ACADEMIC MEDICINE
Journal of the Association of American Medical Colleges

Author: Hanson Janice L. PhD, EdS, MH; Pérez Maria MA; Mason Hyacinth R.C. PhD, MPH; Aagaard Eva M. MD; Jeffe Donna B. PhD; Teherani Arianne PhD; Colson Eve R. MD, MHPE

Title: Racial/Ethnic Disparities in Clerkship Grading: Perspectives of Students and Teachers

DOI: 10.1097/ACM.0000000000004914

Academic Medicine

DOI: 10.1097/ACM.0000000000004914

Racial/Ethnic Disparities in Clerkship Grading: Perspectives of Students and Teachers

Janice L. Hanson, PhD, EdS, MH, Maria Pérez, MA, Hyacinth R.C. Mason, PhD, MPH, Eva M. Aagaard, MD, Donna B. Jeffe, PhD, Arianne Teherani, PhD, and Eve R. Colson, MD, MHPE

J.L. Hanson is professor of medicine, director of education scholarship development, and codirector, Medical Education Research Unit, Washington University School of Medicine in St. Louis, St. Louis, Missouri; ORCID: http://orcid.org/0000-0001-7051-8225.

M. Pérez is a clinical research specialist, Washington University School of Medicine in St. Louis, St. Louis, Missouri; ORCID: https://orcid.org/0000-0002-2809-2504.

H.R.C. Mason is associate professor of public health and community medicine and assistant dean for students, Tufts University School of Medicine, Boston, Massachusetts; ORCID: https://orcid.org/0000-001-6443-4244.

E.M. Aagaard is professor of medicine, Carol B. and Jerome T. Loeb Professor of Medical Education, and vice chancellor and senior associate dean for education, Washington University School of Medicine in St. Louis, St. Louis, Missouri; ORCID: https://orcid.org/0000-0002-5773-0923.

D.B. Jeffe is professor of medicine and codirector, Medical Education Research Unit, Washington University School of Medicine in St. Louis, St. Louis, Missouri; ORCID: https://orcid.org/0000-0002-7642-3777.

A. Teherani is professor of medicine, education scientist, Center for Faculty Educators, director of program evaluation and education continuous quality improvement, and founding codirector,

UC Center for Climate, Health and Equity, University of California, San Francisco, School of Medicine, San Francisco, California; ORCID: http://orcid.org/0000-0003-2936-9832.

E.R. Colson is professor of pediatrics and associate dean for program evaluation and continuous quality improvement, Washington University School of Medicine in St. Louis, St. Louis, Missouri; ORCID: https://orcid.org/0000-0003-3505-5071.

Correspondence should be addressed to Janice L. Hanson, 224 Becker Library, MSC-8021-13-220, 660 S Euclid Ave., St. Louis, Missouri 63110; email: janicehanson@wustl.edu; Twitter: @JaniceEducation.

Supplemental digital content for this article is available at http://links.lww.com/ACADMED/B317.

Acknowledgements: The authors wish to thank the study participants affiliated with 3 academic medical centers: Washington University School of Medicine in St. Louis, the University of California San Francisco, and Albany Medical College. The authors also thank Meihsi Chiang, MSW, at The Brown School Evaluation Center at Washington University in St. Louis for data visualization services.

Funding/Support: This study was funded by the Association of American Medical Colleges (AAMC) Group on Educational Affairs (GEA) National Grant Award.

Other disclosures: None reported.

Ethical approval: All procedures performed were in accordance with the ethical standards of the Institutional Review Board (IRB) at Washington University School of Medicine in St. Louis (myIRB# 202003198).

Abstract

Purpose

Racial/ethnic disparities exist in clinical clerkship grading, yet little is known about medical student and faculty perspectives on why these disparities occur. This study explored what happens during clerkships that might explain grading disparities.

Method

Medical students and clerkship teachers at three U.S. medical schools completed a demographic survey and semi-structured interview. The constant comparative method was used to analyze transcripts by inductively developing codes, grouping codes in categories, and refining codes, descriptions, and group assignments to identify themes. Interpretations of and relationships among themes were iteratively discussed to develop a grounded theory.

Results

Fifty-nine participants (29 medical students, 30 teachers [28 clinical faculty, 2 residents]) were interviewed in 2020. The Social Milieu of Medical Education (relationships, fit, opportunities, and judgments in the clinical-learning setting) was the organizing theme, influenced by 5 additional themes: Societal Influence (experiences in society); Students' Characteristics and Background (personal characteristics and experiences outside medical school); Assessment Processes (collection of student performance data and how data inform grades); Learning Environment (resources available and messaging within the clinical setting), and Students' Interactions and Reactions (interactions with and reactions to peers and teachers). The grounded theory highlights complex, multi-layered aspects of how the social milieu of medical education is shaped by and shapes students' experiences, relationships, and clerkship assessments and promotes clerkship-grading disparities.

Conclusions

Mitigating clerkship-grading disparities will require intervening on inter-related, contextual factors to provide equitable opportunities for students from diverse backgrounds and with varying styles of engagement in clinical-learning settings, along with attending to modifying assessment processes.

