“Sorry, I’m Not Accepting New Patients”: An Audit Study of Access to Mental Health Care

Heather Kugelmass

Abstract
Through a phone-based field experiment, I investigated the effect of mental help seekers’ race, class, and gender on the accessibility of psychotherapists. Three hundred and twenty psychotherapists each received voicemail messages from one black middle-class and one white middle-class help seeker, or from one black working-class and one white working-class help seeker, requesting an appointment. The results revealed an otherwise invisible form of discrimination. Middle-class help seekers had appointment offer rates almost three times higher than their working-class counterparts. Race differences emerged only among middle-class help-seekers, with blacks considerably less likely than whites to be offered an appointment. Average appointment offer rates were equivalent across gender, but women were favored over men for appointment offers in their preferred time range.

Keywords
bias, field experiment, mental health care, race, social class

Despite health care providers’ avowed professional ideal toward egalitarianism, a large body of research has demonstrated that patients’ social characteristics influence the health care encounter. Laboratory experiments have persuasively shown that health care providers’ decisions about patients are shaped by race, social class, and gender biases. However, few studies have measured decisions about access, investigated mental health care providers, or applied a field experimental approach. To my knowledge, no previous work has incorporated all three elements. The present study attempts to remedy this gap in the literature by asking: Do psychotherapists offer equal accessibility to all help seekers regardless of race, class, and gender?

Studies of health care providers’ aversion to, or preferences for, certain help seekers tend to focus on physicians, not psychotherapists, and on diagnosis and treatment rather than access issues (Arber et al. 2006; Kikano, Schiaffino, and Zyzanski 1996; Lutfey et al. 2009; McKinlay, Potter, and Feldman 1996). As with studies of bias among physicians, most studies of biased mental health care decision making examine diagnostic impressions instead of access (Blow et al. 2004; Lee and Temerlin 1970; Loring and Powell 1988; Martin 1993).

Yet, disparities in mental health care access loom large, with African Americans and lower/-working-class individuals facing sizable disadvantages to receipt of treatment, even after controlling insurance coverage (Fiscella et al. 2000; Padgett et al. 1994). Similar to physicians, mental health care providers likely have psychological biases that can contribute to limited access for these negatively stereotyped groups.

1Princeton University, Princeton, NJ, USA

Corresponding Author:
Heather Kugelmass, Princeton University, 107 Wallace Hall, Princeton, NJ 08544, USA.
E-mail: hkugelma@princeton.edu
Studying the behavioral manifestations of clinician biases—regardless of access, diagnosis, or patient management—presents ethical, logistical, and empirical challenges. Consequently, most studies of clinical encounters are constrained by the lack of realism inherent in a laboratory setting. Audit studies—a type of experiment rarely applied to health care providers—are a valuable alternative because they permit the direct observation of decisions that are made in real-world social contexts. In laboratory experiments, recruited health care providers are presented with written descriptions or videos of patients they know to be hypothetical. Audit studies can complement the findings of these studies by targeting providers in their daily professional setting and exposing them to help seekers whom they perceive as real.

The present study employs a two-wave, phone-based audit experiment to detect disparities in psychotherapists’ accessibility to psychological help seekers. It broadens the discussion of clinician bias by targeting a different class of health care providers and measuring a different outcome than most research on this topic.

The results reveal striking differences in psychotherapists’ accessibility to help seekers based on both race and social class. For example, when an identifiably black working-class man with health insurance called 80 therapists in his insurance network to request a weekday evening appointment, only one call elicited an offer. In contrast, 20% of the calls made by a white middle-class woman—with identical insurance coverage—elicited a comparable offer. The profound differences in accessibility revealed by this study are consistent with prior research and theory: providers’ biases affect their behavior in ways that systematically disadvantage some groups of help seekers.

