Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Digital Simulation for Teaching Design Thinking to Medical Students: A Two-Cohort Pre–Post Evaluation

Anna Berbenyuk
Mohammed Bin Rashid University of Medicine and Health Sciences
Areeba Shahid
Mohammed Bin Rashid University of Medicine and Health Sciences
Liko Lu Qi
Mohammed Bin Rashid University of Medicine and Health Sciences
Nour El-Khatib
FAST4
Baptiste Roux
FAST4
Yacine Hadjiat
yacine.hadjiat@dubaihealth.ae

Mohammed Bin Rashid University of Medicine and Health Sciences

Research Article

Keywords: design thinking, medical education, digital simulation, experiential learning, curriculum innovation

Posted Date: November 16th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11268481/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: Competing interest reported. BR is the founder, and NE an employee, of the company that developed the DesignQuest platform evaluated in this study. The company was paid to co-design the simulations and develop the platform for use in the course. The remaining authors declare that they have no competing interests.

Abstract

Background

The growing complexity of healthcare requires clinicians who can collaborate across disciplines and develop human-centred solutions to complex problems. Design thinking (DT) offers a structured framework for developing these competencies, but its integration into medical education remains limited, and evidence on scalable methods for teaching it is sparse. This study evaluated a digital simulation platform for teaching DT in an undergraduate medical curriculum.

Methods

A pre–post evaluation was conducted in two consecutive cohorts of first-year students enrolled in the Innovation and Technologies in Health Sciences course at Mohammed Bin Rashid University of Medicine and Health Sciences. In Phase 1 (2024–2025), students completed three narrative-based simulations; in Phase 2 (2025–2026), students completed two simulations, with a third delivered as the final quiz. Outcomes were change in DT knowledge, engagement with the platform, and, in Phase 1, student satisfaction.

Results

In Phase 1 (n = 88), mean quiz scores fell from 93.6 to 82.5 (mean change − 10.9, 95% CI − 13.7 to − 8.1; P<.001), from a baseline close to the assessment ceiling. In Phase 2, among the 32 of 70 enrolled students who completed both assessments and at least one simulation, mean scores rose from 78.6 to 89.8 (mean change + 11.2, 95% CI 7.0 to 15.5; P<.001; d = 0.95). Improvement did not differ by number of simulations completed (P=.47). In Phase 1, students rated overall satisfaction at 4.05 and usefulness for learning DT at 4.17 on a five-point scale, while suggesting improvements to pacing and scenario realism.

Conclusions

Students found the digital simulations engaging and useful, but knowledge outcomes were inconsistent between cohorts. The Phase 1 decline is most plausibly explained by a ceiling effect, and the Phase 2 gain cannot be attributed to the simulations given the absence of a control group, the selected analytic sample, and the lack of a dose–response relationship. Digital simulation appears to be an acceptable and scalable format for DT teaching; controlled studies with assessments designed to capture DT competencies are needed to establish its effect on learning.

Background

Healthcare is growing more complex as technology, populations and patient expectations change, increasing the need for clinicians who can adapt, collaborate and develop human-centred solutions [1–4]. Traditional medical education, with its emphasis on theoretical knowledge and adherence to protocol, has been less effective at developing these capacities, and many institutions are revising their curricula to strengthen innovation, empathy and systems thinking alongside clinical competence [1, 2, 5, 6].

Design thinking (DT) is an iterative, human-centred approach to problem-solving that moves through empathy, problem definition, ideation, prototyping and testing [2, 7, 8, 11]. Originating in the design disciplines, it has been applied in healthcare to redesign patient experiences, improve clinical workflows and address problems that require a detailed understanding of user needs [2, 5, 7, 9, 10]. In medical and dental education, DT has been associated with greater learner engagement, critical thinking and collaborative problem-solving, consistent with constructivist and experiential learning theory [2, 12]. Its adoption in medical curricula nevertheless remains limited, typically confined to isolated workshops or elective modules [2, 5, 7]. Most existing interventions rely on in-person, project-based formats that are resource-intensive and difficult to scale [2, 7], and evidence on how DT can best be taught, particularly through digital modalities, is sparse [2, 5, 13, 14].