Clinical-grading disparities exist in medical schools across the United States.[1-3] Lee et al. observed that medical students who identified as being from groups historically underrepresented in medicine (URiM) received lower grades than white peers across all clerkships.[2] In another study, after accounting for confounding factors such as the National Board of Medical Examiners (NBME) specialty-specific shelf examination scores, medical student race/ethnicity was independently associated with clerkship grades in several core clerkships.[4] Clerkship-grade disparities also have been associated with the amount of time that medical students spend with their evaluator.[5] Disparities in clerkship grades occur in the context of long-standing concerns about the meaning, value, and validity of grades.[6]

Clinical-grading disparities in medical school could have long-lasting implications for students' specialty choice, residency placements, and career paths.[1,7-9] Changes in pre-clerkship grading to pass/fail at many medical schools and the NBMEs' decision to make United States Medical Licensing Examination (USMLE) Step 1 pass-fail have heightened the importance of clerkship grades in residency selection. Nomination to and membership in the Alpha Omega Alpha (AOA) honor society, largely based on academic performance including clerkship grades, also can affect residency selection,[8,9] despite recognized racial/ethnic disparities in AOA nominations.[1,7]

While disparities in clinical grades and AOA nominations have garnered attention, medical students' and their teachers' perceptions about why racial/ethnic disparities in clerkship grades exist remain under-explored. Data from focus groups with medical students at a single medical school suggested that grading disparities related to both the learning environment and student assessment processes.[4] Medical students' interpersonal communication style[2] and narratives that evaluators use to describe students on clerkship evaluation forms[10] also might contribute to grading disparities. Missing is an in-depth exploration of students' and teachers' perceptions of

how and why racial/ethnic clerkship-grading disparities occur. To address this gap, we developed this study using critical race theory (CRT) as a conceptual framework. CRT describes how race and racism are ingrained in law, power differentials, and education.[11] This lens prompted us to query the role of perceptions about race and racism in our society and the structures and settings of medical education. We also used Milner's recommendations, which encourage education researchers to examine their racial and cultural awareness and positionality in the context of CRT, to guide our investigation.[12]

The aim of our study was to explore the perspectives of medical students and their teachers and assessors regarding why racial/ethnic disparities in clerkship grading occur, as the first step toward identifying and implementing interventions to move toward achieving educational equity.

Method

Study design

We conducted a multi-site qualitative study using a grounded theory approach and an epistemological frame of social constructivism, which posits that people construct meaning from social experiences.[13,14] We chose qualitative research methods to illuminate the details and nuances around what happens and why. The grounded theory approach guided coding and identification of themes in interviews, relationships among the themes, and multilayered influences on disparities in clerkship grading.[13,15] CRT influenced our development of interview questions and codes, and interpretation of the codes and themes that led to the grounded theory. Using Milner's recommendations for self-reflection, the research team engaged in intentional reflexivity and bracketing of the perspectives and biases we brought to the study.[12]

The research team comprised 6 faculty members from 3 academic medical centers (Washington University School of Medicine in St. Louis [WUSM], University of California San Francisco [UCSF], and Albany Medical College [AMC]), and an experienced research project manager at WUSM. All team members had extensive experience and complementary expertise in medical education research with a focus on promoting diversity, equity, and inclusion in academic medicine. The study was approved by the Institutional Review Board at WUSM (WUSM myIRB#202003198). All study procedures were performed in accordance with the ethical standards of the 1964 Helsinki declaration and its later amendments.

Participants and procedures

Study recruitment and enrollment occurred between September and December 2020. We purposefully sampled medical students who were in or had completed clerkships and faculty (including clerkship directors) and residents who taught and assessed students during clerkships (i.e., teachers). Recruitment included broad distribution of emails to students and teachers supplemented by word-of-mouth (snowball) sampling to achieve maximum diversity by participant self-identified race/ethnicity, sex, and medical school until reaching saturation. This recruitment approach supported trustworthiness of data and interpretation and enhanced the potential for transferability of the results to other samples.

We developed a semi-structured interview guide based on the literature and our team's experience evaluating students during clinical clerkships. CRT[12] prompted us to include questions to probe participants' experiences with race/ethnicity prior to medical school and observations of the role of race/ethnicity during clerkships. The interview guide was pilot tested and modified and is available in Supplemental Digital Appendix 1 at http://links.lww.com/ACADMED/B317.

Through recruitment emails, prospective participants were linked to a web-based study screener programmed in Qualtrics (Qualtrics, Provo, UT) to obtain informed consent and determine study eligibility. The project manager (M.P.) contacted eligible participants to schedule interviews and request completion of a brief Qualtrics demographic survey, then conducted interviews via Zoom (Zoom Video Communications, Inc., San Jose, California). The interviewer identifies as Hispanic and neither assesses nor grades medical students. Interviews lasted 30-60 minutes and were audio-recorded, transcribed verbatim, and de-identified. Enrollment ended when data analysis reached saturation (no new codes or themes were emerging and we had a robust understanding of codes, themes, and relationships between them).[13] Study participation was voluntary; participants received a $25 electronic gift card.

Data analysis

The constant comparative method was used to analyze and interpret the data and build a grounded theory.[13,15] We developed codes inductively, grouped codes in categories, and continuously refined code labels, descriptions, and assignment of codes to groups. We examined emerging codes to consider deductive codes from CRT. Three faculty members (J.L.H., E.R.C., H.R.C.M.) and the project manager held weekly coding discussions and periodic additional discussions with the entire research team, until the codebook was well-established. These 3 faculty members independently coded the remaining transcripts. Another team member then reviewed each transcript to confirm agreement between coders. Disagreements were resolved through discussion. Interpretations of groups of codes led to descriptions of themes, identification of key concepts in themes, and relationships among themes, which were iteratively discussed to develop our grounded theory. We ensured trustworthiness through triangulation of data[16] from a diverse sample of participants, data analysis conducted by a diverse team,

intentional self-reflection during data collection and analysis, and member checking for feedback on codes, themes, and emerging theory.[16] Member checking was conducted with 6 diverse participants across all 3 sites (4 teachers and 2 students).