BACKGROUND

Biased Decision Making among Health Care Providers

The past two decades have seen a spate of studies demonstrating that health care providers’ nonclinical biases influence their perceptions of patients and their consequent decisions. Researchers have used experimental methods to measure the extent to which a wide variety of provider behaviors, for example, diagnosis, certainty of diagnosis, treatment recommendations, and patient management, vary as a function of patient attributes (e.g., Arber et al. 2006; Green et al. 2007; Haider et al. 2011; Kikano et al. 1996; Lutfey et al. 2008, 2009, 2010; McKinlay et al. 1997, 1996; Stepanikova 2012). Among the most commonly studied attributes are race, class, and gender. Many studies also consider how provider characteristics (e.g., work setting, specialty, age, race, gender, and years of experience) relate to providers’ decisions about patients. Consistent with this tradition, the present study will experimentally examine the influence of help-seeker race, class, and gender while statistically controlling for provider gender, years of experience, professional degree, and location.

There is ample evidence that racial discrimination during encounters with health care providers continues to persist in the United States (see review by Shavers et al. 2012). Despite health care providers’ explicit endorsement of racial equity, they have a strong prowhite implicit (i.e., nonconscious) bias, similar to that observed in occupationally heterogeneous samples of Americans (Haider et al. 2011; Sabin et al. 2009). Indeed, it is not uncommon for well-educated whites to hold explicitly egalitarian beliefs while harboring nonconscious stereotypes about out-groups (Dovidio et al. 2008). It follows that salient out-group attributes, such as race and social class, can trigger providers’ stereotypes. These stereotypes and other sources of bias, in turn, influence their decisions about whether to extend offers of care.

Van Ryn and Burke (2000) found that physicians ascribe negative characteristics to blacks and lower-class patients. Physicians expressed less affiliative feelings toward black patients and associated lower-class patients with negative personality traits (irrationality, low self-control). Both were perceived as less intelligent than their white and upper-class counterparts and at higher risk for noncompliance with treatment. In a similar vein, Green et al. (2007) found that blacks are implicitly perceived by physicians as less cooperative—both in a medical context and in general. Mental health care providers’ impressions of help seekers are informed by similar racial stereotypes. For example, Abreu (1999) demonstrated that priming therapists with African American stereotypes (vs. no prime) led them to rate a hypothetical patient as more hostile.

The most compelling studies causally link providers’ implicit biases to tangible outcomes via a wide range of experimental methods. In one of the earliest such experiments, McKinlay et al. (1996) presented physicians with a video vignette of hypothetical patients that varied by sex, race, age, health insurance coverage, and socioeconomic level. Using this
method, researchers were able to identify nonclinical sources of variation in subjects’ assessments of diagnosis, treatment, and prognosis. In a more recent study, Stepanikova (2012) used subliminal priming to activate implicit racial biases, which enabled her to determine that patient race and physician stress interact to influence clinical decisions. Notably, although patient race has been more frequently studied than social class, van Ryn and Fu (2003) assert that class is likely to be as strong a determinant of health care provider behaviors. The evidence is clear: stereotypes (race based or otherwise) shape providers’ decisions about care. The present study builds upon this body of evidence in three directions: population (therapists), outcome (access), and method (audit study).

**Psychotherapists**

Health care providers with high levels of professional autonomy are powerful gatekeepers to care. Their level of discretion to restrict access partly depends on the institutional setting in which they operate (Chiarello 2013). While less than 20% of physicians are in solo practice, approximately half of mental health care providers are (American Psychological Association 2009; Kane and Emmons 2013). Psychotherapists—overrepresented in solo practice among mental health care providers—have ample opportunity to make decisions consistent with their biases because they retain exclusive discretion over the provision of their services.

Research suggests that psychotherapists (hereafter also called “therapists”) favor help seekers with the “YAVIS” attributes: young, attractive, verbal, intelligent, and successful (Tryon 1986). Consistent with the YAVIS hypothesis, Teasdale and Hill (2006) found that therapists prefer “psychologically minded” clients and those who share similar values and attitudes. These effects were independent of the demographic characteristics (including race) of the help seekers, but the results were survey based, so social desirability pressures may have influenced the results. In another study, black patients were rated by psychiatrists as “less psychologically minded” as well as “less articulate, competent, [and] introspective” than otherwise equivalent white patients (Geller 1988:124). It is possible, then, that stereotypes linked to blackness are rationalized through, are partially mediated by, or interact with stereotypes associated with a lack of “psychological mindedness,” thereby reproducing discrimination against African Americans in the mental health care sphere.

