



Times of Medical Education Crisis Require New Evaluation Approaches: Proof of Concept of a System-Based Program Evaluation Model

Jorie Colbert-Getz¹ · Candace Chow¹ · Marlana Li² · Tiffany Weber³ · Peter Hannon⁴ · Ibrahim Hammad³ · Brian Good⁵ · Rachel Bonnett⁶ · Mariah Sakaeda⁶ · Jacob Robson⁵ · Sara Lamb¹

Accepted: 10 March 2021 / Published online: 17 March 2021
© International Association of Medical Science Educators 2021

Abstract

There are no program evaluation approaches designed for a crisis, such as the COVID-19 pandemic. It is critical to evaluate the educational impact of COVID-19 to keep administrators informed and guide decision-making. The authors used systems thinking to design an evaluation model. The evaluation results suggest complex interactions between individuals and course level changes due to COVID-19. Specifically, year 1–2 students found more education metrics lacking relative to year 3–4 students, faculty, and course directors. There was no consensus for the value of similar instructional/assessment adaptations. The evaluation model can be adapted by other medical schools to fit systems-based needs.

Keywords Program evaluation · Systems thinking · Crisis

Program evaluation seeks to gather information about the value or merit of a series of activities or events in a systematic way [1] Common program evaluation approaches used in medical education, like Kirkpatrick's hierarchy and the logic model, assume a linear relationship between activities/changes and satisfaction/outcomes, and thus work best for simple programs. Evaluating events in a linear way limits understanding of how multiple factors interact and affect outcomes and do not account for unanticipated issues and macro-level changes [2, 3] and therefore do not work for complicated or complex programs. To handle complexity,

other fields have used systems thinking to design evaluation approaches. [2–5]

The COVID-19 pandemic highlighted complexity in medical education and provides a good illustration for how systems thinking can be applied for evaluation. A key applicable system-thinking feature is the ability to account for the interplay of factors during constant change [2] It is critical to evaluate the inter-relations of factors to inform future decision-making during the pandemic and in preparation for future acute crises. There is little research on how medical education was impacted by prior pandemics and natural disasters, aside from perspective articles [6–8] Two additional articles investigated the impact of Hurricane Katrina on academic performance, but both lacked an evaluation framework. [9, 10] The purpose of this article is to provide a proof of concept for a systems-based evaluation model of the education impact of COVID-19.

Activity

Systems thinking was used to create a model (Fig. 1) for evaluating the educational impact of COVID-19 at the University of Utah School of Medicine during academic years 2019–2020. We identified three layers in the model: individuals (e.g., students, course directors, faculty), changes to instruction and assessment (e.g., shift to virtual lectures and

✉ Jorie Colbert-Getz
Jorie.colbert-getz@hsc.utah.edu

¹ Department of Internal Medicine, University of Utah School of Medicine, Salt Lake City, UT, USA
² Department of Family and Preventative Medicine, University of Utah School of Medicine, Salt Lake City, UT, USA
³ Department of Obstetrics and Gynecology, University of Utah School of Medicine, Salt Lake City, UT, USA
⁴ Department of Neurology, University of Utah School of Medicine, Salt Lake City, UT, USA
⁵ Department of Pediatrics, University of Utah School of Medicine, Salt Lake City, UT, USA
⁶ Office of Education Quality Improvement, University of Utah School of Medicine, Salt Lake City, UT, USA