We used HyperRESEARCH version 4.5.3 (Researchware, Inc., Randolph, Massachusetts) for qualitative analysis and SPSS version 25.0 (IBM Corp., Armonk, New York) to present descriptive statistics of demographic data. Standards for Reporting Qualitative Research guided our reporting of the study.[17]

Results

We screened 133 prospective participants, contacted 99 individuals by email, and interviewed 59 medical students and teachers (see Table 1). Six themes emerged from the data describing participants' perspectives of why racial/ethnic disparities exist in clerkship grades. We identified the *Social Milieu of Medical Education* as the organizing theme; the social milieu includes the relationships, opportunities, and judgments that influence assessment data in the clinical-learning setting. Five additional themes describe what shapes and gives meaning to the *Social Milieu of Medical Education* and what shapes individuals' experiences and assessment of each student: *Societal Influence* (attitudes, biases, and experiences in society); *Student Characteristics and Background* (personal characteristics and experiences of students before and outside of medical school); *Assessment Processes* (the process of collecting student performance data, and the data and judgments about student performance that shape grades); the *Learning Environment* (implicit and explicit messages within the clinical-learning setting and resources available to students); and *Student Interactions and Reactions* (how students interact with and react to peers and teachers). Figure 1 illustrates the theory that resulted from our analysis, which indicates that medical students and teachers bring societal influences and experiences into the social milieu of

medical education, where they affect assessments, interactions and relationships, and the learning environment—which collectively create disparities in grades. Quotes in the text and Table 2 have been edited for ease of reading, without changing content or meaning. Exemplar quotes for all themes appear in Table 2 along with key concepts reflected in each quote.

The social milieu of medical education

Descriptions of the *Social Milieu of Medical Education* encompassed the relationships in clinical-learning settings, including interactions between students, teachers, team members, and patients during patient care, teaching, observation, feedback, and assessment. These relationships and interactions provided the frame for the judgments that influenced assessment. Particularly striking were descriptions of a student's "fit" in the clinical setting. Students described feeling out of place, not part of the "club," or not belonging, and teachers described noticing when a student did or did not seem like a "good fit" for medicine or a particular specialty. The notion of "fit" extended to relationships between students and teachers, with only some students experiencing a comfortable connection with their teachers. Teachers and students alike spoke of how the degree of commonality between students and teachers affected their relationships and the comfort of their interactions, with those who shared common background, interests, interaction styles, and/or personalities relating to one another more comfortably. These relationships, in turn, influenced students' opportunities to participate in the clinical-learning environment. Teachers with whom students had comfortable relationships were perceived as offering more clinical-learning opportunities to ask questions, participate in providing care, demonstrate improvement in clinical skills, and engage with mentors.

> It's just based, again, on the instructor and probably the year they were raised and how they perceive medicine and what medicine looked like

> back in their day. … If you don't fit that criteria or that role, which is probably male Caucasian … they tend to judge you a little bit harder … [Teacher]

Societal influence

Participants described experiences that students, teachers, and patients carried with them into the social milieu of medical education, including assumptions and biases from society regarding gender, sexual orientation, politics, race, ethnicity, culture, privilege, and what they perceived as "fair" or "not fair." Many were acutely aware of how these assumptions, biases, and experiences affected interactions, relationships, opportunities to engage in clinical-care conversations, and judgments about a student's performance. Other participants stated that they were unaware of disparities based on race or ethnicity yet shared examples of differences in observations, attitudes, and judgments based on these social constructs. They carried these attitudes into conversations among teachers, learners, and team members, affecting interpretations of assessment forms and students' performance.

> I guess I would say, everybody carries personal biases. … Just the culture we live in and, maybe, your generation. … [I]f we're talking strictly about race or ethnicity, I would say that's the biggest driving factor, just whatever kind of personal biases people carry. [Student]

Student characteristics and background

Students' family background, experiences prior to medical school, prior mentors, and resources—including financial resources and social support—were described by students and teachers as influencing students' preparation to engage successfully in medical school. Students with physician family members grew up in a medical culture and found it easier to find

commonality with their teachers. First-generation college graduates, however, expressed having less in common with their teachers and more difficulty forming student-teacher relationships. Students and teachers recounted varied reactions to physical appearance (hair, make-up, clothing, attractiveness), language (whether English was their first language, their "accent," the vocabulary and interaction styles of their culture), and voice (loud, soft, high, or low), sometimes forming first impressions of knowledge, skill, and attitudes based on these characteristics. Participants described deliberately changing the pitch of their voice, words and style used to communicate with patients, or physical appearance, to adapt to perceived expectations.

> As a student from an underrepresented background, I feel like my comments have largely landed in my personal characteristics … and not on my clinical abilities. [Student]

Assessment processes

Students and teachers described ways that assessment processes led to biased judgments about a student's performance. Assessment forms included space for teachers' judgments about performance that did not directly reflect the student's knowledge, skills, and attitudes. The assessor's role figured prominently in descriptions of assessment processes, with participants expressing concerns that different assessors provided judgments that were not comparable, leading to grades based on which teacher they happened to work with and the potential for assessors to adjust scores to favor students with particular characteristics. Students spoke of the "randomness" of assessment and learning experiences, with variation in how teachers assessed them, quality of teaching, patients and clinical-learning opportunities, and curriculum that seemed haphazard, all impacting students' abilities to demonstrate that they met or exceeded expectations. Students and teachers alike explained that assessment and grading processes were

not transparent and that grades sometimes seemed to be "made up." Assessment processes were also described as affected by time—not enough time for assessors to get to know students, to complete assessment forms thoughtfully, and for prolonged exposure to opportunities to learn essential skills. The most prominent categories of comments noted the "subjectivity" of assessment data, grades, and gestalt judgments about students. Teachers described their assessments within their gestalts. Students and teachers longed for "objective measures" and consistent standards for performance based on clear criteria, although they described such measures and standards as elusive.

