Purpose
Interest is growing in the use of qualitative data for assessment.\(^1\) Written comments on residents’ in-training evaluation reports (ITERs) can be used by faculty to reliably distinguish between residents,\(^2\) and faculty are adept at interpreting these comments.\(^3\) However, if residents don’t interpret assessment comments in the same way, a valuable educational opportunity may be lost. We therefore explored residents’ interpretations of written assessment comments using a mixed methods approach.

Methods
Twelve PGY2s in Internal Medicine (IM) rank-ordered a set of unknown PGY1 residents from a previous year in IM based solely on their ITER comments. Each PGY1 was ranked by 4 PGY2s and generalizability theory was used to assess inter-rater reliability. Each PGY2 was then interviewed about their rank-ordering process, how they made sense of the comments and how they viewed ITERs in general. Interviews were analyzed using constructivist grounded theory.

Results
Across 4 PGY2 judges the G coefficient was 0.84. For a single judge G=0.56. Resident rankings correlated extremely well with faculty’s at r=0.90. Residents were adept at reading between the lines to construct meaning from the comments, using language cues similarly reported for faculty.\(^3\) In the interviews they discussed the difficulties interpreting vague language and provided thoughts on why they think it occurs (time, discomfort, memorability and the permanency of written records). They emphasized the importance of face-to-face discussions, the relative value of comments over scores, staff-dependent variability in assessment, and the perceived purpose and value of ITERs. They saw particular value in the opportunity to review an aggregated set of comments.

Discussion/Conclusions
Residents understood the “hidden code” in assessment language, and their ability to rank-order residents based on comments matched faculty’s. Residents seemed to accept staff-dependent variability as a reality.\(^4\) Our findings add to the growing evidence that supports the use of narrative comments\(^5\) and subjectivity in assessment.

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

References: