Consortium of British Humanitarian Agencies (CBHA; United Kingdom), now "Start Network," Core Humanitarian Competency Framework.[17] Five competencies that are difficult to assess through knowledge-based testing and more conducive to evaluation in a SimEx were selected and are listed in Figure 2. An additional sixth question was added to indicate readiness for deployment in a humanitarian response.

A team of professional humanitarian health workers ("experts") with prior experience in the field in disaster response or humanitarian aid situations (KJ, HC, PN, MC, AE) derived key measurable behaviors from the five competencies. A preliminary version of the evaluation tool was used in a pilot study in a similar SimEx at a different location in May 2014. A global rating score (Figure 2, Q6) was included given evidence for global rating scales as a valid and reliable method to evaluate technical skills.[29] Based upon user feedback and results, the tool was refined to minimize respondent fatigue and more accurately evaluate student performance.[30,31]

Behaviors representing each competency were graded on a 4-point, numerical Likert scale based on the participant's performance: 1 = very poor, 2 = inadequate, 3 = good, 4 = excellent, and n/a = unable to rate. The global rating scale was scored with a different scale where 1 = definitely not, 2 = probably not, 3 = very probably, and 4 = definitely. The 4-point scale forced the evaluator to rate the participant as "inadequate" or "good" (Figure 3).

Study data and informed consent were collected and entered using an electronic data capture tool (SurveyGizmo online survey software tool; SurveyGizmo; Boulder, Colorado USA).[32] This survey tool is Health Insurance Portability and Accountability Act (HIPAA) compliant and is a secure, web-based application designed to support data capture or surveys by providing: an intuitive interface for data entry; tracking of data and export procedures; and exporting of data downloads to common statistical packages. The electronic platform with these security features was accessible at a minimal cost. Two administrators had password-protected access to the database. Data were de-identified and exported into statistical software package for analysis. The survey was downloaded onto iPads (Apple; Cupertino, California USA) and each evaluator had their own iPad for the duration of the SimEx. Evaluators entered all evaluations into SurveyGizmo on their respective iPads and the data were saved immediately onto the secure SurveyGizmo server. These data were accessible only by the two administrators who had password-protected access.

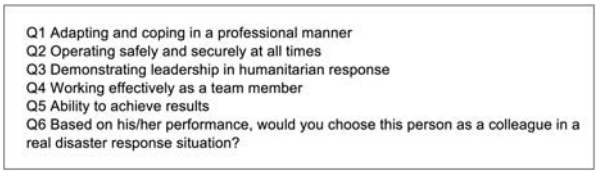

Figure 2. Competencies Used in the Evaluation Tool.

Participants downloaded the electronic data capture tool onto their own mobile devices (iPads or smart phones) and were oriented to the evaluation tool prior to the start of the SimEx. They were trained on how to evaluate each other and themselves with a common understanding of the competencies and Likert scale. On day three, participants were asked to evaluate themselves and their peers before end of the SimEx. These data were stored immediately on the secure Survey Gizmo server and accessed only by the two administrators who had password-protected access.

Evaluation Methods

Evaluators had extensive experience in humanitarian field work, education, or evaluation methods. There was a total of eight evaluators on each day. One evaluator was replaced on the second day, resulting in a total of nine different evaluators over two days.

Evaluators met twice to standardize their common definitions for each competency and the corresponding behaviors that must be demonstrated by participants. Evaluators pilot tested the tool at one SimEx inject prior to starting and compared their individual data pertaining to the behaviors assessed. They ensured familiarity with the use of the electronic system and compared their individual Likert scales to the global rating score. Evaluators were surveyed for their feedback on the application of the tool after the SimEx.

Evaluators were assigned randomly to SimEx injects. A total of six injects were used as evaluation stations. Evaluators attempted to rate every participant on all five competencies at each inject. On the first day of the SimEx, participants were evaluated at two injects by four evaluators at a time for a total of eight ratings of each competency. On day two, participants were evaluated at four skills stations by two evaluators at each station for a total of 16 ratings for each competency. Each participant was identified with a colored and numbered pinney so as to create nameless unique identifiers that were coded by number and color in the electronic tool.