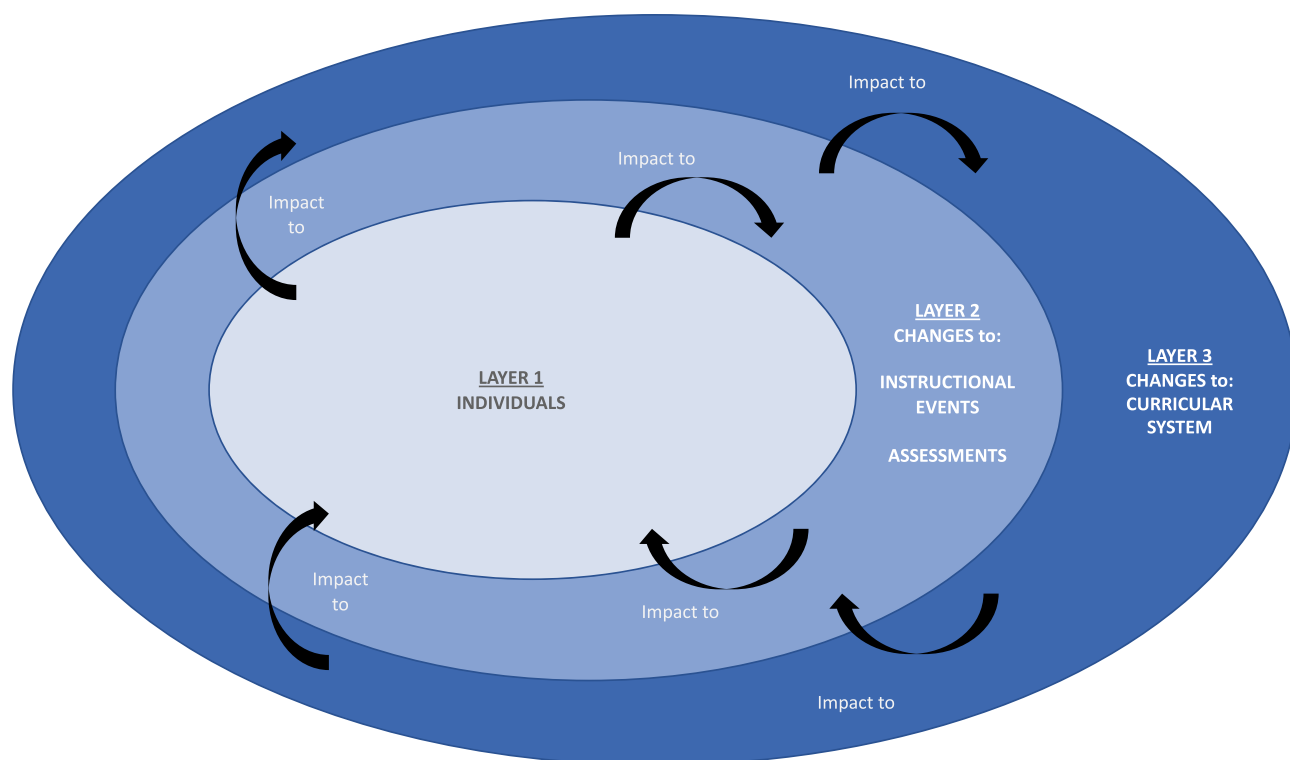


Fig. 1 A systems-based evaluation model for the educational impact of the COVID-19 pandemic

OSCEs), and system level curricular changes (e.g., graduation requirements). The evaluation model was shared with education leaders, refined, and implemented 1 week later. Specifics of layers 1 and 2 were developed and refined with input from course directors.

To evaluate layer 1, we added questions about communication, sense of community, resources/support, and wellness to end-of-course evaluations and end-of-year faculty surveys. Survey items were developed and refined through pilot testing. To evaluate layer 2, we used a value-driven outcomes framework [11] to capture the relative cost, time, and preference for COVID-19 adaptations relative to original in-person versions. We used instructional event type (e.g., small group discussion) and assessment type as the unit of analysis. Layer 2 information helped us understand which adaptations should be continued. Finally, evaluating layer 3, the overall learning impact and cost, will take place after the acute phase of the pandemic is over. Thus, this article will only report data from layers 1 and 2.

