ACS 2025 Surgeons and Engineers: A Dialogue on Surgical Simulation Meeting

O-14

Research Abstracts

Linking Feedback Quality to Procedural Skill Improvement: An (Artificial Intelligence) AI-Driven Approach

Divya Kewalramani, MD; Diego Sanhueza Roman; Sofía Abedrapo Lagos; María Elena Vial Brizzi; Advaith Bongu, MD, FACS; Cristian Jarry Trujillo; Mayur Narayan, MD, MPH, MBA, MHPE, FACS; and Julian Varas Cohen, MD

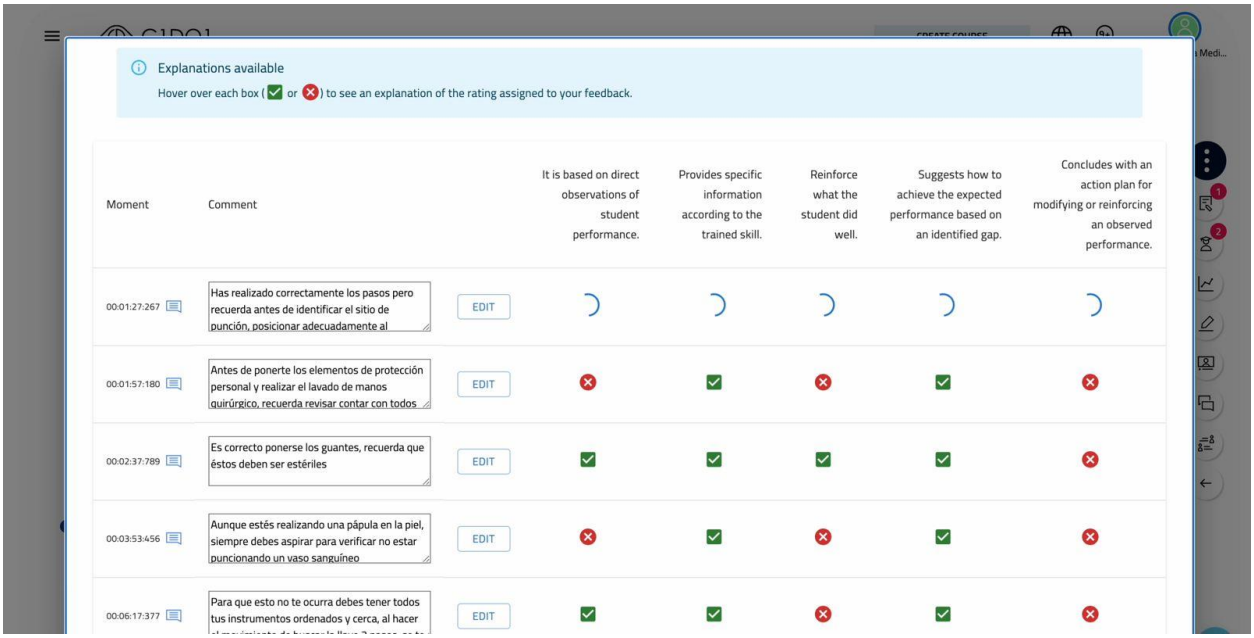
Rutgers Robert Wood Johnson Medical School, New Brunswick, NJ; Pontificia Universidad Católica de Chile, Santiago, Chile

Introduction: High-quality feedback is crucial for developing procedural competencies, but variability in educator expertise and communication challenges complicates standardized delivery. To address this, we developed an AI tool to assess feedback quality based on five criteria: direct observation, specificity, reinforcement of strengths, targeted suggestions, and action plans. This study investigates the relationship between AI-measured feedback quality and improvement in procedural skills.

Methods: Fourth-year medical students performed simulated paracentesis, with recorded videos uploaded to a remote, asynchronous assessment platform. Feedback, consisting of written annotations from trained peer reviewers, identified procedural errors. Cases with initial OSATS scores <20 were analyzed. Feedback quality was assessed using an AI tool, with "Feedback Quality%" (FQ%) calculated as the proportion of criteria met, adjusted for feedback volume, excluding introductory comments. Pearson's correlation and multivariate linear regression analyzed the relationship between ΔOSATS scores and FQ%, adjusting for response length and inter-trial interval.

Results: From 113 uploaded videos, 953 text feedbacks were extracted; after applying all exclusion criteria, 688 feedbacks were analyzed. A positive correlation between FQ% and ΔOSATS ($r = 0.309$, $p < 0.001$) was identified. Feedback quality metrics, (≥3, ≥4 criteria met) showed similar positive correlations ($r = 0.156$, $p < 0.001$; $r = 0.140$, $p < 0.001$). Multivariate regression, adjusted for response length and time between attempts, further confirmed this association ($\beta = 11.61$, $p < 0.001$), accounting for 13.16% of the variance in ΔOSATS. Among the five criteria, fulfilling Criteria 2 and 3 was associated with higher ΔOSATS scores ($r = 0.238$, $p < 0.001$; $r = 0.109$, $p < 0.01$). AI-analysis matched or exceeded human evaluations in 83% cases, indicating a conservative bias.

Conclusions: Our findings highlight the importance of targeted, criteria-driven feedback in enhancing skill acquisition, demonstrating the potential of AI tools to standardize and optimize feedback delivery in medical education.


Explanations available
Hover over each box (or) to see an explanation of the rating assigned to your feedback.
Moment
Comment
It is based on direct observations of student performance.
Provides specific information according to the trained skill.
Reinforce what the student did well.
Suggests how to achieve the expected performance based on an identified gap.
Concludes with an action plan for modifying or reinforcing an observed performance.
00:01:27:267
Has realizado correctamente los pasos pero recuerda antes de identificar el sitio de punción, posicionar adecuadamente al
EDIT
00:01:57:180
Antes de ponerte los elementos de protección personal y realizar el lavado de manos quirúrgico, recuerda revisar contar con todos
EDIT
00:02:37:789
Es correcto ponerse los guantes, recuerda que éstos deben ser estériles
EDIT
00:03:53:456
Aunque estés realizando una pápula en la piel, siempre debes aspirar para verificar no estar puncionando un vaso sanguíneo
EDIT
00:06:17:377
Para que esto no te ocurra debes tener todos tus instrumentos ordenados y cerca, al hacer
EDIT