Data were loaded anonymously to a secure server where evaluations were coded instantaneously to allow for appropriate protection of confidentiality in the field. Once an evaluation was uploaded to the server, the evaluation could not be accessed by other evaluators, ensuring anonymity. The administrator of the survey was able to remotely track incomplete surveys, which allowed for reminders and improvement in survey participation and completion. A summary of the evaluators' assessment, peer-assessments, and their self-evaluation were aggregated for each participant and delivered at the end of the SimEx.

Data Analysis

Data analysis using SPSS statistical software (SPSS-IBB SPSS Statistics 20, IBM Corporation; Armonk, New York USA) included statistical comparisons of competency scores and global

https://doi.org/10.1017/S1049023X1700005X Published online by Cambridge University Press

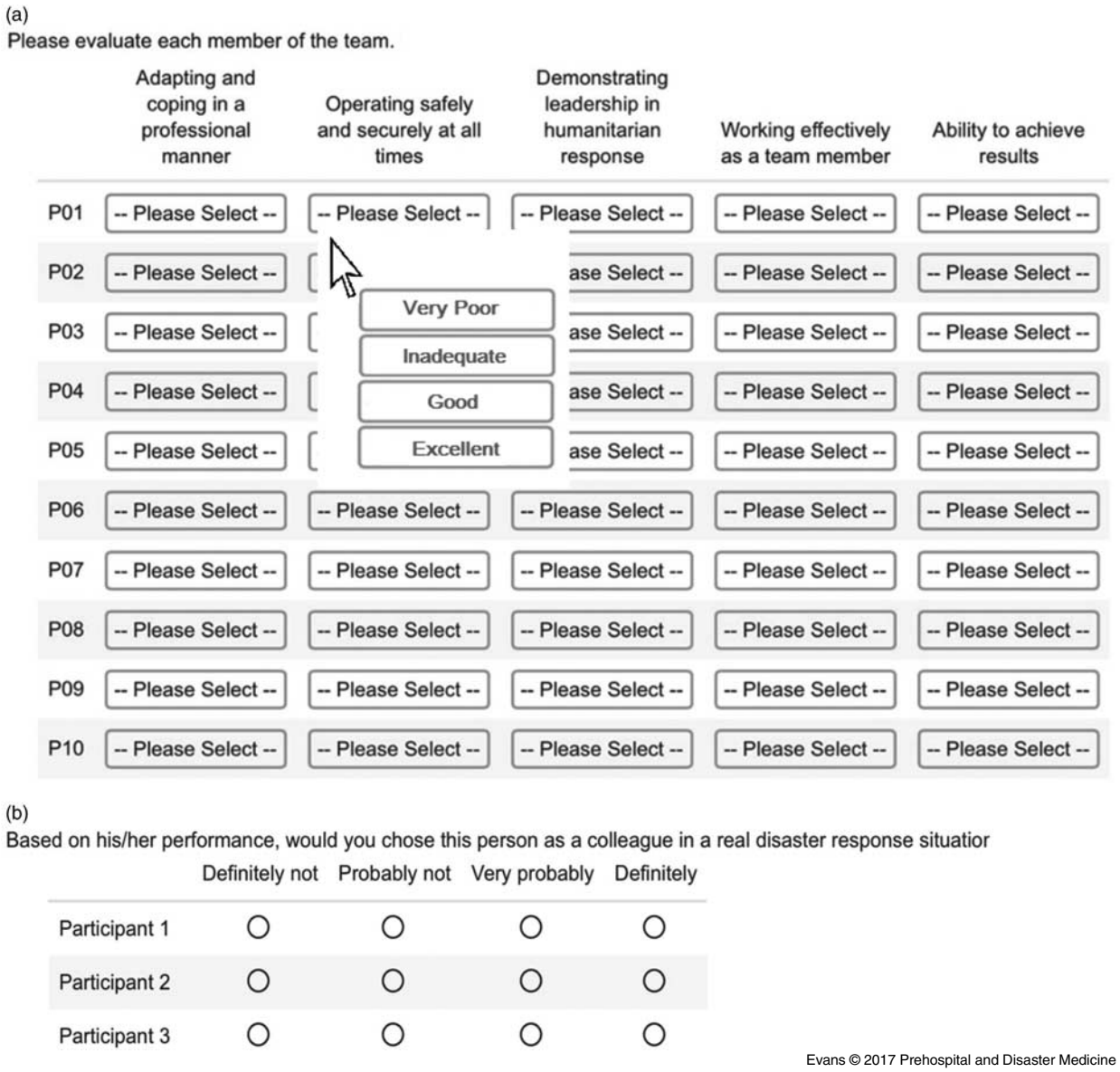

Figure 3. Offline Electronic Data Competency-Based Evaluation Tool as Presented on a Tablet.
(a) For each team of nine to 10 participants, five competencies were ranked based on participant performance using a Likert scale of 1-4: 1 = very poor, 2 = inadequate, 3 = good, 4 = excellent, or not applicable.
(b) A final question was asked as a global assessment for each participant.

rating scores between evaluator assessments, peer-evaluations, and self-evaluations. Comparisons were made using one-way ANOVA, and Levene's test was used to ensure equality of variance between groups. A paired t-test compared competency ratings between the first and second day of the simulation. Results were considered significant when the probability of making a Type I error was less than five percent ($P < .05$).

Results

A total of 63 participants enrolled in, and completed, all three days of the SimEx. There were seven teams, each comprised of eight to 10 participants. Participant demographics are summarized in Table 1.

All student participants (63) were evaluated. A total of 1,008 individual evaluations were completed. There were 90 (9.0%) missing evaluations due to scheduling issues at inject stations or due to technical difficulties with the electronic data capture tool. All 63 participants evaluated themselves and each of their teammates using the same survey.

Electronic Evaluation Tool

The electronic survey tool was displayed on a smart phone or tablet as shown in Figure 3. The entire survey fit on the tablet screen. The evaluator could enter scores on several participants at a time, flipping between surveys. In order complete all team member evaluations at one inject station, evaluators had on average one minute per survey question, or six minutes to observe each individual.