The influence of help-seeker social class on the perceptions and behavior of psychotherapists was studied extensively in the 1960s and 1970s. The results indicated that therapists’ initial impressions of low-class help seekers are tainted with a negative bias, informed by stereotypes of low-class help seekers as hostile and untreatable (see review by Lorion 1974). The effect of social class has also been highlighted in a recent vignette-based experiment, which found that therapists in training perceived hypothetical working-class and poor help seekers as more unpleasant to work with (Smith et al. 2011). Biases such as these—conscious or non-conscious—could operate in subtle ways to influence therapists’ decisions regarding if and how to respond to help seekers’ requests for care.

**Access**

Studies consistently show that African Americans have higher rates of unmet need for mental health care than whites; similarly, poor and near-poor Americans have lower rates of mental health service usage (Broman 2012; Snowden and Yamada 2005; Wang, Lane, et al. 2005). Blacks are no less likely (and sometimes more likely) to express a willingness to seek mental health care than whites are (Schnitkner, Pescosolido, and Croghan 2005; Shim et al. 2009). Moreover, some studies have found racial gaps and class gaps in receipt of mental health care even among the insured (Fiscella et al. 2000; Padgett et al. 1994). These findings point toward the existence of provider-generated obstacles to access.

Yet, research on stereotypes’ influence on access is rare. The majority of rigorous research on disparate treatment by physicians has focused on discriminatory decisions made during or following a clinical encounter, not prior to it (Fennell 2005). Similarly, investigations of provider bias in the mental health care sphere typically center on clinical impressions (e.g., Blow et al. 2004; Loring and Powell 1988; Young and Powell 1985). However, questions of bias in diagnosis and patient management are secondary to questions of access because the former presuppose a clinical encounter that only a subset of disadvantaged help seekers will obtain if there exists systematic bias in access. I posit that racial and class bias influence decision making at the pre-encounter stage—namely, at the first request for care. The scant research on this topic supports differences by help-seeker race and class, even in the absence of financial incentives for discrimination (Olah, Gaisano, and Hwang 2013; Wang, Berglund, et al. 2005). Like medical care providers, therapists’ decisions regarding whether
and how to respond to help seekers’ initial requests for care likely depend on the perceptions of the social categories ascribed to them.

Audit Studies

One critique of early work on health care providers’ perceptions of patients is that it was measured by self-reports, which are confounded by social-desirability responding and can capture only conscious biases. As attention turned from conscious to non-conscious bias, researchers increasingly employed Implicit Association Tests (e.g., Haider et al. 2011; Krieger et al. 2010; Sabin et al. 2009; Sabin, Rivara, and Greenwald 2008) and subliminal priming (e.g., Stepanikova 2012). Clinicians’ underlying biases have been linked to behavioral outcomes through the use of ratings of doctor–patient interactions (Penner et al. 2010), written vignettes (e.g., Green et al. 2007; Haider et al. 2011; Kikano et al. 1996), or videotaped vignettes (Arber et al. 2006; Lutfey et al. 2009, 2010; McKinlay et al. 1996, 1997) depicting hypothetical patients.

Although vignette-based experiments have provided strong evidence of the influence of patients’ race and class on providers’ behaviors, they are necessarily artificial and simplified representations of complex decision-making contexts. They do not enable conclusions about the extent to which results would generalize to a large, randomly selected sample of subjects in real-world settings. A real-world setting is particularly important for the study of psychotherapists, whose everyday decisions about access are unmonitored by colleagues or staff. A non–convenience sample is valuable because, as noted by Dovidio et al. (2008), health care providers may be reluctant to voluntarily participate in research on the topic of racial bias; such studies have the potential for legal and personal repercussions, should bias be uncovered. Low participation rates among health care providers in laboratory studies are common, with some as low as 2% (Stepanikova 2012). Audit studies do not require the consent of subjects; participation rates therefore are, by definition, 100%.