Digital simulation offers one possible solution. Well established in clinical skills training, simulation-based education bridges theory and practice and supports the development of complex competencies such as critical thinking and teamwork [1, 5, 6, 13, 18]. Current platforms allow interactive, narrative-driven scenarios in which learners work through each stage of the DT process in a safe, feedback-rich environment [1, 5, 6, 13, 15]. Such environments also support Schön's concept of reflection-in-action, encouraging learners to think adaptively as scenarios evolve [19]. Few studies, however, have evaluated whether digital simulation improves medical students' understanding of DT, and the literature has called for research examining knowledge, engagement and the scalability of such approaches [2, 5, 13, 21].

This study evaluated a digital simulation platform for teaching DT within the undergraduate medical curriculum at Mohammed Bin Rashid University of Medicine and Health Sciences (MBRU), across two consecutive student cohorts, assessing knowledge outcomes, engagement with the platform and student perceptions.

Methods

Study design and setting

This was a pre–post evaluation conducted in two consecutive cohorts at MBRU, Dubai, United Arab Emirates. The study was embedded in the Innovation and Technologies in Health Sciences course, which introduces first-year medical students to innovation and DT as approaches to healthcare challenges, and was delivered in the 2024–2025 (Phase 1) and 2025–2026 (Phase 2) academic years. Each participant served as their own comparator; randomization was not feasible within a single compulsory course.

The study had three aims: to assess change in DT knowledge between the start and end of the simulation period; to describe engagement with the platform and its relationship to knowledge gain; and to assess student satisfaction with the simulations.

Participants

All students enrolled in the course were eligible, and no exclusion criteria were applied. Phase 1 included the 88 students who completed both assessments and all three simulations. Of the 70 students enrolled in Phase 2, the paired analysis included the 32 who completed both the pre-course and final quizzes and at least one simulation.

A power analysis for a two-tailed paired t-test ($\alpha = .05$, power = 0.80) indicated that 90 participants would be required to detect a small within-subject effect ($d = 0.3$), and 41 to detect a moderate effect ($d = 0.45$). The Phase 1 sample approached the former; the Phase 2 analytic sample was below both thresholds.

Intervention

The intervention comprised narrative-based digital simulations delivered through the DesignQuest platform. Each simulation presented a healthcare challenge set in an interactive clinical vignette, requiring students to work through the five stages of DT (empathize, define, ideate, prototype and test). The simulations provided real-time feedback and allowed repeated attempts, so that students could revise decisions and observe their consequences.

In Phase 1, students completed three simulations (A, B and C) as scheduled course activities between Weeks 7 and 13, facilitated by faculty to ensure consistent delivery and integration with the wider DT curriculum. In Phase 2, the scenarios were adapted to reflect local healthcare challenges; students completed two simulations during the course, and a third was delivered as the final quiz.

Data collection

Knowledge was assessed at the start and end of the simulation period. In Phase 1, the pre-test (Week 7) and post-test (Week 13) each comprised 10 multiple-choice and short-answer questions aligned with the course learning outcomes. Different questions were used at each time point, drawn from the same question bank and selected to be of comparable difficulty; equivalence was not empirically verified. Both assessments contributed to in-course grades. In Phase 2, knowledge was assessed with a pre-course quiz and a final quiz delivered in simulation format.

Engagement data, comprising the number of simulations attempted, completed and repeated, were captured automatically by the platform. In Phase 1, a satisfaction survey was distributed after the final in-class presentations (Week 14), comprising Likert-scale items on usefulness, engagement and learning

value, and open-ended questions. No survey was administered in Phase 2. All data were anonymized before analysis.

Outcome measures

The primary outcome was change in DT knowledge, measured as the difference between pre- and post-intervention quiz scores. Secondary outcomes were engagement with the simulations, knowledge gain by number of simulations completed, and, in Phase 1, self-reported satisfaction.