Written feedback from the evaluators post-SimEx indicated that most (7/9) evaluators agreed that the tool was easy to use, the Likert scale was appropriate, and that they did not experience any technical difficulties in its application. Most (7/9) evaluators

https://doi.org/10.1017/S1049023X1700005X Published online by Cambridge University Press

	N = 63
Age (average)	29.0 (SD = 7.9)
Women (number)	39 (57.0%)
Degrees (n):	
Bachelors	44 (69.0%)
Masters	16 (25.3%)
MD	10 (15.9%)
Other (ie, RN, Paramedic, Lawyer, or Certification)	5 (7.9%)
Previous Experience:	
In Own Profession (average, years)	3.25 (SD = 1.3)
In Global Paid Work (N)	12 (19%)
Median (months)	5.5 (SD = 19.2)
In Global Volunteer Work (N)	27 (43%)
Median (months)	2.4 (SD = 5.5)
Previous Deployment for Emergency Response	9 (14%)
Previous Experience working with NGOs (N)	31 (49%)
Median (months)	9.9 (SD = 19.6)

Table 1. Participant Demographics
Abbreviations: MD, Medical Doctor; NGO, nongovernmental organization; RN, Registered Nurse.

reported that they had insufficient time to evaluate each individual. They indicated that smaller team sizes or longer scenarios at the SimEx inject stations might allow for more accurate individual evaluations. They felt the tool accurately reflected performance in the key humanitarian competency areas that were being evaluated. They also noted that the global rating score allowed them to provide what they perceived as an accurate overall reflection of the participant's interpersonal interactions and competencies related to having the participant as their colleague in the field.

Participant Performance

There was no significant association between age of the participant and performance scores. However, participants with a Medical Doctor (MD) degree scored significantly higher on all competencies as assessed by evaluators compared to participants without an MD ($P < .05$). Peers evaluated participants with an MD higher on competencies C1, C2, and C5 ($P < .05$), but there were no significant differences between self-evaluation scores between MDs and other participants. Furthermore, participants with humanitarian-related work experience of one month or more scored higher than participants without work experience on all competencies as assessed by evaluators ($P < .05$). Those participants with previous humanitarian experience scored higher on C2 (Operating safely and securely at all times) and C4 (Working effectively as a team member) as assessed by evaluators ($P < .05$).