> I think that the current [assessment] system is actually more subjective than the previous system. It tends to be more subjective in making judgments about how to assign final grades. So we have honors and pass—or honor and high-pass minimums, and then we discuss how to push people … up to the next-highest class or category of grade, without a lot of data based on pretty sparse comments that are present in the record that are completely divorced from actual experience with the individuals. It takes away the bias that's introduced by actually knowing the person—and introduces bias that's based on speculation and not actually knowing the person. [Teacher]

Many participants spoke about the data that determined grades. They provided descriptions of competencies or knowledge, skills, and attitudes, and expressed a desire for a set of skills that all students would be expected to learn. Also mentioned were exams, oral presentations, OSCEs (objective structured clinical examinations), and overall "student knowledge." Potential for bias and subjectivity was described for each type of data, along with the potential for the data to

disadvantage some students, particularly students who came to medical school with little experience in medical environments or with cultural experiences very different from other students and teachers. Oral presentations were perceived as the main opportunity for students to demonstrate knowledge, skills, and attitudes to their assessors. Standardized exams, often thought of as "objective" data, were said to favor students with past opportunities to learn how to perform well on exams. Most participants recognized possible flaws of knowledge assessments such as shelf examinations, while a few viewed these as the only objective assessments available.

> They do an OSCE. … Is that rating affected by people's biases? We hope that it isn't because the way that we do it is we literally just look at a checklist and check it off to make sure. So, we try and make it as standard as possible. But there's certainly … room for it to be affected [by bias]. [Teacher]

The learning environment

Students and teachers varied in their awareness of explicit racial/ethnic disparities. Regardless of awareness of disparities, however, the societal attitudes, biases, and perceptions discussed above were said to affect the clinical-learning climate. Students specifically mentioned words and behaviors they experienced as microaggressions, lack of trust that reports of mistreatment would remain anonymous, attitudes of collaboration or competition among clinical team members, and expectations for how they would relate to patients who looked like the student as sources of their discomfort within the learning environment.

> If you're African American and you have a patient who's really been … either outright racist or has been just dropping microaggressions … and you're working with a supervisor who doesn't quite understand that or how

> it feels … it can be very difficult because they don't understand why you're getting angry or why this interaction is hard. [Student]

Students also described a pervasive "hidden curriculum" about how to get good grades during clerkships, with implicit messages and unwritten rules that led to success. Students described varying levels of understanding of these rules for success, dependent, in part, on their pre-medical school experiences with the medical community and their experience with peers and teachers during medical school. One approach that students and teachers described as helping to address inequities was providing resources such as support or education during the transitions to medical school, clerkships as a whole, or certain clerkships. These resources were seen as a way of promoting equity by preparing students for the next part of their medical education journey.

> I think another element to that is making it very explicit what the hidden curriculum is to everybody…. The way I've tried to do this personally is, someone forwarded me some documents that some older student had written, which were wonderful. And I've basically emailed them to every single person I can think of whenever I can just so that people at least have this sheet of paper that they can look at and be like, "Okay, this is giving me a model of how I should present in the morning when I'm rounding on patients" or, "This is a model of how I should phrase my questions so it doesn't sound like I'm an idiot, but it sounds like I'm curious and want to know more, but I already know a lot of things," right? ... If you kind of know how to play the game and what to do, you'll go on and do better. [Student]

Student interactions and reactions

Students and teachers described judgments and, in turn, assessments, based on engagement, interaction style, and work style in the clinical setting. Engagement included cognitive engagement (asking questions), physical engagement (volunteering for tasks), emotional engagement (engaging in conversation), and engagement with patients. Students' interaction styles were described as timid, affable, charismatic, likeable, outgoing, humble, shy, quiet, humorous, introverted, extroverted, confident, or not confident. Students perceived that their interaction styles affected their teachers' assessments of them, and teachers often acknowledged that this was likely. Students' work styles were described in terms of whether or not they appeared to work hard, prepare well, and be a team player.

> For some reason, the team that I was with, towards the end of working with them, told me that they thought that I didn't have initiative … because I wasn't picking up enough patients, when they hadn't really oriented me to how to do that, and I was also actively asking the team for more patients to follow. And at the end of the clerkship, I got a high pass, which is fine, but my shelf grade—I got, like, 97 percentile on the shelf—and that didn't alter … how my medical knowledge was ranked in their grading rubric. [Student]

Students' engagement and interaction styles (not necessarily their clinical skills) affected teachers' impressions of and interactions with them, which in turn affected students' comfort to participate in the clinical-learning setting, creating a cycle—their style, a reaction from others, a reaction in response of more or less engagement. Sometimes students described feeling as though they had to try harder or were held to a higher standard than their peers, or thinking that others

assumed they had certain experiences or cultural knowledge because of their appearance. Some students experienced emotional reactions to the social milieu, such as feeling unworthy to be in medical school ("imposter syndrome"), feeling at risk of being judged poorly based on stereotypes and perhaps performing less well than they were capable of, or having uncomfortable emotions such as anger or depression. Other students reported developing a higher tolerance for annoying attitudes and comments in the learning environment in response to prior similar experiences, and deciding it was in their best interest to overlook such things.

> I think that the imposter syndrome, almost every med student has it, but the more minority group label you have, the more imposter syndrome you have. [Student]

Discussion

Our findings describe medical student and teacher perspectives about clinical-grading disparities, with six themes contributing to this phenomenon. The theoretical model that emerged from our analysis illustrates how thematic concepts inter-relate to create the social milieu of medical education and the resulting racial/ethnic clinical-grading disparities (see Figure 1). We describe themes discretely for clarity of the concepts, although throughout the interviews, the descriptions of concepts in each theme overlapped across themes. The theory indicates that the interactions and judgments that contribute to grades take place in a milieu of relationships between medical students and teachers in the clinical-learning setting. This social milieu is shaped, in part, by clerkship assessment and grading processes and by the societal context of race, ethnicity, culture and associated biases, which everyone brings into the learning setting and the relationships that occur there. Pre-medical school experiences of race/ethnicity and family influence students' entry into relationships with teachers, peers, and patients and the impressions that others form

about them; these relationships, in turn, influence students' opportunities to participate clinically. Students' styles of engaging in the clinical-learning setting further affect their relationships and opportunities to participate, and reactions to their styles further influence impressions and judgments about them and their future comfort to participate. This feeds back to the clinical-learning setting, where relationships among teachers and students form more or less easily, judgments occur about students' "fit" with medicine and teams, and gestalt impressions set the stage for judgments about a student's performance. No one part of this complex social setting explains clerkship-grading disparities independently; each set of characteristics and experiences interacts with the others. This theory and the rich, complex qualitative data on which it is based suggest the complexity and urgency of need for change, partly in clerkship-assessment processes and grading, and perhaps more importantly, in the nuanced and influential relationships and interactions in clinical-learning settings.