Audit studies, a type of field experiment, enable researchers to systematically test for otherwise unobservable discrimination in access. In audit studies, subjects (psychotherapists) are unknowingly exposed to auditors (help seekers) who are equivalent on all characteristics save for those manipulated by the research (race, class, and gender). In audit studies of racial discrimination, for example, if the racial-minority auditors receive lower rates of access to opportunities than the white auditors, then discrimination is said to have occurred (Pager and Shepherd 2008).

Audit studies—which can be written, in person, or telephone based—have supplied ample evidence of race and class discrimination in labor and housing markets, among others (see reviews by Pager 2007; Pager and Shepherd 2008; Quillian 2006; Ross and Turner 2005). Massey and Lundy (2001) were among the first to employ a telephone-based audit study; they found large differences by race and class in access to rental agents and housing (see also Fischer and Massey 2004; Purnell, Idsardi, and Baugh 1999). Their study has several important implications for the present research. First, it demonstrated that Americans can readily recognize black speakers over the phone and can identify them as of middle- or lower-class origins (see also Doss and Gross 1994; Feagin 1994; Purnell et al. 1999). Second, it experimentally confirmed that black and lower-class individuals continue to experience subtle forms of discrimination by professionals, even without in-person contact. And third, most of the discrimination they observed occurred through blocking access rather than other means, such as charging higher prices.

Recently, phone-based audit studies have been used to detect discrimination among gatekeepers in the health care sector as well, with a focus on socioeconomic status (e.g., Bisgaier and Rhodes 2011; Saloner et al. 2015). For example, one recent audit found that office staff are less likely to offer a primary care appointment to lower-class patients than to middle-class ones, even under Canada’s universal health care system (i.e., insurance coverage held constant; Olah et al. 2013). Mental health care providers may similarly discriminate by class independent of insurance status. Based on the evidence reviewed from both field and laboratory experiments, I predict lower rates of access to African American and working-class help seekers as compared to white and middle-class help seekers.

DATA AND METHODS

Sample

The subjects of this experiment were psychotherapists in New York City selected through systematic sampling from the directory of a single large health insurance provider’s HMO plan. New York City was chosen because it is a racially and socioeconomically diverse urban area with a high concentration of mental health practitioners. The sample is
restricted to licensed psychotherapists who have a solo practice and a PhD or PsyD degree (i.e., it does not include psychiatrists, who have medical degrees). Isolating solo practitioners permits direct identification of potential biases of and discrimination by providers, unmediated by office staff. In this way the study maintains comparability to laboratory experiments on health care providers.

Information about the psychotherapists’ race is not available. However, it is unlikely that there was a sufficiently large number of racial-minority psychotherapists in the sample to influence the results. Blacks constitute only 3% of licensed psychologists nationwide (Michalski, Mulvey, and Kohout 2010). Statistics specific to New York City are unavailable, and it is possible that the population of psychologists is more racially diverse in this metropolitan area. If so, the inclusion of black therapists would merely mute observed discriminatory effects.

**Experimental Protocol**

Voice-over artists recorded scripted messages using racially distinctive names and adopting specified race- and class-based speech patterns. I drew from the large pool of survey takers on Amazon Mechanical Turk (an online crowdsourcing marketplace) to select final recordings for the experiment. Both of the white female help seekers were voiced by one actress, as were both of the black female help seekers. Similarly, one male actor voiced both white conditions, and another voiced both black conditions. This allowed me to control unobserved variation that could have introduced error to estimates of differences between classes.

Survey takers on Mechanical Turk listened to the recordings of the voice messages and responded to a series of questions about them. These questions facilitated the selection of voices with the highest race and class agreement and authenticity. They also confirmed that audio quality was high across all recordings. These recordings served as the experimental manipulation.

On the recordings, the help seekers each mentioned symptoms of depression or anxiety, named the same health insurance plan, requested an appointment, and indicated a preference for a weekday evening. The help seeker requested that the therapist leave a voicemail indicating available appointment slots. The phone numbers provided to the therapists corresponded to unique voicemail boxes for each fictitious help seeker. Two middle-class scripts and two working-class scripts were employed. The four scripts were designed to be substantively equivalent but were dissimilar on minor details so as not to arouse suspicion that could compromise the psychotherapists’ blindness to the experiment. Script presentation order was randomized and counterbalanced across study waves. (See the appendix for call scripts in the online supplemental material at http://jhsb.sagepub.com/supplemental.)

Three help-seeker characteristics were systematically manipulated: social class (middle or working), gender (female or male), and race (black or white). I
employed a partial within-subjects design, with each psychotherapist randomly exposed to two of eight conditions: help seekers (callers) of the same class and gender but a different race (Figure 1).

A sample of therapists from Empire Blue Cross Blue Shield’s in-network directory was randomly divided into four groups. During Wave 1, the groups received messages from one of the following help seekers: middle-class female, middle-class male, working-class female, working-class male. Race was alternated within each group. Therapists were called until 80 messages per condition were placed.

During Wave 2, one month later, each of the therapists received a similar message from a different caller of the same social class and gender but a different race. For example, if a therapist was exposed to the white middle-class male condition in November, he was exposed to the black middle-class male condition in December. Therefore, a total of 640 calls were placed (2 per each of the 320 subjects).

Research assistants placed calls at night to minimize the number of therapists who answered their phones. Potential subjects with office staff, or those in group practices, were systematically screened out of the sample. Messages were left for, and received from, therapists only. Google Voice (an Internet-based call management service) was used to place calls (i.e., play the recordings for answering machines) and to collect returned voicemail messages.

**Variables**

**Independent Variables.** This study manipulated the race, social class, and gender of the help seeker. Race (black or white) was communicated through racially distinctive names and linguistic styles. Social class (middle or working class) was conveyed primarily by vocabulary and grammar. The fictitious black working-class help seekers spoke in Black English Vernacular, which is a linguistic pattern consisting of black-inflected pronunciation as well as nonstandard grammar and diction. It communicates low socioeconomic status (Rahman 2008). Black Accented English, used by the middle-class black callers, is differentiated from White Standard English by black-inflected pronunciation only (Rahman 2008). The fictitious working-class white help seeking callers used low-level vocabulary and grammar that do not comply with the rules of Standard English. They spoke with a heavy New York City accent.

**Covariates.** I considered four characteristics of the psychotherapists: type of doctoral degree (PsyD or PhD), office location, gender, and number of years in practice. Type of degree and office location are publicly available online through Empire Blue Cross Blue Shield’s records. Seventeen percent of therapists in the sample held a PsyD, which is oriented to applied clinical practice, as compared to the more research-oriented PhD. Office location was measured by two variables: New York City borough and distance in miles from midtown Manhattan. In the interest of parsimony, regression models controlled for location in the form of dichotomized borough: Manhattan (78%) was coded 1 and outer boroughs coded 0. For subjects with gender-ambiguous names, gender was determined based on the therapist’s voice on the answering machine message. Fifty-six percent of the sample were female.

Number of years in practice (mean = 18 years) was derived from the therapist’s licensure year, which is publicly available through the New York State’s Office of Professions. This may be interpreted as a proxy for age, clinical experience, or length of exposure to the New York City pool of help seekers.

Years were aggregated into three categories: fewer than 10, 10 to 29, and 30 or more.

**Accessibility.** The main dependent variable was therapist accessibility. Accessibility refers to the extent to which, after responding to the help seeker, the psychotherapist enables access to his or her services. It was operationalized through an appointment offer rate. Each of the 287 voice messages received from a therapist was assigned one of five mutually exclusive codes: (1) The message clearly stated that there are no appointments available, for example, “I am not accepting any new patients.” Twenty-nine percent of messages received fell into this category. (2) A similar response, which characterized 6% of the total messages received, indicated that no appointments were available during the requested time frame but did not address the possibility of appointments during the day or on weekends (e.g., “Sorry I don’t have any availability during the time that you want”). (3) A third type of messages did not address the request for appointments at all, such as “Please call me back.” This category, which constituted 31% of all responses, poses a challenge for assessing therapists’ intent with regard to the eventual offer of appointments. It is possible that these messages are left by therapists reserving judgment about appointment offers before conversing with the help seeker. (4) A fourth category, composing 14% of the sample, included all messages that referenced some availability but either outright denied appointments during the
preferred period or did not address that possibility (e.g., “I have some openings on Tuesday after-
noon”). (5) In the last category are responses that contained an implied or an unambiguous offer of an appointment during one or more weekday evenings (e.g., “I can see you Monday or Wednesday at 6
p.m.”). Twenty percent of messages met this criterion.6

For statistical analyses I considered two vari-
tions of accessibility. The first variable represents any appointment availability. It is a dichotomous variable where messages of type 4 or 5 described above are coded 1. Messages of type 1, 2, or 3, as well as calls that did not elicit a response (noncall-
backs), are coded 0. The second variable (preferred appointments) is a dichotomous indicator of avail-
ability during the time frame requested by the help seeker (weekday evenings). Only messages of type 5 were coded 1 for this variable, which represents very favorable responses. This second variable sets a higher threshold than the first; that is, any response coded 1 for the second was also coded 1 for the first. The effects of race, class, and gender on each outcome are presented in the section that follows.

Responsiveness. Responsiveness is a secondary dependent variable, intended to identify one of the steps contributing to the rates of appointment offers. Responsiveness was measured by callback rates. Responses of any type (i.e., all categories 1–5 as described above) were compared to the failure to respond at all. A callback is a necessary but not suf-
ficient condition for an appointment offer. A mes-
sage for a therapist was said to have elicited a callback if the therapist left a voicemail for the help seeker.7

RESULTS

Descriptive Statistics

Across all conditions and waves, a total of 287 return messages were received from therapists, repre-
senting 44% of calls placed (n = 640) by help seekers.5 Fifteen percent of the 640 calls placed elic-
ited an appointment offer (n = 97). This translates to 34% of return messages (n = 287) received by help seekers.

Figure 2a illustrates descriptive results pertaining to therapist accessibility, by class and race. It shows that white middle-class help seekers have a sizable advantage over the other groups, with 28% (n = 45/160) of their calls to therapists resulting in an appointment offer. In contrast, only 17% (n = 27/160) of the calls placed by black middle-class help seekers did. This is striking in light of the partial within-subjects design of this study; these pairs of matched black and white help seekers called the same individual therapists. The racial disparities were more pronounced for middle-class men than for middle-class women. All middle-class pairs far surpassed the working-class callers, who, regard-
less of race, had a success rate of only 8% (white, n = 13/160; black, n = 12/160). Race appears to influence access but only for middle-class help seekers. Figure 3 shows that the patterns are similar for women and men. Figure 4 visually displays this information in a different way. The point estimates on this graph represent the odds of each character being offered any appointment relative to the odds of the middle-class white female help seeker. It shows, for example, that the lower-class male help seekers were approximately 20% as likely to receive an offer as the middle-class white female help seeker.

Figure 2b displays results for very favorable (i.e., preferred appointment time frame) responses, which constitute 9% (n = 57) of the 640 calls placed. Of all the messages that offered any appointment (n = 97), 59% provided appoint-
ment availability within the preferred time frame. The patterns across race and class are remarkably similar to those seen in Figure 2a. Among middle-
class callers, whites were strongly favored over blacks; this disparity is even larger than emerged in Figure 2a. Therapists expressed a clear prefer-
ence for middle-class help seekers over working-
class ones. Only one preferred appointment was offered to the working-class black man; this is out of a total of 80 therapists to whom he reached out. Figure 5 displays the results disaggregated by gender; here, the female help seekers received more offers than the men under all race and class conditions.

Logistic Regression

Each of the four models presented in Table 1 is com-
posed entirely of dummy variables. Findings are reported as odds ratios, which were obtained by expo-
nentiating the regression coefficients. The white middle-
class male help seeker serves as the referent category, with an odds ratio of 1.0. The observations in the data set are not independent, so the standard errors are corrected for clustering on therapist ID.9 The inclusion of variables for script and study wave do not alter the results of these models and are
therefore omitted from the models in the interest of parsimony.

Models 1a (any appointment) and 1b (preferred appointment) explored differences in accessibility across experimental conditions; they were conditional on race, class, and gender. These models confirm the class and race effects that were evident in Figures 2 and 3. Blacks, on average, are approximately 40% less likely to receive an appointment offer than whites ($p < .05$), and working-class callers are almost 70% less likely to receive an appointment offer than middle-class callers ($p < .001$). Models stratified by gender, not displayed in this table, confirm that those racial differences exist only among the middle-class help seekers.