Statistical analysis

Continuous variables were summarized as means and standard deviations. Within-participant change in quiz scores was assessed using paired t-tests, with effect sizes expressed as Cohen's d for paired data. In Phase 1, first-attempt scores were compared between simulations using pairwise tests; in Phase 2, improvement was compared between students completing one and two simulations using the Mann–Whitney U test. Open-ended survey responses were reviewed and grouped into recurring areas of feedback. Statistical significance was set at P<.05. Analyses were performed using Microsoft Excel and IBM SPSS Statistics version 29.

Ethical considerations

The study was conducted in accordance with the ethical principles of the World Medical Association Declaration of Helsinki. Ethical approval was obtained from the Institutional Review Board of Mohammed Bin Rashid University of Medicine and Health Sciences (MBRU IRB-2024-449). All students provided informed consent electronically. Data were anonymized using unique identifiers, and only aggregated findings are reported.

Results

Knowledge outcomes

In Phase 1, the mean quiz score fell from 93.6 (SD 5.7) before the simulations to 82.5 (SD 9.3) afterwards, a mean change of − 10.9 points (95% CI − 13.7 to − 8.1; P<.001). In Phase 2, the mean score rose from 78.6 to 89.8, a mean change of + 11.2 points (95% CI 7.0 to 15.5; t = 5.39; P<.001; d = 0.95). The median improvement in Phase 2 was 12.5 points, ranging from a decrease of 10 points to an increase of 35 points (Table 1).

Table 1
Knowledge outcomes in both phases

Pre-intervention score, mean (SD)	**Phase 1 (n = 88)**	**Phase 2 (n = 32)**
	93.6 (5.7)	**78.6**
Post-intervention score, mean (SD)	82.5 (9.3)	89.8
Mean change (95% CI)	−10.9 (− 13.7 to − 8.1)	+ 11.2 (7.0 to 15.5)
P value	< .001	< .001
Cohen's d	−0.82	0.95

CI: confidence interval; SD: standard deviation. Standard deviations were not available for Phase 2.

Engagement with the simulations

In Phase 1, students completed a mean of 1.22 attempts on Simulation A, 1.85 on Simulation B and 1.01 on Simulation C (Fig. 1). First-attempt scores differed between simulations, increasing from A to B (mean difference 3.2, 95% CI 0.5 to 5.9; P=.02), from B to C (4.4, 95% CI 1.0 to 7.9; P=.01), and from A to C (6.8, 95% CI 3.2 to 10.4; P<.001) (Table 2).

Table 2
Pairwise comparisons of first-attempt simulation scores, Phase 1

Comparison	**Mean difference**	**95% CI**	**P value**
Simulation B vs A	3.2	0.5 to 5.9	.02
Simulation C vs A	6.8	3.2 to 10.4	< .001
Simulation C vs B	4.4	1.0 to 7.9	.01

CI: confidence interval.

In Phase 2, 50 students completed Simulation 1 and 29 completed Simulation 2. Mean scores on each simulation are shown in Table 3. Because the two simulations were completed by different subsets of enrolled students, the comparison is cross-sectional and does not represent a within-student trajectory. No student scored below 50 on Simulation 2, compared with 16% on Simulation 1.

Table 3
Simulation performance, Phase 2

Simulation	**Students, n**	**Score, mean (SD)**
Simulation 1	50	67.8 (17.2)
Simulation 2	29	77.6 (11.8)

SD: standard deviation.

Knowledge gain by number of simulations

In Phase 2, 26 of the 32 students in the paired analysis (81%) completed one simulation and 6 (19%) completed two. Mean improvement was 11.9 points among those completing one simulation and 8.2 points among those completing two; the difference was not statistically significant (Mann–Whitney U = 62.5; P=.47) (Table 4).

Table 4
Knowledge gain by number of simulations completed, Phase 2

Simulations completed	Students, n (%)	Mean improvement, points
One	26 (81)	+ 11.9
Two	6 (19)	+ 8.2

Mann–Whitney U = 62.5; P=.47.