Our findings align with the tenets of CRT[11] and illuminate the interrelationships and intersections that cross race, ethnicity, gender, sexual orientation, and class, and the prominence of these social constructions in medical education. The structures, attitudes, biases, and experiences in our society related to race, ethnicity, and culture permeate our educational institutions, access to resources, opportunities, and social relationships. Students, teachers, and patients all bring this background into the clinical environment, where they shape the relationships and judgments that form clinical-learning opportunities, assessment data, and, consequently, grades.[11,18] CRT guided us to ask participants how experiences in society, outside medical school, influenced their experiences in clinical-learning settings. We learned from participants to notice how racism is reflected in perceptions of student performance that are based on cultural norms, and how this creates inequity.

Our findings add to the literature by explaining the "why" and "what" behind the "how many" of previous quantitative studies.[1-3] We have come to a "profound grasp of the obvious."[19] The rich narratives from our robust and diverse sample of medical students and teachers also illuminate the interconnectedness of experiences reported separately in the literature. Effects of microaggressions from peers, teachers, and patients,[20] for example, affect students' comfort to participate in clinical-learning activities. Students' background experiences affect ease of interacting with teachers, which influences opportunities to participate in clinical care and presentations.[21] Our findings add to concerns about opportunities for implicit biases to inform gestalt judgments—and the difficulty of creating assessment systems that do not allow gestalt judgments to frame assessment data and incorporate biases of which assessors and grading committees are unaware.[6]

Our data describe, in a new way, the patterns of engagement and interaction that students learn to follow in clinical-learning settings, and how these patterns are affected by societal and social environments. Students and teachers describe engagement of students as forming the basis for judgments about a student's knowledge, effort, sense of responsibility, professionalism, and ability to fit in and work on a team. These observations and judgments are somewhat separate from actual clinical performance and skill; nevertheless they form the basis of the gestalt view of a student and the ratings and written comments that create assessment data.

Addressing racial/ethnic clerkship-grading disparities will require creating equity of opportunity in social relationships and interactions in the clinical-learning setting—opportunity for students to participate in clinical care, to acquire and demonstrate clinical skill and talent, to form supportive relationships among students and teachers, and to experience appreciation of their varied styles of engagement and interaction.[22] Our results suggest that assessment processes must

focus on documentation of clinical performance in ways that do not rely on first impressions and gestalt judgments as assessment data. Changing the social milieu of medical education will require faculty development that goes beyond identifying racial/ethnic biases, to developing strategies to support teachers in identifying their individual biases and specific ways that these biases influence their relationships with students with backgrounds and experiences different from their own. In particular, faculty development can help teachers create equitable opportunities for all students to participate in patient care, receive constructive feedback on clinical performance, and demonstrate progress. Peer support among teachers may help them hold one another accountable and develop new insights, much like qualitative researchers hold each other accountable to identify and bracket biases during qualitative analysis. Another way to adjust the social milieu of medical education may involve providing peer-to-peer support while navigating transitions to and between clerkships, making messages about how to succeed explicit and accessible to all students, regardless of their background and prior experiences in medical settings. Next steps include developing and studying multi-faceted interventions to address these aspects of equity in clinical-learning settings.

Our study has limitations. The interview could reflect biases of research team members and affect participant responses. To guard against these sources of bias, the interviewer used consistent open-ended questions, allowing respondents to guide the conversation. Although we purposefully sampled at one private and two public institutions in the northeastern, midwestern and western geographic regions of the United States, we could not include participants from all possible racial/ethnic backgrounds and experiences. We used our professional networks to recruit participants, which may have added bias to which participants enrolled. We did, however, achieve a robust sample size, and we sampled until we reached deep understanding of a complex

set of themes. Our racially and ethnically diverse research team also managed bias with thoughtful reflexivity and careful attention to hearing and incorporating one another's interpretations. We ensured that all transcripts were analyzed by at least two researchers and resolved all disagreements. We described our sample and data in detail, so that others can make decisions about the transferability of our findings to their learning settings.

Conclusions

This study describes a complex, inter-related set of factors that contribute to racial/ethnic disparities in clerkship opportunities, assessment, and grades. As medical schools strive to become more equitable, diverse and inclusive, it is imperative that medical education redouble efforts to mitigate these disparities and move toward achieving equity.

References

1. Teherani A, Hauer KE, Fernandez A, King TE, Jr., Lucey C. How small differences in assessed clinical performance amplify to large differences in grades and awards: A cascade with serious consequences for students underrepresented in medicine. Acad Med. 2018;93(9):1286-1292.
2. Lee KB, Vaishnavi SN, Lau SK, Andriole DA, Jeffe DB. "Making the grade:" Noncognitive predictors of medical students' clinical clerkship grades. J Natl Med Assoc. 2007;99(10):1138-1150.
3. Low D, Pollack SW, Liao ZC, et al. Racial/Ethnic disparities in clinical grading in medical school. Teach Learn Med. 2019;31(5):487-496.
4. Colson ER, Pérez M, Blaylock L, et al. Washington University School of Medicine in St. Louis case study: A process for understanding and addressing bias in clerkship grading. Acad Med. 2020;95(12S Addressing Harmful Bias and Eliminating Discrimination in Health Professions Learning Environments):S131-S135.
5. Ingram MA, Pearman JL, Estrada CA, Zinski A, Williams WL. Are we measuring what matters? How student and clerkship characteristics influence clinical grading. Acad Med. 2021;96(2):241-248.
6. Hanson JL, Rosenberg AA, Lane JL. Narrative descriptions should replace grades and numerical ratings for clinical performance in medical education in the United States. Frontiers in Psychology. 2013;4:668.
7. Wijesekera TP, Kim M, Moore EZ, Sorenson O, Ross DA. All other things being equal: Exploring racial and gender disparities in medical school honor society induction. Acad Med. 2019;94(4):562-569.