The University of Utah School of Medicine Institutional Review Board deemed this study exempt. For layer 1, individuals rated if each communication, sense of community, resources, and wellness survey item in Table 1 was exceptional, adequate, or lacking. A threshold of >24% lacking indicated concern for the Curriculum Committee. For layer

2, course directors indicated the time (more, less, the same), cost (more, less, the same), and their preference for each adaptation relative to the original in-person version. Students indicated if they preferred in-person, synchronous, or asynchronous learning. Students could also indicate no preference, or that it depended on the topic. Student preference for original vs. adapted was determined with a 70% majority threshold (combined with percent no-preference/depends on topic). An overall value determination (high, moderate, low, or no added value) for each instruction event type and assessment type at the course level was based on the number of metrics (cost, time, course director preference, student preference) favoring the adapted version. Instruction events/assessments with at least three value metrics favoring the adaptation were considered high value, two value metrics favoring the adaptation were deemed moderate value, and one value metric favoring the adaptation was considered low value.

Results

Response rates were above 90% for students and course directors and 71% for faculty. Table 1 provides the percent of students, course directors, and faculty who rated each

Table 1 Percent of University of Utah School of medicine students, course directors, and course faculty perceiving education metrics were lacking during COVID-19 in academic year 2019–2020

| Education metric | Year 1 students 97% (121/126) response rate | Year 2 students 97% (123/127) response rate | Year 3 students 99% (105/106) response rate | Year 4 student 98% (122/124) response rate | Course direc- tor 92% (25/28) response rate | Course faculty 71% (111/156) response rate |
|---|---|---|---|--|---|--|
| Communication items | | | | | | |
| Communication from Dean's Office Leadership | 6% (7) | 32% (39) | 7% (7) | 4% (5) | 0% | 2% (2) |
| Two-way communication with Dean's Office Leadership | 7% (6) | 30% (31) | 2% (2) | 19% (23) | 4% (1) | 2% (2) |
| Sense of Community Items | | | | | | |
| Sense of community with other students in course | 34% (41) | 28% (33) | 9% (9) | 3% (4) | 21% (5) | 10% (11) |
| Sense of community with faculty in course | 20% (24) | 17% (21) | 5% (5) | 3% (4) | 9% (2) | 10% (11) |
| Resources | | | | | | |
| School of Medicine resources and support for course | 17% (20) | 15% (18) | 7% (7) | 11% (13) | 4% (1) | 2% (2) |
| Technology/WIFI at home for remote course sessions | 4% (5) | 19% (23) | 6% (6) | 1% (1) | 0% | 5% (6) |
| Personal space at home for remote course sessions | 25% (30) | 25% (30) | 10% (10) | 1% (1) | 8% (2) | 7% (8) |
| Wellness | | | | | | |
| Personal ability to preserve your mental health | 26% (32) | 27% (33) | 9% (9) | 3% (4) | 13% (3) | 10% (11) |

Excludes n/a which is why percentages can vary with the same *n* within a column

education metric item as lacking. Greater than 24% of year 2 students found all communication items, sense of community with faculty, personal space for remote sessions, and personal ability to preserve mental health to be lacking. Greater than 24% of year 1 students found sense of community with students, personal space at home for remote sessions, and personal ability to preserve mental health to be lacking. There were no items rated as lacking by 24% or more of year 3 students, year 4 students, course directors, or course faculty.

For the purposes of this paper, we have provided value metrics for two course types (Foundational Science course, clerkship) in Table 2. There were 36 instruction event types and 32 assessment type adaptations due to COVID-19. Almost half of the adaptations (48%, 33) did not have added value over the original versions, 34% (23) had low value,

13% (9) had moderate value, and 1% (1) had high value. A higher percentage of assessment adaptations were considered low-high value (63%, 20) relative to instructional event adaptations (38%, 13). There was a higher percentage of year 3–4 clerkship course adaptations considered low-high value (57%, 20) relative to years 1–2 course adaptations (42%, 13). Finally, there was no course director preference consensus on similar adaptations; student preference also varied for similar adaptations across courses.

