



GEORGETOWN UNIVERSITY
Center for Research & Fellowships



**Examining the Impact of Race and Gender on Psychiatric Diagnosis and Treatment
Equity: A Computerized Case-Based Survey Instrument**

Ella Shi

Georgetown University Laidlaw Scholars Programme

August 14, 2026

Background

the 1960s, a large body of epidemiological literature has established that Black patients are diagnosed with psychotic disorders at disproportionately higher rates than their white counterparts^{1,2}. Moreover, Black patients are less likely to be prescribed medication for diagnosed psychotic disorders². In the cases where they *are* prescribed, Black patients are more likely to receive an older first-generation antipsychotic medication when compared to their white counterparts³. There is strong evidence of racial bias at each step of the psychiatric evaluation process, from establishing a diagnosis to developing treatment recommendations⁴⁻⁷. The central importance of equity in diagnostic and treatment recommendations is highlighted in the literature, which points out that, in addition to prescription access, diagnostic labels also influence a patient's interactions within broader social systems⁸ (e.g., incarceration, education, occupations, etc.).

Despite this evidence, there is a lack of understanding of the specific underlying mechanisms, clinical elements, causal pathways, and cultural misconceptions (and biases) that drive these disparities at the level of the individual clinical encounters. Therefore, psychiatrists need studies that examine clinical decision-making to better understand the main drivers of racial biases that result in disparate diagnoses and treatment recommendations for Black (and Hispanic) patients, in hopes that eventually, psychiatrists can be trained and educated (trainees and faculty) toward more culturally responsive and equitable clinical practices. Moreover, as medical training programs pilot anti-racism and social justice curricula, programs need to assess behavior changes (e.g., diagnostic and prescription practices) as a result of these educational interventions.

Case-based assessments (or clinical vignettes) are a well-established educational tool in medicine, and they have been shown to document diagnostic bias^{9,10}. Mock clinical vignettes give the investigators the ability to change specific aspects of a case to test the effect of key clinical

elements on the diagnostic decisions. Additionally, case-based assessments allow for the control of certain structural barriers that would be encountered in real-world situations (such as differential access to care). Case-based assessment tools have been employed to demonstrate bias at the level of individual clinical decision-making in various medical specialties. For example, using scripted clinical vignettes (with standardized patients), Schulman et. al (1999) demonstrated that cardiologists are less likely to diagnose chest pain as myocardial infarction and less likely to recommend thrombolytic therapy if patients are Black and female-identifying, even after controlling for lab values.¹¹ Hoffman et. al (2016) utilized mock, or written, case vignettes to show that students' explicit beliefs about racial differences significantly influence their clinical decision-making; in this case, undervaluing and undertreating the Black-identifying patients' pain compared to the white patient.¹² To date, there are no comparable studies in psychiatry that examine bias at the level of clinical decisions using case vignettes.

Further, there are currently no studies in psychiatry that demonstrate specific anti-bias training that changes clinician diagnostic and prescription habits. Broader medical education literature illustrating that standardized patient-based training may help reduce the likelihood of clinicians expressing bias and stereotypes is emerging¹³; in addition, video-based educational interventions have been shown to significantly reduce the stigma associated with certain mental health diagnoses (eg., depression)¹⁴, which highlights the potential for case-based videos (involving standardized patients) as an educational tool to address bias.

This paper serves as a progress report. It documents the design and development phase preceding data collection.

Objectives

The overall goal of this study is to establish and test specific underlying mechanisms of diagnostic and treatment bias using a case-based survey instrument that integrates standardized patient videos. The project has three specific aims: first, to develop a novel computerized case-based survey instrument designed to assess clinical reasoning biases in response to skepticism expressed by patients of different demographic groups; second, to pilot test the instrument with experts in psychotic disorders; and third, to implement the instrument across a range of learners in the medical pipeline, from undergraduate medical students to late-career faculty.

