Abstract: Purpose
To examine challenges encountered in patient note scoring by raters from three medical schools and discuss strategies for overcoming those challenges.

Methods
The study used a mixed method approach. Four patient notes of 72 randomly selected rising 4th-year students from three medical schools were each scored by two different raters. The notes used the same format as that of the United States Medical Licensing Examination Step 2 Clinical Skills. A 4-point scale was used to assess five domains of clinical competency, including Documentation of History (DOC_HX), Documentation of Physical Examination (DOC_PE), Differential Diagnosis (DDX), Justification of Differential Diagnosis (DDJ), and Workup (WP). Intra-class correlation coefficients (ICC) for each domain and case were calculated to evaluate inter-rater reliability. A generalizability study was conducted to examine the reproducibility of the scores. Responses to a survey on raters’ scoring experiences were examined using content analysis to discern an in-depth understanding of the scoring process and pathways for future improvement.

Results
Twelve raters participated in the study and scored a total of 288 patient notes. Eight (67%) responded to the survey. ICCs for ratings on domains ranged from .48 (DDJ) to .83 (DOC_PE), and those for ratings on cases ranged from .64 to .84, indicating varied domain difficulty and case specificity in scoring. The highest percentage of total score variance was associated with student competency (49%), with a phi coefficient of .49. Raters’ feedback highlighted the need for a more specific description for each scale anchor, a clearer answer key for each case, and more scoring discussions between raters before the task.

Conclusions
Different rater training methods used by different schools may lead to different challenges experienced by raters. Descriptors with numeric specifics for scale anchors and answer keys matching more closely with the questions asked are the primary areas needing future improvement.

Level of Audience: Mid-career
Focus of Presentation: UME