Figure 4(a) shows a breakdown of participant scores by competency for evaluators, peers, and self-evaluations. For all competencies and the global rating score, the average evaluator scores were significantly lower than the peer-evaluations and the self-evaluations ($P < .05$). Similarly, for all competencies, self-evaluations were lower than the peer-evaluations; these differences were statistically significant ($P < .05$) for competencies C1 and C3.

Figure 4(b) displays the breakdown of Likert scale scores by evaluators. The evaluators' average Likert scale ratings for C1, C3, C4, and C5 were significantly higher than the global rating score ($P < .01$). For peer evaluations, only C2 was not significantly different from the global rating score.

Evaluator assessment scores improved significantly from Day 1 to Day 2 for C1, "Adapting and coping in a professional manner" ($P < .05$). However, performance ratings decreased on Day 2 for both the global rating scale ($P < .05$) and C4, "Working effectively as a team member" ($P < .01$; Figure 4(c)).

Discussion

This study presents a novel, electronic evaluation tool to assess individual performance in five of six humanitarian competency domains in a 3-day humanitarian SimEx. The evaluation tool provides a standardized approach to the assessment of humanitarian competencies that are difficult to evaluate through knowledge-based testing in a classroom setting. It also permits the simultaneous evaluation of a number of participants working in teams in a dynamic, simulated, humanitarian emergency.

Since the development and general acceptance of the CBHA's, now Start Network's, Humanitarian Competency Framework in 2011,[32–34] "competency-based humanitarian training" has become a catchphrase in the humanitarian sector. The competencies in the

https://doi.org/10.1017/S1049023X1700005X Published online by Cambridge University Press