Discussion

This is the first study of a systems-based evaluation model in a time of crisis. Similar to gathering feedback alone with surveys, level 1 data provided us with insight on how

Table 2 Instructional and assessment event adaptations due to COVID-19 for a year 1 foundational science course and year 3

| Instructional or Assessment Type | Original Version | COVID-19 Adaptation | Cost for Adapted version relative to Original In-Person Version | Course Director/Staff Time for Adapted version relative to Original In-Person Version | Course director preference for original or adapted version | > 70% Students preference for original or adapted (combined with no preference/depends on topic) | Overall value of adaptation for future courses after COVID-19 era ^a |
|--|----------------------------|---|---|---|--|--|--|
| Large Group Lectures in a Year 1 Course | In-person, recorded | Online via Zoom/ uploaded to YouTube (Synchronous & Asynchronous), Some Asynchronous only | No difference | No difference | Original | Original | No added value |
| Case Based Learning Sessions in a Year 1 Course | In-person with facilitator | Online via Zoom with facilitators (Synchronous) | More | No difference | Original | Original | No added value |
| Team Based Learning Sessions in a Year 1 Course | In-person, not recorded | Online via Zoom/ video chat (Synchronous) | No difference | More | Adaptation | Adaptation | Moderate value |
| Pathology/Microbiology Labs in a Year 1 Course | In-person | Remote via Zoom/ uploaded to YouTube, handout (Synchronous & Asynchronous) | No difference | More | Original | Original | No added value |
| Large Group Review Sessions in a Year 1 Course | In-person | Online via Zoom/ uploaded to YouTube (Synchronous & Asynchronous) | No difference | Less | Original | Original | Low value |
| Didactics in a Year 3 Clerkship | In-person | Recorded (Asynchronous) | No difference | Less | Adaptation | Adaptation | High value |
| Case Based Learning Sessions in a Year 3 Clerkship | In-person, Case Review | Online via Zoom with faculty (Synchronous) | No difference | More | Adaptation | Adaptation | Moderate value |
| Formative Quizzes in a Year 1 Course | In-person | At home (Asynchronous) | No difference | No difference | Adaptation | Not asked ^b | Low value |
| Final Exam in a Year 1 Course | In-person, proctored | At home, not proctored (Asynchronous) | Less | No difference | Original | Not asked ^b | Low value |
| Objective Structure Clinical Exam (OSCE) in a Year 3 Clerkship | In-person | Online via Zoom (Synchronous) | Less | No difference | Original | Original | Low value |

Table 2 (continued)

Clerkship in academic year 2019–2020 at the University of Utah School of Medicine

| Shelf Exam in a Year 3 Clerkship | In-person, Proctored | At home, proctored via Zoom (Synchronous) | No difference | More | Original | Original | No added value |
|-------------------------------------|----------------------|--|---------------|------|----------|----------|----------------|
|-------------------------------------|----------------------|--|---------------|------|----------|----------|----------------|

^aInstruction events and assessments with at least three value metrics favoring the adaptation were considered high value, two value metrics favoring the adaptation were moderate value, one value metric favoring the adaptation was low value, and no metrics favoring the adaptation was considered to have no added value

^bStudents were not asked about their preference for assessment adaptations in year 1–2 courses

to better support our students. Sense of community was challenging to foster, especially while simultaneously enforcing stay at home orders, physical distancing, and not wanting to overwhelm students with too many online meetings. For both students and faculty, sense of community was endangered by online learning and changes in social interaction. We also identified aspects that needed to be addressed beyond normal course support, such as having enough personal space and WIFI at home and the ability to tend to one’s mental health. Based on these findings, we arranged for on-campus building space so students who needed to complete a final examination outside of their home could do so.

Level 1 data also highlighted how differently year 1–2 and year 3–4 students experienced course changes which has made us more mindful of their varying educational environments, course-loads, and developmental needs. For instance, year 1–2 students interact in much larger groups than year 3–4 students and may therefore have felt the sudden loss of those interactions more acutely. Moreover, early in medical school peer, student-faculty relationships and a sense of community are intensely developing in comparison with latter stages when students have had years to form peer bonds, build a support network, and acclimate to the culture of the school. Additionally, year 1–2 students were in three courses at the same time and may have found increasing disruptions at home trying to study relative to year 3–4 students in one rotation during COVID-19.