Student perceptions

In the Phase 1 survey, mean ratings on a five-point scale were 4.05 for overall satisfaction and 3.85 for engagement (Table 5). Clarity of instructions received the highest rating (4.29), followed by usefulness for understanding DT principles (4.17). Realism of scenarios and balance between theory and practice were each rated 3.97, and usefulness for the final course project 3.85.

Table 5
Student ratings of the digital simulations, Phase 1

Survey item	Rating, mean (SD)
Overall satisfaction with the simulation experience	4.05 (0.6)
Level of engagement during the activity	3.85 (0.7)
Clarity of instructions provided	4.29 (0.5)
Realism of simulation scenarios	3.97 (0.7)
Balance between theory and practice	3.97 (0.6)
Usefulness for understanding design thinking principles	4.17 (0.6)
Usefulness for completing the final course project	3.85 (0.7)

Rated on a five-point Likert scale (1 = strongly disagree; 5 = strongly agree). SD: standard deviation.

Open-ended responses addressed three areas. On the learning experience, students described the simulations as interactive and engaging and valued applying the DT stages to realistic scenarios in a guided environment; one wrote that “the feedback after each decision made it easier to learn.” On design and usability, most found the interface intuitive, although some reported that cases progressed slowly or

contained unrealistic options, and suggested additional visual cues or branching feedback. Suggestions for improvement included more detailed clinical contexts, a wider range of healthcare settings and faster navigation between stages. Most respondents favoured retaining the simulations in future iterations of the course.

Discussion

Principal findings

This study evaluated a digital simulation platform for teaching DT to first-year medical students across two cohorts. Students rated the simulations positively for clarity, engagement and usefulness, but knowledge outcomes were inconsistent: scores declined significantly in Phase 1 and rose significantly in Phase 2. Neither result can be attributed to the simulations with confidence, for different reasons in each cohort.

Interpreting the knowledge outcomes

The Phase 1 decline should not be read as a loss of learning. The pre-test mean of 93.6 lay close to the maximum score, leaving little room for improvement and substantial room for decline. Although pre- and post-test questions were drawn from the same bank and selected to be of comparable difficulty, their equivalence was not empirically verified, and with scores clustered near the ceiling, modest differences in item difficulty or random variation in the questions sampled would be expected to produce lower post-test scores through regression to the mean. A 10-question instrument is particularly susceptible to this effect, since each item contributes ten percentage points. The decline is therefore more plausibly an artefact of measurement than a consequence of the intervention, and it indicates that the pre-test was insufficiently discriminating for this cohort.

The Phase 2 gain of 11.2 points was large, but several features of the design prevent it from being attributed to the simulations. There was no control group, so the improvement cannot be separated from the effect of the wider course, which delivered DT teaching through other activities over the same period. The analysis included 32 of 70 enrolled students, selected on the basis of completing both assessments and at least one simulation; students who engaged with the course may have improved irrespective of the simulations. Both assessments contributed to grades, which may have motivated preparation for the final quiz. Most importantly, there was no dose–response relationship: students who completed two simulations improved by less than those who completed one, although the difference was not significant and the two-simulation group comprised only six students. Had the simulations been the main driver of improvement, greater exposure would be expected to produce greater gains.

The divergence between phases may also reflect differences between cohorts, including a baseline mean 15 points lower in Phase 2, and changes to the simulations and assessment format between years. Because the two phases differed in more than one respect, the contrast between them cannot be used to identify which change, if any, was responsible.

Engagement with the simulations

In Phase 1, first-attempt scores increased across the sequence from Simulation A to C, consistent with students becoming more familiar with the DT process and with the platform. Engagement, measured as repeated attempts, was highest for Simulation B and lowest for C. One possibility is that the content of Simulation C placed greater demands on students at a late stage of the course, consistent with cognitive load theory, although the present data do not allow this to be tested, and the higher first-attempt scores on Simulation C suggest that students may simply have needed fewer attempts to reach a satisfactory result.

In Phase 2, the mean score rose from 67.8 on Simulation 1 to 77.6 on Simulation 2, and no student scored below 50 on the second simulation. This resembles a learning effect, but because the two simulations were completed by different subsets of students, it cannot be interpreted as the progression of individual students and may reflect differences in who completed each simulation.