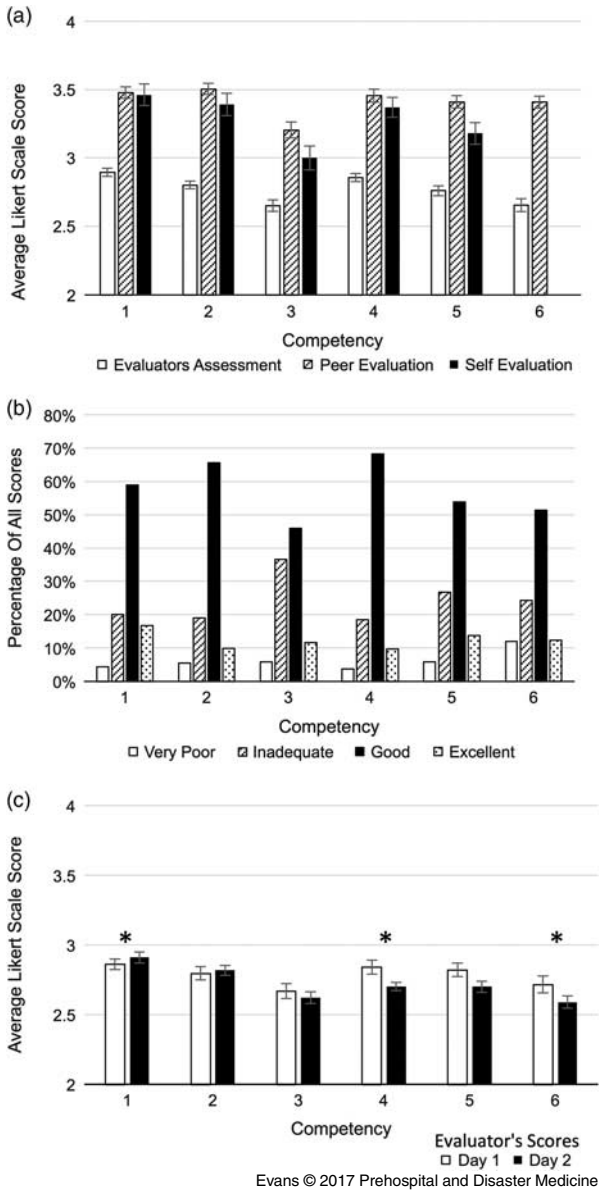

Figure 4. (a) Summary of Evaluators Scores by Competency. **(b)** Summary of Self, Teammate, and Evaluator's Assessments. **(c)** Scores by Evaluators Day 1 Compared to Day 2 of the Simulation Exercise.
Note: For C1 to C5, Likert Scale Score 1 = very poor, 2 = inadequate, 3 = good, 4 = excellent. For the global rating score, 1 = definitely not, 2 = probably not, 3 = very probably, 4 = definitely. C1 - Adapting and coping in a professional manner. C2 - Operating safely and securely at all times. C3 - Demonstrating leadership in humanitarian response. C4 - Working effectively as a team member. C5 - Ability to achieve results. The global rating score question was "Based on his/her performance, would you choose this person as a colleague in a real disaster response situation?"
* Indicates difference is statistically significant $P < .05$.

Start Network Framework are divided amongst six domains with several sub-categories in each. These domains are:

1. Understanding humanitarian contexts and applying humanitarian principles;
2. Achieving results;
3. Developing and maintaining collaborative relationships;
4. Operating safely and securely at all times;
5. Managing yourself in a pressured and changing environment; and
6. Demonstrating leadership in humanitarian response.

The competencies presented in the Framework originally were intended to serve as benchmarks for self-assessment.[28,32-36] Subsequently, the Framework's competencies were incorporated into some humanitarian training programs that then represented themselves as being "competency-based."[36,37] Although the Framework presents a list of competencies, it lacks specific guidelines on what is required to demonstrate competencies and how to assess achievement of competencies. Most competencies listed are difficult to assess using a traditional knowledge-based testing approach. Currently, there is no standard approach to the evaluation of humanitarian competencies in humanitarian training programs. Furthermore, most "competency- based" training programs do not include an evaluation component and only issue a "certificate of completion" based on attendance. This is especially the case for those programs that include simulation, partly due to the difficulty in evaluating participants in dynamic settings. The evaluation tool developed and tested here presents a method to evaluate and measure the Framework's competencies outside of the classroom, in a simulated setting. When combined with a knowledge-based test in the classroom, this presents an approach to a comprehensive, competency-based assessment approach that provides a relatively objective measurement of competency with respect to the competencies listed in the Framework. This results in an evaluation method that can benchmark student performance individually and when compared to peers. It also provides an indication of suitability for deployment in a humanitarian setting.

Outcomes of the evaluation tool presented here showed a significant difference between those students who had professional backgrounds, such as a medical degree, compared to non-professional backgrounds. This may be explained by the fact that physicians have more exposure to simulation-based, communication skills and an organized approach to problem solving.

Physicians from this cohort also were more likely to report work experience of one month or more, or prior experience in a global health setting, both of which correlated with higher performance.

Over the course of the 3-day SimEx, some scores improved significantly for some of the competencies, suggesting that some competencies are learned and practiced in a simulated setting and improve over time. This raises the question of the number of times an individual must perform an act before they are rated as being "competent." The medical literature suggests anywhere from one to 60 times, depending on the competency.[38,39] For example, the Accreditation Council for Graduate Medical Education (ACGME; Chicago, Illinois USA) requires surgeons to perform basic laparoscopy 60 times prior to graduating.[39-40] However, this relates to mastery of this skill through repetition rather than competence.[41-44] There seems to be a general agreement in the medical accreditation community that competence can come after any number of repetitions, and even before proficiency or mastery is achieved.[38]

Other competency scores decreased over the duration of the SimEx, suggesting that evaluator ratings may take time to become accurate. Evaluator assessments may have been too high on the first day, and once evaluators observed many learners, they

https://doi.org/10.1017/S1049023X1700005X Published online by Cambridge University Press

adjusted their application of the Likert scale. Conversely, this decrease may indicate that once students became fatigued as the SimEx went on, they became less competent. This highlights the need for evaluation of competency over a period of time and at different times during the SimEx. More study is required to understand the number of times that a competency must be demonstrated and the length of time that students must be observed in a SimEx setting to arrive at the most reliable score. Different competencies may require more or fewer demonstrations and evaluators may require more standardization in their interpretation of competency. Finally, intra-rater and inter-rater reliability must be factored into the competency scores.