8. LaGrasso JR, Kennedy DA, Hoehn JG, Ashruf S, Przybyla AM. Selection criteria for the integrated model of plastic surgery residency. Plast Reconstr Surg. 2008;121(3):121e-125e.
9. Rinard JR, Mahabir RC. Successfully matching into surgical specialties: An analysis of national resident matching program data. Journal of Graduate Medical Education. 2010;2(3):316-321.
10. Rojek AE, Khanna R, Yim JWL, et al. Differences in narrative language in evaluations of medical students by gender and under-represented minority status. J Gen Intern Med. 2019;34(5):684-691.
11. Ladson-Billings G. Just what is critical race theory and what's it doing in a nice field like education? International Journal of Qualitative Studies in Education. 1998;11(1):7-24.
12. Milner HR. Race, culture, and researcher positionality: Working through dangers seen, unseen, and unforeseen. Educational Researcher. 2007;36(7):388-400.
13. Corbin J, Strauss A. Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory. 4th ed. Los Angeles, CA: Sage Publications; 2015.
14. Charmaz K. Constructing Grounded Theory—A Practical Guide Through Qualitative Analysis. Los Angeles, CA: Sage; 2006;208.
15. Glaser BG, Strauss AL. The Discovery of Grounded Theory: Strategies for Qualitative Research. New Brunswick, NJ: Aldine Publishing Company; 1967.
16. Lincoln Y, Guba E. Naturalistic Inquiry. Los Angeles, CA: Sage; 1985.
17. O'Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA. Standards for reporting qualitative research: A synthesis of recommendations. Acad Med. 2014;89(9):1245-1251.

18. Parker L. Critical Race Theory in rducation and qualitative inquiry: What each has to offer each other now? Qualitative Inquiry. 2015;21(3):199-205.
19. Hurley RE. Qualitative research and the profound grasp of the obvious. Health Serv Res. 1999;34(5 Pt 2):1119-1136.
20. Ackerman-Barger K, Boatright D, Gonzalez-Colaso R, Orozco R, Latimore D. Seeking inclusion excellence: Understanding racial microaggressions as experienced by underrepresented medical and nursing students. Acad Med. 2020;95(5):758-763.
21. Hauer KE, Lucey CR. Core clerkship grading: The illusion of objectivity. Acad Med. 2019;94(4):469-472.
22. Teherani A, Perez S, Muller-Juge V, Lupton K, Hauer KE. A narrative study of equity in clinical assessment through the antideficit lens. Acad Med. 2020;95(12S Addressing Harmful Bias and Eliminating Discrimination in Health Professions Learning Environments):S121-S130.

[Figure legend]

Figure 1 Theoretical model of inter-related influences on racial/ethnic disparities in clerkship grades. White circles represent themes that emerged from the data. The black circle represents the phenomenon that was studied. Arrows represent the direction of influence between the themes and of the inter-related themes to racial/ethnic disparities in clerkship grades. Connections between the inter-related themes of "assessment processes," "student interactions and reactions," and "the learning environment" represent bi-directional influences (i.e., potential feedback loops) between each pair of themes that are nested within the Social Milieu of Medical Education.

Table 1

Participant Demographic Characteristics Among Medical Students and Teachers Who Completed Interviews About Disparities in Clerkship Grades Across 3 Academic Medical Centers in 2020 (N = 59)

Characteristics	Medical students (n = 29)	Teachers (n = 30)
Sex		
Male	12	14
Female	16	16
Other response	1	0
Race/ethnicity		
Non-Hispanic White	9	13
Asian	12	6
Underrepresented in medicine[a]	7	10
Unknown or more than one race	1	1
Sexual or gender minority		
Yes	10	6
No	18	24
Prefer not to answer	1	0
Degree program		
MD	26	---
MD and another program[b]	3	---
Phase in curriculum		
Currently in clerkships	14	---
Post-clerkship phase	15	---
Years at institution[c]		
0-5 years	---	9
6-10 years	---	5
11-15 years	---	3
16-20 years	---	3
More than 20 years	---	10
Age		
20-29	27	2
30-39	1	9
40-49	0	7
50-59	0	7
60-69	0	5
Prefer not to answer	1	0
Highest level of education of first parent/guardian		
Less than college degree	4	5
College degree	5	2
Master's degree	9	8
Doctoral degree	11	15

Highest level of education second parent/guardian[d,e]		
Less than college degree	5	6
College degree	9	12
Master's degree	8	5
Doctoral degree	7	3
Country of origin		
Born in the United States	20	25
Born outside the United States to parents who were U.S. citizens	2	2
Born outside the United States to parents who were not U.S. citizens at the time	6	3
Prefer not to answer	1	0
U.S. citizenship/permanent residency		
Yes	27	30
No	2	0
Total current debt		
$0 no debt	7	15
$1-$49,000	3	1
$50,000-$99,000	3	0
$100,000-$149,000	7	1
$150,000-$199,000	2	2
$200,000 or more	5	8
I prefer not to say	2	3
Prior federal Pell grant awardee		
Yes	6	8
No	23	22

[a]Includes participants who self-identified as Black/African American, Hispanic/Latino(a), or American Indian or Alaska Native race/ethnicity.
[b]Three students were enrolled in a joint program, including a Bachelor's/MD program, Master's Degree/MD program, and MD/PhD program.
[c]Two residents were between their second and third years of residency training at the time of study participation. Students referred to residents as teachers and residents participated in assessing students, so we included residents as participants in the teacher group.
[d]Four teachers did not respond to the item about a second parent/guardian's highest level of education.
[e]Two teachers responded to the item about a third parent/guardian's highest level of education indicating in both cases receipt of a doctoral degree.