The primary objective is to identify the clinical decision-making factors that produce diagnostic bias when clinicians evaluate mental health symptoms across racial and gender identities. Specifically, do clinicians interpret *reasonable* distrust or skepticism expressed by a Black patient as a sign or symptom of psychosis? The term "reasonable" is used in a defined technical sense: a reaction to lived experience of discrimination or to historical memory of systemic discrimination. The secondary objective is to determine whether patient gender, alongside race, shapes how clinicians interpret a patient's expressed mistrust of the medical system.

Hypotheses

It is hypothesized that clinicians tend to interpret reasonable mistrust expressed by Black patients as a sign of psychosis, particularly under conditions of diagnostic ambiguity, and that this contributes to overdiagnosis of psychotic disorders. It is believed this is due to the following possible mechanisms: clinicians may misunderstand the historical roots of distrust in Black communities; patient race may activate an association to psychosis-adjacent constructs such as paranoia, bizarreness, and severity; race may reduce diagnostic confidence, producing a kind of

diagnostic clouding; or race may activate stigma and negative attitudes about patient characteristics such as likelihood of adherence and symptom fabrication.

Methods

Script Development

The script portrays a 30-year-old patient with no known past medical history who was brought into the emergency department by police for “aggressive behavior” and making threats towards their co-workers. Given the existing literature on vignette-based research approaches to examine diagnostic biases, every parameter of the verbal script was chosen to serve a specific inferential purpose.

The patient monologue will be scripted to last two to three minutes, providing a realistic portrayal of symptoms and concerns. The script will intricately detail affect and body language for standardized patients, ensuring consistency across various identities. The case includes minimal medical information and deliberately excludes racialized diagnoses such as sickle cell disease or cystic fibrosis. Such details could introduce a second pathway by which race could influence judgment, as a clinician might reason from epidemiology rather than from stereotype, making the effect of race difficult to isolate.

The script will deliberately maintain diagnostic ambiguity, presenting a challenging case for physicians. Importantly, the video will exclude the clinician, focusing solely on the patient to emphasize the impact of biases on diagnostic accuracy. The vignette will strategically incorporate overt skepticism from the patient, including skepticism towards the medical system. This careful scripting ensures that biases related to race, skepticism, and diagnostic challenges are effectively captured.

Artificial Intelligence Video Generation

The project was originally designed for the clinical vignettes to be produced through live filming with standardized patients and a professional videography team. However, since the study was initiated, advances in generative artificial intelligence technologies have made it possible to create highly realistic clinical video scenarios that closely approximate professionally filmed encounters. Based on these technological developments, AI-generated videos are sufficiently realistic to serve as appropriate experimental stimuli for testing the study hypotheses regarding clinical decision-making and diagnostic bias.

This modification offers several practical advantages. AI-generated videos substantially reduce the logistical complexity associated with coordinating standardized patients, filming schedules, locations, and videography personnel. The revised approach streamlined production, reduced scheduling barriers, and expedited completion of the video vignettes, allowing the project to progress more efficiently while preserving the overall study design and research objectives. AI generation allows precise specification of affect, pacing, posture, and down to the level of blink rate and breath intake, holding the entire encounter constant while varying only phenotype.

Following this modification, a private professional video company was utilized to assist in creating the AI-generated clinical vignettes based on the developed script. The role of the contractor was limited exclusively to technical production of the video stimuli under the direction and oversight of the study investigators.

Character selection was guided by the literature on racial phenotypicity bias, which shows that stereotype activation tracks the degree to which an individual displays Afrocentric or Eurocentric features, not merely the racial category assigned to them.¹⁵ Characters were therefore

chosen whose skin tone, facial features, hair texture, and apparent age vary substantially across conditions, so that the manipulation is strong enough to detect an effect if one exists. Figure 1 shows the final characters, or patients, chosen. The Black characters have tightly curled hair, a darker skin tone, and fuller lips. The white characters have thinner lips, straighter hair texture, and a narrow nose.

Figure 1. Patients as portrayed through artificial intelligence generation.