Scores for the global rating scale question "Would you work with this person in the field" by both evaluators and peers were significantly lower than scores for the competencies rated. This is in contrast to the literature that finds global rating scales to correlate well with detailed evaluations of technical skills.[23] It may be that the competencies chosen for assessment in this study did not present a realistic measure of the individual's competency for humanitarian fieldwork. For example, although C1 broadly attempts to evaluate professionalism, there are limitations that current competency-based evaluation systems have in taking into account directly the preparation of deployed aid workers to interact adequately with local contexts and cultures.[45]

As the humanitarian sector looks towards professionalization, including training and certification, the utility of the evaluation tool presented here is considerable. For example, the World Health Organization (WHO; Geneva, Switzerland) has recently released guidelines for Emergency Medical Response Teams, stating that in order for teams to be included on the roster, team members must have training.[46] Competency-based training that incorporates classroom-based and SimEx components, and that include an evaluation method, will gain importance for not only personal performance bench-marking, but also for team and organizational deployment. Real-time evaluation of humanitarian workers while engaging in humanitarian relief would improve the timeliness of monitoring and evaluation of disaster and relief response. It also would highlight the more urgent areas for human research training and support.[8,10] More broadly, competency-based evaluation potentially offers accountability to the global community that training and simulation alone cannot guarantee.

Limitations

Since this is the first experience of using this competency-based evaluation tool in a humanitarian SimEx, this study does not ensure the tool is generalizable to real-time evaluations of humanitarian action, nor to other SimEx. However, with greater use of this competency-based evaluation tool, its reliability and transferability to other situations can be studied.

Team sizes of eight to 10 pose challenges to accurately evaluate individuals at injects since large team size allows for non-participation by some individuals, thus making it harder to assess them. Ideally, teams would be smaller, in groups of four to six, allowing each evaluator 10 minutes of observation by evaluators per individual.

Competency ratings did not account for the amount of participation by individuals, but rather focused on performance alone. Further study is required to determine the most effective amount of time required to observe one individual effectively and the ideal number of evaluations to provide a credible assessment of the performance of each student.

The survey tool platform required an electronic survey application to be downloaded via a Weblink when Internet access was available. This extra step required planning and coordination on multiple devices. Using electronic devices for survey data capture allows for more accurate and complete data analysis due to constraints that can be added by the designer. They also can be instantaneously modified in the field. However, electronics require sufficient charge, and battery power or battery packs must be provided. The offline survey platform also encountered multiple complicated steps to edit previously submitted evaluations. In future studies, to overcome these challenges, other platforms or higher-end versions of the survey tool should be used for easy retrospective editing.

Conclusion

This report presents a competency-based assessment of potential and practicing humanitarians using a novel electronic evaluation tool in a field-based SimEx. It allows the generation of multiple data points on observed competencies using a globally recognized humanitarian competency Framework that can be aggregated on the spot, providing an immediate benchmark of student performance. There is an opportunity to advance the use of this tool in future humanitarian training exercises and potentially in real time, in the field. This could impact the efficiency and effectiveness of humanitarian operations.