Student perceptions

The consistently positive ratings indicate that students found the simulations clear, engaging and relevant to learning DT, and the open-ended responses identified specific refinements to pacing, realism and navigation. These findings support the acceptability of the format, which is a prerequisite for adoption. Satisfaction does not, however, establish learning, and the contrast between positive perceptions and the Phase 1 decline illustrates the need to measure both.

Comparison with prior studies

Technology-enhanced simulation has been associated with improved knowledge and skills in health professions education when compared with no intervention [3], and narrative and virtual simulation are increasingly used in undergraduate medical training [13]. Reviews have, however, noted that the effects of simulation depend on design features such as feedback, repetition and debriefing, and that much of the evidence comes from uncontrolled studies [15, 16]. In the present study, positive perceptions were likewise not consistently accompanied by objective gains. Within DT education specifically, evidence remains limited and largely descriptive [2, 7, 12].

Implications

Two implications follow. First, the assessment of DT should be aligned with its aims. A short knowledge quiz captures familiarity with the DT stages but not the competencies the method is intended to develop, such as empathy, problem framing and iterative reasoning. Reflective assignments, portfolio-based assessment, structured observation of team-based challenges and analysis of decision-making within the simulations themselves may provide more valid measures [17, 20]. Pre-tests should also be calibrated to avoid ceiling effects.

Second, the simulations' acceptability and capacity for delivery to large cohorts make them a practical candidate for scaling DT teaching. Establishing whether they improve learning will require designs that

can separate their effect from that of the wider curriculum, such as comparisons between simulation and non-simulation groups or stepped introduction across cohorts, together with structured debriefing to support reflection on the experience [20].

Limitations

The study used a pre–post design without a control group in both phases, which prevents attribution of change in knowledge to the simulations. The Phase 2 paired analysis included fewer than half of enrolled students, introducing potential selection bias, and was underpowered. The knowledge assessment comprised 10 questions, showed a ceiling effect in Phase 1, and had not been validated psychometrically, and pre- and post-test equivalence was not empirically verified. The assessment format and simulation content differed between phases, limiting comparison between cohorts. Satisfaction data were available only for Phase 1. Finally, the study was conducted at a single institution, and the findings may not generalize to other curricula or settings.

Conclusions

Across two cohorts, first-year medical students found digital simulations for teaching DT clear, engaging and useful, supporting the acceptability of the approach. Knowledge outcomes were inconsistent: the Phase 1 decline most plausibly reflects a ceiling effect in a short assessment, and the Phase 2 gain, while large, cannot be attributed to the simulations given the absence of a control group, the selected analytic sample and the lack of a dose–response relationship. Digital simulation is a practical and acceptable format for delivering DT teaching at scale, but its effect on learning remains to be established. Future studies should use controlled designs and assessment methods that capture DT competencies rather than recall of the DT process.

Abbreviations

CI, confidence interval; DT, design thinking; MBRU, Mohammed Bin Rashid University of Medicine and Health Sciences; SD, standard deviation

Declarations

Ethics approval and consent to participate: The study was conducted in accordance with the ethical principles of the World Medical Association Declaration of Helsinki. Ethical approval was obtained from the Institutional Review Board of Mohammed Bin Rashid University of Medicine and Health Sciences (MBRU IRB-2024-449). All students provided informed consent electronically.

Consent for publication: Not applicable.

Availability of data and materials: The datasets generated and analysed during the current study are available from the corresponding author on reasonable request.

Competing interests: BR is the founder, and NE an employee, of the company that developed the DesignQuest platform evaluated in this study. The company was paid to co-design the simulations and develop the platform for use in the course. The remaining authors declare that they have no competing interests.

Funding: This study was funded by Dubai Health Innovations. The funder had no role in study design, data analysis or interpretation, or the decision to publish.

Authors' contributions: AB and YH designed the Phase 1 study, and AB managed operations. AS contributed to the design and operations of Phase 2. BR and NE contributed to the design of the simulations, developed the DesignQuest platform and conducted the data analysis. All authors reviewed and approved the final manuscript.