Survey Development

In a Qualtrics survey, respondents are randomized to view one of the four videos and then complete approximately 25 closed- and open-ended items drawn from the prominent clinical reasoning theories listed below. The survey is split into multiple sections. 1) Clinical decision-making captures how the clinician reasons toward a diagnosis and a plan, under both a timed constraint and untimed conditions. 2) Patient characteristics capture the clinician's beliefs about the patient. 3) Respondent demographics capture who the clinician is and how that could contribute to diagnostic bias. 4) Explicit bias captures consciously held racial preferences.

Theoretical Frameworks

Situated cognition serves as the umbrella framework. It describes clinical decision-making as shaped by clinical context, including physician factors, patient factors, and specific features of the individual encounter (e.g., inpatient vs. outpatient).¹⁶ Section 2 of the survey, Patient Characteristics, draws on this framework by capturing how a clinician's judgment about the patient is shaped by contextual features of the encounter, such as appearance, affect, and demeanor in the video.

Dual process theory distinguishes clinical reasoning as a product of two interacting systems: System 1, fast and intuitive pattern recognition, and System 2, slower analytic reasoning.¹⁷ Stereotypes and biases most quickly appear when operating in System 1. As such, section 1 of the survey, clinician decision-making, is split up into a timed and untimed condition. The timed condition pushes the clinician to intuitively generate a hypothesis and to quickly produce a differential diagnosis based on pattern recognition. The untimed condition allows the clinician to deliberately assess and quantify their diagnostic certainty. Comparing the two shows whether a biased first impression gets corrected, kept, or justified once the clinician has time to think.

Signal detection theory characterizes clinical reasoning as a process of separating meaningful data (or signals) from non-meaningful data (or noise).¹⁸ By understanding what information is being retained as salient, it can be determined whether the same details are being weighed differently depending on who the patient is. The survey therefore asks respondents to name the information that most informed their diagnosis and to explain how they reached their differential. These responses are coded for which details were treated as signals and compared across the four vignettes.

Finally, van Ryn et. al (2011) model how implicit and explicit racial biases and stereotypes generate diagnostic bias, where interactions occur between both patient and clinician characteristics.¹⁹ By collecting clinician attributes and attitudes alongside diagnostic responses, the study can determine whether differences in judgment vary with characteristics of the clinician as well as of the patient.

Next Steps

With design and development complete, recruitment begins in fall 2026. Feedback will first be collected from external experts in psychotic disorders to establish content validity, followed by pilot testing to assess feasibility. Broader dissemination will draw from across the medical education pipeline, from medical students to late-career faculty, recruited both within the MedStar Georgetown Department of Psychiatry and externally through partner institutions (pilot sample $n=30$; national sample $n\geq 100$). Analysis will include both a quantitative (ANOVA, chi-squared) and a qualitative portion (inductive vs deductive thematic coding). By pinpointing which aspects of clinical reasoning are most vulnerable to bias, we aim to identify the specific cognitive mechanisms behind diagnostic bias, yielding a behavioral measure for anti-racism and bias-reduction training and supporting more equitable diagnostic practice.

Acknowledgements

This project was made possible by the support of the Laidlaw Foundation and Georgetown University's Center for Research and Fellowships. Thank you to my Principal Investigator, Dr. Williams, for his mentorship, guidance, and encouragement throughout. Thank you also to Rachel, Ashley, Dr. Saunders, and Dr. Krishnan for their feedback and project management support.

