Supplemental Digital Appendix 1
Examples of Contributions to Inequality and Assessment

Examples of Contributions to Inequity in Assessment		
Component of Assessment	**Contribution to Inequity in Assessment**	**Example**
Individual	If women, URIM learners or SOGI minorities are differentially prepared to demonstrate achievements, unrelated to aptitude.	• Cultural backgrounds that reward introvert behaviors or discourage challenging elders may disadvantage learners in environments that reward charisma and debate
Intrinsic: Assessment for Learning	If URM students do not have equal opportunities to be observed and coached for improvement.	• Unconscious bias on the part of supervisors • Lack of time for direct observation and feedback or no opportunity for the learner to try again after feedback is provided and before assessment is submitted • Failure of supervisor to recognize the impact of microaggressions or stereotype threat or to consider the complexity of assigned patient in assessing performance of learner. • Failure of supervisor to compare demonstrated performance against school-based standardized criteria; instead using personal or normative standards.
Intrinsic: Assessment of Learning	If grades are assigned: • in a manner that amplifies small differences in assessed performance that are not educationally significant; • based on assessment elements that are known to disadvantage URML; • without considering comparative complexity of work; • without using strategies to mitigate potential of unconscious bias on the part of some observers; • in a manner in which early deficiencies in performance that are corrected during the educational experience weigh against the learner.	• Normative grade assignment such that non-significant differences in scores lead to outsized difference in grades • Selecting criteria that do not predict future performance or for which only low-validity measures exist • Overweighting of knowledge exam scores, known to disadvantage URM learners. • Overemphasis on numeric rating scales leads to loss of information that would allow committees to interpret descriptions of assessed performance in the context of case complexity. • Assigning equal weight to quantitative performance descriptors obtained early in and at the end of the educational experience in considering the final grade, rather than recognizing the level of competency attained • Failure to collect adequate sample size of performance assessments under different circumstances
Instrumental: Assessment for Ranking	If those ranking students for future opportunities focus on elements of assessment that do not predict future performance.	• Ranking for residency programs based on standardized exam scores
Contextual: Conduct of Learning Experiences	If circumstances prevent URM learners from demonstrating true achievement or aptitude.	• URM learners distracted by microaggressions , stereotype threat or substandard care for minority population • URM learner assigned or voluntarily take on the care of more complex patient (i.e. those with limited English proficiency • If URM learners lack cultural capital & are unaware of unwritten rules of performance. • URM learners lack the educational opportunities to maximize performance because of unequal resources

Supplemental Digital Appendix 2
Examples of Contextual Inequity

Performance negatively impacted by racism	AM is a 2nd year medical student who wears a hijab as a reflection of her Muslim faith. During her first rotation on family and community medicine, a patient referred to her as a terrorist on rounds and refused to allow her to participate in her care. AM was very shaken up by this experience and found it difficult to concentrate on the discussion on rounds. She couldn't answer the question her attending posed to her and she saw that she was disappointed in her response.
Performance negatively impacted by stereotype threat	GS is a third-year African American medical student interested in pursuing a career in plastic surgery. She has been warned that she must earn honors in her core surgical clerkship to have a chance to fulfill this goal. She is dismayed to find out from students in the class ahead of hers that the very senior and influential attending to whom she has been assigned is known to make disrespectful comments about minority and poor patients and what he refers to as 'affirmative action' students—those from racial and ethnic minorities. Students are too scared to report this attending to departmental or school leaders because of the role he plays in residency selection. She is anxious about whether she has a chance to perform well and be judged fairly or whether she will reinforce his attitudes about minority students.
Assessment negatively impacted because of use of unendorsed or unpredictive criteria	LK is a 2nd year woman resident in internal medicine. She has superb exam scores and is well liked by her patients, peers and other health professionals on the ward. She is passed over for chief resident because she is "too soft spoken" for a leader.
Performance and assessment negatively impacted by unequal social capital	JH is the only woman learner on the surgical service. She has a particularly difficult time with one attending physician who asks a series of obscure questions in the operating room, questions that only she seems not to be able to answer. When talking with another resident after the procedure, she becomes aware that the men are all preparing one another for this attending while changing in the locker room before the surgical procedure.
Assessment negatively impacted by overemphasis of performance rather than growth over the course of the rotation	FT is on a clerkship where the attending faculty change every day or two. The Department asks each faculty member to rate the students on their performance for the short time interval in which they supervise the student. At the beginning of the clerkship, FT had difficulty with some of the core physical exam maneuvers required for this specialty but he worked hard to improve and by the midpoint he was told his technique was excellence. Nonetheless, the assessment of his early preceptors that he needed more work on technique led to a lower grade for the clerkship.
Assessment negatively impacted by failure to consider complexity of work	CL is a 3rd year Latinx student who comes from a rural background and also speaks fluent Spanish. She has been certified as an interpreter by the health system in which she is doing her clinical clerkships. In contrast to the other student on the team, she has been assigned multiple patients with limited English proficiency because of her language skills and is often called upon to interpret for other patients on the team. She was happy to serve this important role for her patients and her team but when she reviewed her evaluation, she read that her faculty gave her lower scores in practice based learning and medical knowledge because "she spent more time with patients than she did in the library."
Assessment negatively impacted by lack of direct observation	EM is an early clerkship student who is excited about pursuing a career in orthopedic surgery. When she receives her surgery clerkship evaluation, she is shocked to see that she has received low grades on her skills in history taking and physical exam because no one observed her doing a history or physical during her month on service.