Acknowledgements: We thank Liko Lu Qi, MBRU medical student, for her support of the project.

References

1. Romancenco A, Saratila I, Ababii I, Rojnoveanu G, Dandara O, Spinei L. Bridging theory and practice: enhancing medical education through simulation-based training methods. Mold J Health Sci. 2024. https://doi.org/10.52645/mjhs.2024.2.09.
2. Chuenjitwongsa S, Amir L, Jessani A, Samaranayake L, Osathanon T. Integrating design thinking into dental education. Front Oral Health. 2025;6. https://doi.org/10.3389/froh.2025.1547335.
3. Cook DA, Hatala R, Brydges R, Zendejas B, Szostek JH, Wang AT, Erwin PJ, Hamstra SJ. Technology-enhanced simulation for health professions education: a systematic review and meta-analysis. JAMA. 2011;306(9):978–88. https://doi.org/10.1001/jama.2011.1234.
4. Brown T, Wyatt J. Design thinking for social innovation. Stanf Soc Innov Rev. 2010;8(1):30–5. https://doi.org/10.48558/58Z7-3J85.
5. Cant R, Cooper S. Use of simulation-based learning in undergraduate nurse education: an umbrella systematic review. Nurse Educ Today. 2017;49. https://doi.org/10.1016/j.nedt.2016.11.015.
6. Kaddoura R, Faraji H, Otaki F, et al. High-fidelity simulation versus case-based tutorial sessions for teaching pharmacology: convergent mixed methods research investigating undergraduate medical students' performance and perception. PLoS ONE. 2024;19. https://doi.org/10.1371/journal.pone.0302609.
7. Gottlieb M, Wagner E, Wagner A, Chan T. Applying design thinking principles to curricular development in medical education. AEM Educ Train. 2017;1. https://doi.org/10.1002/aet2.10003.
8. Liedtka J. Why design thinking works. Harv Bus Rev. 2018;96(5):72–9.
9. Dym CL, Agogino AM, Eris O, Frey DD, Leifer LJ. Engineering design thinking, teaching, and learning. J Eng Educ. 2005;94(1):103–20. https://doi.org/10.1002/j.2168-9830.2005.tb00832.x.
10. Meinel C, Leifer L, Plattner H. Design thinking: understand–improve–apply. Berlin, Heidelberg: Springer; 2011. https://doi.org/10.1007/978-3-642-13757-0.

11. Ausubel DP, Novak JD, Hanesian H. Educational psychology: a cognitive view. 2nd ed. New York: Holt, Rinehart and Winston; 1978.
12. Guo Y, Wang X, Gao Y, Yin H, Chen T. Flipped online teaching of histology and embryology with design thinking: design, practice and reflection. BMC Med Educ. 2024;24. https://doi.org/10.1186/s12909-024-05373-7.
13. Wu Q, Wang Y, Lu L, Chen Y, Long H, Wang J. Virtual simulation in undergraduate medical education: a scoping review of recent practice. Front Med. 2022;9. https://doi.org/10.3389/fmed.2022.855403.
14. Hamstra SJ, Brydges R, Hatala R, Zendejas B, Cook DA. Reconsidering fidelity in simulation-based training. Acad Med. 2014;89(3):387–92. https://doi.org/10.1097/ACM.0000000000000130.
15. Issenberg SB, McGaghie WC, Petrusa ER, Gordon DL, Scalese RJ. Features and uses of high-fidelity medical simulations that lead to effective learning: a BEME systematic review. Med Teach. 2005;27(1):10–28. https://doi.org/10.1080/01421590500046924.
16. McGaghie WC, Issenberg SB, Petrusa ER, Scalese RJ. A critical review of simulation-based medical education research: 2003–2009. Med Educ. 2010;44(1):50–63. https://doi.org/10.1111/j.1365-2923.2009.03547.x.
17. Kolb DA. Experiential learning: experience as the source of learning and development. 2nd ed. Upper Saddle River: FT; 2014.
18. Aggarwal R, Mytton OT, Derbrew M, et al. Training and simulation for patient safety. BMJ Qual Saf. 2010;19(Suppl 2):i34–43. https://doi.org/10.1136/qshc.2010.042358.
19. Sweetman MA. Schön's theory of reflective practice. In: Routledge Companion to Occupational Therapy. Routledge; 2025. pp. 711–9. https://doi.org/10.4324/9781003526766-57.
20. Boud D, Keogh R, Walker D. Reflection: turning experience into learning. Routledge; 2013.
21. Norman G. Teaching basic science to optimize transfer. Med Teach. 2009;31(9):807–11. https://doi.org/10.1080/01421590903049814.