References

1. Schwartz RC, Blankenship DM. Racial disparities in psychotic disorder diagnosis: a review of empirical literature. *World J Psychiatry*. 2014;4(4):133–140. doi:10.5498/wjp.v4.i4.133
2. Faber SC, Khanna Roy A, Michaels TI, Williams MT. The weaponization of medicine: early psychosis in the Black community and the need for racially informed mental healthcare. *Front Psychiatry*. 2023;14:1098292. doi:10.3389/fpsy.2023.1098292
3. Kuno E, Rothbard AB. Racial disparities in antipsychotic prescription patterns for patients with schizophrenia. *Am J Psychiatry*. 2002;159(4):567–572. doi:10.1176/appi.ajp.159.4.567
4. Olbert CM, Nagendra A, Buck B. Meta-analysis of Black vs. White racial disparity in schizophrenia diagnosis in the United States: do structured assessments attenuate racial disparities? *J Abnorm Psychol*. 2018;127(1):104–115. doi:10.1037/abn0000309
5. Neighbors HW, Trierweiler SJ, Ford BC, Muroff JR. Racial differences in DSM diagnosis using a semi-structured instrument: the importance of clinical judgment in the diagnosis of African Americans. *J Health Soc Behav*. 2003;44(3):237–256.
6. Eack SM, Baborik AL, Newhill CE, Neighbors HW, Davis LE. Interviewer-perceived honesty as a mediator of racial disparities in the diagnosis of schizophrenia. *Psychiatr Serv*. 2012;63(9):875–880. doi:10.1176/appi.ps.201100388
7. Gara MA, Vega WA, Arndt S, et al. Influence of patient race and ethnicity on clinical assessment in patients with affective disorders. *Arch Gen Psychiatry*. 2012;69(6):593–600.
8. Link BG, Phelan JC. Conceptualizing stigma. *Annu Rev Sociol*. 2001;27:363–385. doi:10.1146/annurev.soc.27.1.363

9. Peabody JW, Luck J, Glassman P, Dresselhaus TR, Lee M. Comparison of vignettes, standardized patients, and chart abstraction: a prospective validation study of 3 methods for measuring quality. *JAMA*. 2000;283(13):1715–1722. doi:10.1001/jama.283.13.1715
10. Kales HC, Neighbors HW, Blow FC, et al. Race, gender, and psychiatrists' diagnosis and treatment of major depression among elderly patients. *Psychiatr Serv*. 2005;56(6):721–728. doi:10.1176/appi.ps.56.6.721
11. Schulman KA, Berlin JA, Harless W, et al. The effect of race and sex on physicians' recommendations for cardiac catheterization. *N Engl J Med*. 1999;340(8):618–626. doi:10.1056/NEJM199902253400806
12. Hoffman KM, Trawalter S, Axt JR, Oliver MN. Racial bias in pain assessment and treatment recommendations, and false beliefs about biological differences between blacks and whites. *Proc Natl Acad Sci U S A*. 2016;113(16):4296–4301. doi:10.1073/pnas.1516047113
13. Blatt B, LeLacheur SF, Galinsky AD, Simmens SJ, Greenberg L. Does perspective-taking increase patient satisfaction in medical encounters? *Acad Med*. 2010;85(9):1445–1452. doi:10.1097/ACM.0b013e3181eae5ec
14. Amsalem D, Martin A. Reducing depression-related stigma and increasing treatment seeking among adolescents: randomized controlled trial of a brief video intervention. *J Child Psychol Psychiatry*. 2022;63(2):210–217. doi:10.1111/jcpp.13427
15. Maddox, Keith. (2004). Perspectives on Racial Phenotypicality Bias. *Personality and Social Psychology Review*. 8. 383-401. 10.1207/s15327957pspr0804_4.
16. Wilson BG, Myers KM. Situated cognition in theoretical and practical context. In: Jonassen D, Land S, editors. *Theoretical Foundations of Learning Environments*. Mahwah, NJ: Erlbaum; 1999. pp. 57–88

17. Bellini-Leite SC (2022) Dual Process Theory: Embodied and Predictive; Symbolic and Classical. *Front. Psychol.* 13:805386. doi: 10.3389/fpsyg.2022.805386
18. Swets, J. A., Dawes, R. M., & Monahan, J. (2000). Psychological science can improve diagnostic decisions. *Psychological Science in the Public Interest*, 1(1), 1–26.
doi:10.1111/1529-1006.001
19. van Ryn M, Burgess DJ, Dovidio JF, et al. THE IMPACT OF RACISM ON CLINICIAN COGNITION, BEHAVIOR, AND CLINICAL DECISION MAKING. *Du Bois Review: Social Science Research on Race*. 2011;8(1):199-218. doi:10.1017/S1742058X11000191