Table 2
Exemplar Quotes and Key Concepts for Themes Among Medical Students and Teachers

Key concepts	Quotations (with attributions to students or teachers)
Theme 1: Social Milieu of Medical Education	
• Randomness • Climate • Experiences	I think, honestly, if I had to sum up what I think affects grading the most, I think it's luck, the learning environment, and the background of the student. You're lucky if you get a team or an attending who loves to teach— and is willing to teach. You're lucky if you get an attending and a resident in [a] team who works well together and they can agree on the same treatment plan, and they have good communication. You're lucky if you're on a team where the residents love—the interns love what they're doing and they're not super stressed. (student)
• Subjectivity/Gestalt	… Some people may be perceived as … fitting in more because that's the way that this culture has been forever and so the quote "honors student" then … fits into that mold better, and so automatically they have that wind at their back. (teacher)
• Relationships • Commonalities	I could see where maybe if a resident or a medical student have a little bit more common factors and start talking more, that medical student may have a little bit more benefit because they spend a little bit more time with the residents and attendings. So that helps with their grade. … So, the attending or the resident learns a little bit more about that medical student. So it helps with their grading process. (teacher)
• Relationships • Race/Ethnicity/Culture • Family background • Assessor role • Commonalities	I feel like I'm always constantly knocking on the door like, "Please let me in. Please—" and I see when my preceptors, for instance—you know, you have a White male connect with a White male student. They're, like, buddy buddies. You know, they—talk about their families. … And so you have that social link that enables opportunity for them to bond better, to communicate better, to learn from each other better. It's just— … I don't have access to that door. … When we're graded, you know, who are you going to remember better, the student who you didn't really know much about their family … or the student who you chatted with every other day about their family, about something that you had in common—a student that you could relate to? … This is part of what causes the disparities in the grading experience and experiences of students who are underrepresented, minorities in medicine. (student)
• Subjectivity/Gestalt • Opportunities to participate • Interaction style • Work style	If you get along with the team, you're gonna be called a team player. And, you're gonna be perceived as more of a, you know, a positive member of the team. You'll be given more opportunities to, excel. … If you don't mesh with the team, if there's a personality clash, then that student will feel … isolated, marginalized. They won't be given opportunities to … showcase their talent. (teacher)
• Student appearance • Gender • Race/Ethnicity/Culture	I feel like the way patients interact with me, I definitely feel is different for me as a tall white man. Definitely patients answer my questions easier ... than they would a woman or a woman of color. So I don't think—I think the information I get and the respect I have from patients is probably better than them, which may inadvertently lead to my better scores … just because when a tall white man walks in, most of the patients … assume that I'm a higher status. (student)
Theme 2: Societal Influence	
• Hidden curriculum • Fit • Bias	… there are also traditions within medicine … sort of process traditions as well as leadership traditions … that can also … cement some of those … biases. … [O]n the walls of medical schools you see classes and classes of prior graduates and, you know, before a certain era, it was all White men except for a couple White women … so that's just who was in medical school classes and so that's the implicit message that people get. (teacher)

• Bias	I think we draw generalizations very, very quickly in our brains. It takes effort to disabuse us of those … associations and, if we are stressed, if we are placed in an environment where we need to make decisions quickly, where it's really, really busy, where we're remediating another member on the team and don't have as much time to spend with the student, then we may just make summary judgments. (teacher)
• Intersectionality • Bias • (Structural) racism	And then when the student actually arrives I think they are also fighting all of that historical implicit bias in their preceptor's or attending's mind—like, is this student really qualified? … There's nothing I have in common with this person, but I'm supposed to embrace them and teach them … I think it-it can be hard, and it leads to those kind of disparities that you see in grading system. (teacher)
• Race/Ethnicity/Culture • Student's language • Bias • Interaction style	I definitely think there's a difference in terms of how a person of color is graded. … As a person of color myself … I have this like underlying feeling that I have to perform better, and be more on top of things and adjust my own vocabulary and adjust my own attitude and approach in order to match what is traditionally expected, and even though that may not be like my own personal style. (student)
Theme 3: Assessment Processes	
• Assessment forms and systems • Exams	So as much as we would hope that standardized exams would, quote, even the playing field for students, there's good data that supports the fact that standardized exams … in general have bias against certain students. … We still use that assessment, even knowing that standardized tests are not performed on by students equally among races, socioeconomic background, etcetera. (teacher)
• Transparency • Feedback • Randomness • Assessment forms and systems	… there's not a lot of transparency—on how grading is done. Even in clerkships where we're provided with sort of a rubric that they're gonna use, people apply that rubric very inconsistently. Honestly to me, the clerkship grades are, kind of, a black box. You do your best and you do what you think is, desired. And maybe you get some in-the-moment feedback to modify what you're doing. But truly it's just kind of a question mark. (student)
• Subjectivity/Gestalt • Assessment forms and systems	I know there's been a lot of attempts to kind of make that a more of a systematic process. Like, breaking it down into individual attributes. Like, how good is the student at oral presentations, written presentations, and demonstrating professionalism? Kind of like a checklist almost and rating them along that. But again, I—my sense is that this—what happens is more kind of a[n] overall halo effect. (student)
• Assessment forms and systems • Fairness • Relationships	So I think the fairness is sometimes hindered because the attendings may only see us for 30 minutes a day, and you may do one presentation. So if you give one presentation and it's not as good, that has a huge impact versus, you know, you can do excellent the rest of the day … (student)
• Assessment forms and systems	I think it's … an unfortunate feature of human psychology. I think it is. I don't know what it is. But that we make generalizations and inferences, quite quickly on the basis of our prior experiences, and that sometimes to the detriment of individual learners as we assess them. (teacher)
Theme 4: Student characteristics and background	
• Family background • Student's resources • Race/Ethnicity/Culture • Hidden curriculum	… when you're thinking about these standardized exams … students don't always have the same access to the resources as one another. You have students coming from a wide variety of backgrounds. Some people have—their family has more resources for them to be able to purchase certain studying equipment. Some people come from families or physicians, and it—they have people that know how to study, and that's a big part of it, and who know how to manage their time in