Figures

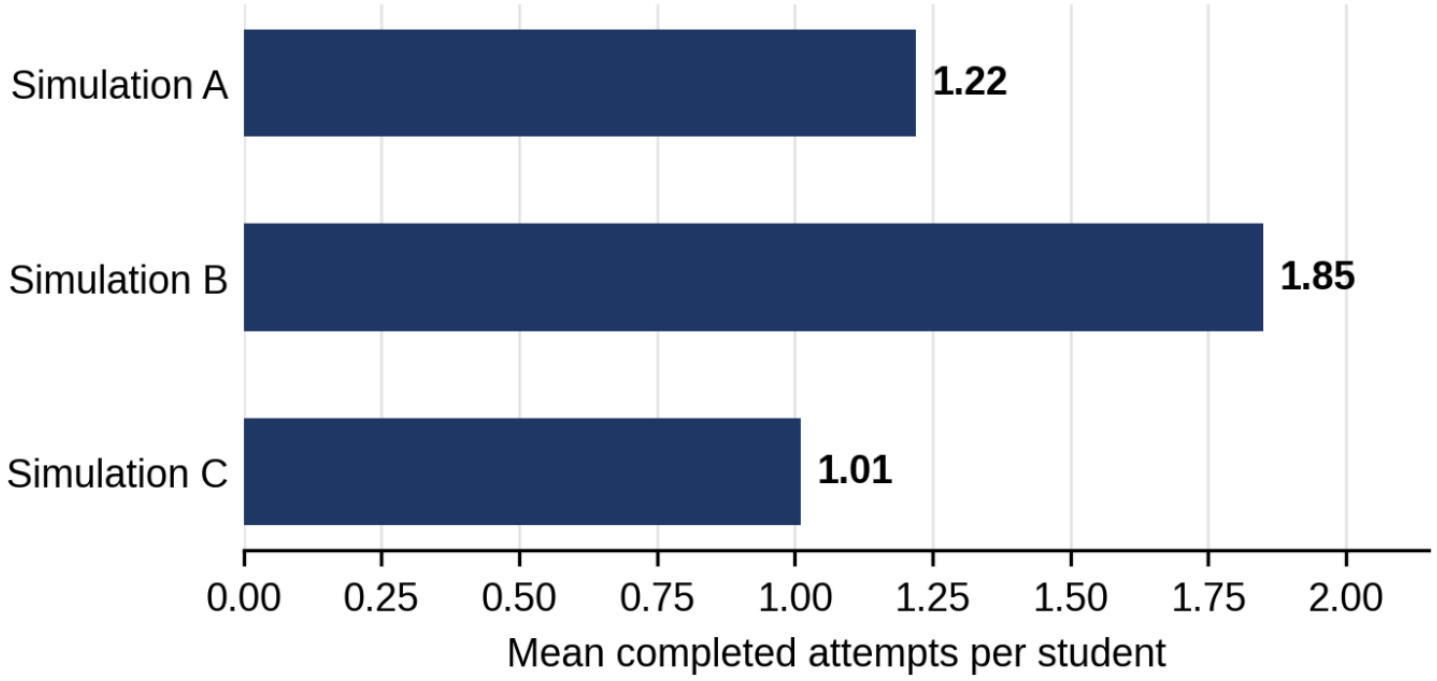


Figure 1

Mean number of completed attempts per student for each simulation, Phase 1 (n=88).