References

1. United Nations Office of Humanitarian Affairs. Global Humanitarian Overview 2016: A consolidated appeal to support people affected by disaster and conflict. Published 2016. https://www.humanitarianresponse.info/fr/programme-cycle/space/document/global- humanitarian-overview-2016-consolidated-appeal-support-people. Accessed January 28, 2016.
2. United Nations Population Fund. State of the World's Population Report 2015. Published 2015. http://www.unfpa.org/sites/default/files/sowp/downloads/State_of_World_Population_2015_EN. pdf. Accessed January 28, 2016.
3. Department of International Development -United Kingdom. Humanitarian Emergency Response Review. Published 2011. https://www.gov.uk/government/uploads/ system/uploads/attachment_data/file/67579/HERR.pdf. Accessed January 4, 2016.
4. United Nations Department of Social and Economic Affairs. World Urbanization Prospects: The 2014 Revision Highlights. Published 2014. http://esa.un.org/unpd/wup/Highlights/WUP2014-Highlights.pdf. Accessed January 28, 2016.
5. Active Learning Network for Accountability and Performance in Humanitarian Action. The State of the Humanitarian System in 2015 Report. Published 2015. http://www.alnap.org/resource/21036.aspx. Accessed January 28, 2016.
6. Inter-Agency Standing Committee. IASC is the United Nations' primary mechanism for interagency coordination of humanitarian assistance. https://interagencystanding-committee.org/. Accessed, January 28, 2016.
7. World Health Organization. Health Emergency Risk Management Framework and Improving Public Health Preparedness. Published 2012. http://apps.who.int/iris/bitstream/10665/78359/1/WHO_HSE_HEA_HSP_2013.1_eng.pdf. Accessed January 28, 2016.
8. Cranmer H, Chan J, Kayden S, et al. Pathway to professionalization: a competency-based evaluation of humanitarian aid workforce personnel during a humanitarian crisis simulation exercise. *Prehosp Disaster Med*. 2014;29(1).
9. Johnson K, Idzerda L, Baras R, et al. Standardized training for humanitarian workers: the first step toward professionalization of the humanitarian community. *Disaster Med Public Health Prep*. 2013;7(4):369-372.
10. Enhancing Learning and Research for Humanitarian Assistance. Professionalizing the humanitarian sector: a scoping study. http://www.elrha.org/wp-content/uploads/2015/01/Professionalising_the_humanitarian_sector.pdf. Accessed January 4, 2016.
11. Enhancing Learning and Research for Humanitarian Assistance. Global Survey on Humanitarian Professionalization. Published 2014. http://euhap.eu/upload/2014/06/global-humanitarian- professionalisation-survey.pdf. Accessed January 4, 2016.
12. Okuda Y. The utility of simulation in medical education: what is the evidence? *Mt Sinai J Med*. 2009;76(4):330.
13. Public Health Emergency Pre-Deployment Training in Tunisia. World Health Organization Regional Office for the Eastern Mediterranean. http://www.emro.who.int/surveillance-forecasting-response/surveillance-events/public-health-emergency-pre-deployment-training-in- the-region.html. Accessed January 28, 2016.