Layer 2 data revealed no clear consensus on preference or perceived value of adaptations made by course directors, which at first glance seemed counterintuitive. However, the lack of consensus reminds us why a systems-thinking approach for evaluating educational activities, particularly during uncertain times, is so important. We suspect that perceptions vary greatly because of the constantly changing and differing contexts for students and faculty. Additionally, adaptations made in years 3–4 seemed to have more value than those in years 1–2, which could have impacted the overall receptibility and stress of students or vis-a-versa. Overall, there was much interplay between layers 1 and 2, demonstrating that our layer 2 data is inherently affected by our layer 1 data, underscoring why a linear evaluation model is less than ideal for this situation.

As others have outlined,[2–5] using systems-thinking principles is helpful in designing evaluations of complex educational situations.[12] To design our nimble approach, we used the lens of systems thinking from the onset, considered inter-relations of factors, and captured diverse stakeholder input. Given the novelty of the COVID-19 crisis, results from this case study provide insight for other medical schools. There is potential for transferability of the systems-based model to other schools because it can be adapted to their needs. We expect to see more studies

utilizing systems-based evaluation approaches in the future since the landscape of medical education is becoming more complex.

Declarations

Ethical Approval The University of Utah School of Medicine deemed this study exempt, Exemption Umbrella: Assessment of Medical Education within the MD Curriculum, IRB# 00109278.

References

1. Patton MQ, Asibey E, Dean-Coffey J, Kelley R, Miranda R, Parker S, Fariss Newman G. American Evaluation Association 2014 What is Evaluation? Statement Accessed June , 2020. <http://www.eval.org/p/bl/et/blogid=2&blogaid=4>
2. Dyehouse M, Bennett D, Harbor J, Childress A, Dark M. A comparison of linear and systems thinking approaches for program evaluation illustrated using the Indiana Interdisciplinary GK-12. *Eval Program Plann.* 2009;32(3):187–96.
3. Gates EF. Learning from seasoned evaluators: implications of systems approaches for evaluation practice. *Evaluation.* 2017;23(2):152–71.
4. Patton MQ. What is essential in developmental evaluation? On integrity, fidelity, adultery, abstinence, impotence, long-term commitment, integrity, and sensitivity in implementing evaluation models. *Am J of Eval.* 2016;37(2):250–65.
5. Mowles C. Complex, but not quite complex enough: the turn to the complexity sciences in evaluation scholarship. *Evaluation.* 2014;20(2):16–172.
6. Rieder MJ, Salvadori M, Bannister S, Kenyon C. Collateral damage: the effect of SARS on medical education. *Clinical Teacher.* 2004;1(2):85–9.
7. DiCarlo, , et al. Survival and recovery: maintaining the educational mission of the Louisiana state university school of medicine in the aftermath of hurricane Katrina. *Acad Med.* 2007;82(8):745–56.
8. de Arzola OR. Emergency preparedness and hurricane Maria: the experience of a regional academic medical center in southwest Puerto Rico. *J of Grad Med Ed.* 2018;10(4):477.
9. Krane NK, DiCarlo RP, Kahn MJ. Medical education in post-Katrina New Orleans: a story of survival and renewal. *JAMA.* 2007;298(9):1052–5.
10. Crawford BE, et al. Impact of hurricane Katrina on medical student academic performance: the Tulane experience. *The Am J of the Med Sci.* 2008;336(2):142–6.
11. Kawamoto K, Martin CJ, Williams K, Ming-Chieh T, Park CG, Hunter C, et al. Value driven outcomes (VDO): a pragmatic, modular, and extensible software framework for understanding and improving health care costs and outcomes. *J Am Med Inform Assoc.* 2015;22:223–35.
12. Systems in Evaluation Topical Interest Group of the American Evaluation Association. (2018). Principles for Effective Use of Systems Thinking in Evaluation Practice. Accessed February 21, 2021. <https://www.systemsinevaluation.com/wp-content/uploads/2018/10/SETIG-Principles-FINAL-DRAFT-2018-9-9.pdf>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.