	studying. … So you're giving everybody the same questions, but you're not giving everybody the same opportunity to learn. (teacher)
• Fit • Race/Ethnicity/Culture • Experiences	It's more culture—than it is race, per se. Many white students come with greater clinical experience. They've been on the wards. They know people who help them shadow, and so they fit in more rapidly to the culture. They know the expectations better, and so meet them more rapidly. (teacher)
• Hidden curriculum • Family background	… The ability to navigate into your clinical settings is really difficult, right? … I'm a child of immigrant parents, neither of which are physicians—it's a very different world to navigate, and there are many people who pick up on situational awareness very quickly and adapt very easily. There are many who have physician family and households, and so they have a one-up 'cause they already know what to expect. They've been coached— … they've grown up in the world. (teacher)
Theme 5: The Learning Environment	
• Hidden curriculum • Climate • Fit	And then, all of a sudden, you're plunked into this environment that you don't know where you fit in this—in the ecosystem, the clinical realm. And in addition to that, you're forced to develop completely different study techniques. And that is very disorienting. (teacher)
• Climate • Feedback • Randomness • Opportunities to participate	I've seen this, too, where a resident and attending both dislike each other a lot and that creates extra tension. And then you might get a more cohesive team, where everybody agrees really well. So, I think, a lot of that—like things that are out of your control as a medical student can really impact your experience and your evaluation because, if you are intimidated, or nervous, or anxious all the time, it's much harder to perform well versus if you're feeling good and enjoying your environment, and you're lucky enough to be paired with people who give you constant feedback—and opportunities to improve. (student)
• Commonalities • Race/Ethnicity/Culture	… [T]he energy just, like, changes when I'm with another minority student in that we are kind of … "We are in this together, and we're gonna, really look out for one another." And it doesn't feel that way when it's with a white classmate or something like that. … [I]t almost feels like in opposition. Or, that's when I feel most worried or [laughter] afraid of not performing well or not being able to feel my authentic self. (student)
• Race/Ethnicity/Culture • Microaggressions • Climate • Experiences	I noticed that a lot of those who were struggling were students of color. … I think the clinical learning environment may be such that some students might have a hard time focusing, either because of micro-aggressions or as a result of focusing on what just happened, it might be hard for students to actually be present and learn the content—such as, the instructions that one person might give to the student. (teacher)
• Hidden curriculum • Race/Ethnicity/Culture • Transitions • Mentorship	The biggest transition is navigating the social and professional maze that is the beginning of clerkship year. And I think a lot of that comes from just having peers that are showing you, "Here, this is the way. This is how you're supposed to be presenting. This … is what you're supposed to be doing on rounds. This is how you're supposed to act in the OR." A lot of that you learn on the fly, but a lot of that, you've heard from your peers or you're picking up as you go from people you're working with. (student)
Theme 6: Student interactions and reactions	
• Student appearance • Professional behavior • Fit	… when I think of professionalism, I think of like showing up on time. I think of doing everything that was required of you as a team member. I think of communicating effectively, eye contact with people, being accountable. … But I don't think things like … if someone tells you your hair's too curly, or if someone tells you you should shave your beard. Or you shouldn't wear this color with that color 'cause it doesn't look as good. … I think those should just be like if someone dresses a certain way because of their culture, that's their style. That's their flair, and they're able to do their job effectively, then that just makes them a little more unique, and it makes them a little more who they are without having to worry about

	trying to fit into one culture … as a, you know, person of color in my late 20s. If I don't know somebody, but they are different—if they're White, I'm always kind of prepared to be a little more what would be considered like, I don't know, prim and proper and … I guess lose a little bit of my personality in order to kind of fit in, and that changes my ability to perform in the clerkship setting. (student)
• Hidden curriculum • Work style	… I put a lotta effort into sort of playing the game. … I always come off as like, a hard-working, smart person who's always reading up and studying, and always volunteering information to the team even when I'm not asked for it. So I think it was a tremendous amount of effort on my own part. (student)
• Engagement • Language	[A] lot of the softer skills are stuff that maybe you wouldn't consider in an evaluation. ... Like how long somebody spends with a patient. Or if they speak a different language and are able to connect with them, or have shared cultural things with them. Like those aren't ever appreciated. (student)
• Interaction style • Engagement • Bias • Student's comfort to participate • Opportunities to participate • Relationships	I think it's hard when you are the only … Latinx person on your team or Black person on your team. [E]ven if you're really bubbly and outgoing in every other situation, it's really hard to have that voice or find your voice in a situation where you're the only one that looks like you. Especially if you pick up vibes from the team … [Y]ou hear stories all the time about medical students being on a team and the attending may say something that's offensive … like, "Oh, I wish this patient would just learn to speak English," or something like that. And so, that's not going to encourage your medical student, who may speak another language at home, [to] wanna speak up in that setting, right? (teacher)
• Race/Ethnicity/Culture • Language • Bias • Interaction style	… I definitely think there's a difference in terms of how a person of color is graded … [A]s a person of color myself, I know that I have this underlying feeling that … I have to perform better and be more on top of things and adjust my own vocabulary and adjust my own attitude and approach in order to match what is traditionally expected, and even though that may not be my own personal style. (student)
• Interaction style • Race/Ethnicity/Culture	People who come from Asian households are more timid and quiet, and that sometimes reflects negatively when you're on rotations because it's good to be more confident and assertive. And that's just not a value that a lot of Asian households teach their kids. (student)

Figure 1