Cross-Specialty Expansion of ENTRUST: A Serious Game-Based Virtual Patient Platform to Teach and Assess Clinical Decision-Making

Principal Investigators: Dana T. Lin, MD and Cara A. Liebert, MD
Co-Investigators: James Korndorffer, Jr., MD, MHPE; Edward Melcer, PhD, Amber W. Trickey, PhD, MS, CPH
Priority Areas Addressed: Collaboration; Rigorous approaches to innovation; Impact and sustainability
Grant Application Type: Teaching and Mentoring Innovation Grant

I. Specific Educational Aims. The goal of this project is to expand the application of ENTRUST, an innovative serious game-based virtual patient simulation platform, to non-surgical specialties to assess clinical decision-making competence for their respective Entrustable Professional Activities (EPAs). This proposal addresses three of the four priority areas, including collaboration, rigorous approaches to scholarship or innovation, and impact and sustainability. The core project team is comprised of an interdisciplinary collaboration between surgical educators, computer programmers, user interface designer/artists, and biostatisticians. Additional lines of collaboration will be fostered with residency leadership and medical educators in other GME training programs to expand and customize the use of ENTRUST in their specialty-specific contexts. The project holds strong potential to have a lasting future impact on how EPAs are taught and assessed at Stanford and at the national level.

II. Project Rationale. *Background*: Medical education is increasingly adopting EPAs as the framework and strategy for verifying clinical competence and autonomy.[1,2] However, few tools exist to objectively assess trainees' performance on EPAs, especially in the critical domain of clinical decision-making. Serious games and virtual patient simulations have the benefit of testing learners' competency through realistic and immersive standardized scenarios that allow for analysis of game play patterns and key points in decision-making. *Preliminary Work*: Our team designed ENTRUST, featuring a user-friendly Case Authoring Portal and immersive Assessment Platform. An overview of ENTRUST can be viewed at https://www.youtube.com/watch?v=Ve1nhBAjkF8. To date, ENTRUST has been successfully piloted and studied as an assessment platform for surgical decision-making as an assessment for the American Board of Surgery Inguinal Hernia EPA and as a high-stakes examination for the College of Surgeons of East, Central, and Southern Africa (COSECSA), the largest accrediting body for surgery in Sub-Saharan Africa. Our data showed a strong positive correlation between ENTRUST performance and years of training (p<0.001) among surgery residents,[3] as well as a high correlation with COSECSA examinees' traditional Objective Structured Clinical Examination (OSCE) scores (p<0.001), providing initial validity evidence[4] for its use as an objective assessment for surgical decision-making. Users rated ENTRUST highly on the System Usability Scale (77.7), akin to well-established software tools.

III. Approach. In collaboration with our multidisciplinary team, the platform will be further innovated and expanded as an assessment platform to evaluate clinical decision-making competence for EPAs in other medical specialties. We will collaborate with medical educators in other medical specialties to author and pilot specialty-specific EPA-aligned case scenarios and assist implementing ENTRUST in their respective programs. With each pilot, we will collect validity evidence (content, response process, relationship to other variables) according to Messick's framework.[5]

Aim 1: Assessment Platform Refinement and Expansion. We will refine the existing functionality, optimize playability, and create new assets and environments to allow expansion to additional medical specialties, such Internal Medicine, Pediatrics, Emergency Medicine, Critical Care, and

Anesthesia. Heightened security features (implementation of security rules, guards, access credentials, data validation and encryption) will support its use in higher-stakes examination contexts, and an expanded administrative dashboard (account management, role assignments, tracking of individual and aggregate user performance data) will allow the platform to serve multiple institutions and training programs.

Aim 2: Specialty-Specific Case Development and Deployment in GME Training Programs. We will collaborate with faculty stakeholders in two other medical specialties at Stanford to author specialty-specific EPA-aligned case scenarios, develop scoring algorithms, and deploy ENTRUST to evaluate trainees' clinical decision-making competency. Cases will be iteratively developed and aligned with EPA definitions to query decision-making identified as critical for entrustment. Cases and scoring algorithm will be vetted by an expert panel.

IV. Timeline and Plan for Implementation

	October 1, 2022 – July 30, 2023									
	Months									
Tasks	1	2	3	4	5	6	7	8	9	10
Assessment platform refinement/expansion (Aim 1)	X	X	X	X						
Specialty-specific case development (Aim 2)				X	X					
Deployment in GME Training Programs (Aim 2)						X	X			
Data Analysis								X	X	
Dissemination and Stemmler Grant Application									X	X

V. Anticipated Work Product. The proposed project will expand and further refine the ENTRUST Platform for use in assessment of clinical decision-making for EPAs across GME specialties such as Internal Medicine, Pediatrics, Emergency Medicine, and Anesthesia. **It is our goal that ENTRUST will become an assessment tool available to all medical educators in the Teaching and Mentoring Academy at Stanford for application within their specific learner population.**

VI. Evaluation Plan. *Data Collection*: As new specialty-specific ENTRUST modules are piloted in their respective training programs, we will continue to collect validity evidence[4-5] for ENTRUST as an assessment tool. An accompanying survey will query demographics, PGY-level, gaming experience, and system usability. Scores will be analyzed for individual cases and combined modules to confirm consistent usability across GME specialties and populations. A backend event log will allow for analysis of game-play patterns that distinguish novice from expert within each specialty. *Data Analysis*: Internal structure will be evaluated by measuring the internal consistency and inter-case reliability. We will conduct bivariate and multivariate analyses to measure associations between total score, PGY-level, training experience, and demographics. Multivariable regression models will be used to identify predictors of total score. Item analysis will be performed on multiple-choice questions to determine item difficulty and item discrimination.

VII. Dissemination of Results. We plan to submit this project for presentation locally at Stanford School of Medicine Annual Medical and Bioscience Education Day and nationally at the AAMC Learn Serve Lead conference. We plan to apply for an NMBE Stemmler Foundation grant to collect further validity evidence for the platform as a summative assessment of clinical decision-making. We anticipate submission of multiple peer-reviewed publications as a result of this innovation. This innovative serious game-based assessment platform is adaptable and scalable with potential to be used by other specialties in GME or UME, and high-stakes accrediting bodies in medical education such as NBME, ABS, ACGME, and COSECSA.

VIII. Appendix: References

1. ten Cate O, Scheele F. Competency-based postgraduate training: can we bridge the gap between theory and clinical practice? *Acad Med*. 2007;82(6):542-547.

2. ten Cate O, Chen HC, Hoff RG, Peters H, Bok H, and van der Schaaf M. Curriculum development for the workplace using Entrustable Professional Activities (EPAs): AMEE Guide No. 99. 2015. *Medical Teacher*, 37:11, 983-1002.

3. Lin DT, Melcer EF, Keehl OG, Trickey A, Eddington H, Merrell SB, Korndorffer JR, Liebert CA. ENTRUST: A Serious Game-Based Virtual Patient Platform to Assess Entrustable Professional Activities in Surgical Education. *J Am Coll Surg*. 2021;233(5):S224. doi: 10.1016/j.amcollsurg.2021.07.462

4. Cook DA, Zendejas B, Hamstra SJ, Hatala R, Brydges R. What counts as validity evidence? Examples and prevalence in a systematic review of simulation-based assessment. *Adv in Health Sci Educ*. 2014 19:233-250.

5. Messick S. Standards of Validity and the Validity of Standards in Performance Assessment. *Educ Meas Issues Pract*. 1995;14(4):5-8. doi:10.1111/j.1745-3992.1995.tb00881.x