14. Humanitarian Response Intensive Course. Harvard Humanitarian Initiative. http://hhi.harvard.edu/education/workshops/hric. Accessed January 28, 2016.
15. Lateef F. Simulation-based learning: just like the real thing. *J Emerg Trauma Shock*. 2010;3(4):348-352.
16. Canadian Disaster and Humanitarian Response Training Program. Humanitarian Studies Initiative, McGill University. https://www.mcgill.ca/familymed/global-health/courses/hsir/disaster-and-humanitarian-response-program. Accessed January 28, 2016.
17. Start Network. Core Humanitarian Competencies Framework. Published 2012. file:///Users/melaniecoutu/Downloads/CBHA-Core-%20Humanitarian-%20Competencies-Framework-colour.pdf. Accessed January 28, 2016.
18. Karpicke JD, Blunt JR. Retrieval practice produces more learning than elaborative studying with concept mapping. *Science*. 2011;331(6018):772-775.
19. World Health Organization. Health Action in Crises: Primary Health Care in Crises. Published 2009. http://www.who.int/hac/about/annual_report/annual_report_2008.pdf. Accessed January 28, 2016.
20. Savoldelli GL, Naik VN, Hamstra SJ, Morgan PJ. Barriers to use of simulation-based education. *Can J Anaesth*. 2005;52(9):944-950.
21. Welke TM, LeBlanc VR, Savoldelli GL, et al. Personalized oral debriefing versus standardized multimedia instruction after patient crisis simulation. *Anesth Analg*. 2009;109(1):183-189.
22. Larsen DP, Butler AC, Roediger HL. 3rd, Test-enhanced learning in medical education. *Med Educ*. 2008;42(10):959-966.
23. Roediger HL, Karpicke JD. The power of testing memory. Perspectives on psychological. *Science*. 2006;1(3):181-210.
24. Roediger HL, Karpicke JD. Test-enhanced learning: taking memory tests improves long-term retention. *Psychol Sci*. 2006;17(3):249-255.
25. Bradt DA, Drummond CM. Professionalization of disaster medicine—an appraisal of criterion-referenced qualifications. *Prehosp Disaster Med*. 2007;22(5):360-368.
26. University of Laval program: Maîtrise en sciences de l'administration - gestion du développement international et de l'action humanitaire. https://www2.ulaval.ca/international/pole-dexcellence-sur-linternational/programmes-detudes/developpement-international-et-action-humanitaire.html. Accessed January 22, 2016.
27. McGill Humanitarian Studies Initiative (HSI). http://www.mcgill.ca/familymed/global-health/courses/hsir. Accessed January 22, 2016.
28. Humanitarian U Core Professional Humanitarian Training Program. www.humanitarianu.com. Accessed January 22, 2016.
29. Kamper SJ, Maher CG, Mackay G. Global rating of change scales: a review of strengths and weaknesses and considerations for design. *J Man Manip Ther*. 2009;17(3):163-170.
30. Egleston BL, Miller SM, Meropol NJ. The impact of misclassification due to survey response fatigue on estimation and identifiability of treatment effects. *Stat Med*. 2011;30(30):3560-3572.
31. Cosgrave J, Ramalingam B, Beck T. Real-time evaluations of humanitarian action: an ALNAP guide. ALNAP. http://www.alnap.org/resource/5595. Accessed January 21, 2016.
32. Consortium of British Humanitarian Agencies. Evaluation of Consortium of British Humanitarian Agencies pilot. Published 2012. file:///C:/Users/ejohn/Downloads/1448.pdf. Accessed January 23, 2016.
33. Consortium of British Humanitarian Agencies. Context Humanitarian Staff Development Project Program Guide. http://www.contextproject.org/index.php. Accessed January 23, 2016.
34. Consortium of British Humanitarian Agencies. Humanitarian Capacity Building Program: Objective 1 Final Report. Published 2010. http://www.start-network.org/wp-content/uploads/2014/01/Humanitarian-Capacity-Building- Programme-%E2%80%93-Objective-1-Final-Report.pdf. Accessed January 23, 2016.
35. Johnson K. Professionalizing Humanitarian Action. In: Caroline Abu-Sada, (ed). *Dilemmas, Challenges, and Ethics of Humanitarian Action: Reflections on Médecins Sans Frontières' Perception Project*. Montreal, Quebec, Canada: McGill-Queen's University Press; 2012: 104-115.
36. Disaster Ready courses. www.disasterready.org. Accessed January 24, 2016.
37. LaRoche Global Development and Humanitarian Aid Training. http://www.laroche.edu/International_Outreach/Global_Development_and_Humanitarian_Aid_T raining/?gclid=CLS6q4_sgcsCFVRbhgodbEkNsw. Accessed January 24, 2016.
38. Lenchus JD. End of the "See One, Do One, Teach One" era: the next generation of invasive bedside procedural instruction. *J Am Osteopath Assoc*. 2010;110(6):340-346.
39. Hasson HM. Core competency in laparoendoscopic surgery. *J Soc Lap Surg*. 2006;10(1):16-20.
40. Accreditation Council for Graduate Medical Education. https://www.acgme.org/acgmeweb/. Accessed January 23, 2016.
41. McFadden CL, Cobb WS, Lokey LS, Cull DL, Smith DE, Taylor SM. The impact of a formal minimally invasive service on the residents' ability to achieve new ACGME guidelines for laparoscopy. *J Surg Educ*. 2007;64(6):420-423.
42. Dulan G, Rege RV, Hogg DC. Proficiency-based training for robotic surgery: construct validity, workload, and expert levels for nine inanimate exercises. *Sure Endosc*. 2012;26(6):1516-1521.
43. McGaghie WC, Issenberg B, Cohen E, Jeffrey B, Wayne D. Does simulation-based medical education with deliberate practice yield better results than traditional clinical education? A meta-analytic comparative review of the evidence. *Acad Med*. 2011;86(6):706-711.
44. Ericsson KA. Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains. *Acad Med*. 2005;79(10):70-81.
45. Eichbaum Q. The problem with competencies in global health education. *Acad Med*. 2015;90(4):414-417.
46. Humanitarian Health Action: Emergency Medical Teams and World Health Organization. World Health Organization. http://www.who.int/hac/techguidance/preparedness/emergency_medical_teams/en/. Accessed January 23, 2016.