In Model 2b one sees help-seeker gender emerge as significant, with women favored over men at a rate of two to one ($p < .05$). Race and class continue to be significant, with magnitudes similar to Model 2a. The disparity between white and black is slightly larger (odds ratio = .54, $p < .05$) and between middle and lower class is slightly smaller (odds ratio = .35, $p < .01$). It is evident that the best appointments (weekday evenings) are reserved for white middle-class women.

To examine the race-by-class interactions evident in Figures 2 and 3, I ran an additional series of models (not displayed here), on both outcome measures, that included three race–class groups and alternated the omitted category. As expected, there are significant interactions between race and class, with differences emerging between all groups except the black and white working-class callers.

After adjusting for clustered standard errors and gender, both black and white working-class help seekers had approximately one fifth the odds of being offered an appointment at any day or time, relative to the white middle-class help seekers ($p < .001$). For weekday evenings, the working-class odds were each approximately one quarter of the white middle class ($p < .001$).

I explored the potential effects of therapist traits: gender, type of doctoral degree, location, and years in practice. Therapist characteristics did not influence appointment offers, with one exception: help seekers were considerably more likely to receive an appointment offer from therapists who practice in Manhattan rather than the outer boroughs of New York City. This is the case for any appointment (odds ratio = 2.77, $p < .01$) as well as a preferred appointment (odds ratio = 3.69, $p < .05$). However, office location does not mediate relationships between the outcomes of interest and race or class, with coefficients and significance levels that change very little from Models 2a and 2b.$^{10}$

**Responsiveness (Callbacks)**

In this study, receipt of an appointment offer is the ultimate outcome. I conducted supplementary analyses to provide additional perspective on the noncallback cases. In the accessibility measures, noncallbacks were considered equivalent to categories 1 through 3 (i.e., nonaccessibility). In the responsiveness measure, responses of any type are compared to the failure to respond at all. This contrast enables me to determine if
the discrimination could be traced to a failure to call back help seekers.

All else equal, blacks were less likely to receive a callback than whites (odds ratio = .76, p < .05), and working-class help seekers were less likely to receive a callback than middle-class ones (odds ratio = .65, p < .05). (There were no differences by gender.) The differences in callback rates were more muted than the effects presented in the accessibility analyses, suggesting that the discriminatory effect is not driven entirely by a lack of response. The percentages of help-seeker calls that elicited a therapist callback are displayed in Figure 6.

**DISCUSSION**

**Summary**

The results presented here provide strong prima facie evidence of racial and class discrimination by psychotherapists. This field experiment largely confirms the hypotheses that help seekers who are black or working class are at a disadvantage with regard to psychotherapists’ accessibility. These results comport with extant studies that demonstrate the persistence of discrimination by health care providers, despite the assumption that those who select such professions have a strong commitment to equity. Moreover, this research demonstrates that audit studies of health care providers can be executed in ethical, precise, and low-cost ways; this powerful method need not be relegated to the realms of real estate or labor markets.

White middle-class help seekers were significantly more likely to be offered an appointment than black middle-class, black working-class, or white working-class help seekers. Black middle-class help seekers also have a considerable advantage over black working-class and white working-class help seekers. Therapists were more accommodating of female help seekers’ request for a weekday evening appointment than they were for men’s requests. The most remarkable disparities in accessibility exist between the white middle-class and the

**Figure 3.** Percentages of Help-seeker Calls (n = 80 per group) That Elicited at Least One Appointment.  
*Note:* Race, class, and race-by-class interaction effects are statistically significant. Gender differences are not.

**Figure 4.** Odds of receiving any appointment, relative to middle-class white female (n = 640).  
*Note:* The referent group is the middle-class white female, with an odds ratio of 1.0 (horizontal line). Ninety-five percent confidence interval bands are displayed.
working-class help seekers. Therapists’ aversion to working-class help seekers overall is color-blind; that is, their appointment offer rates are indistinguishable.

Accessibility rates were virtually indistinguishable between black and white working-class callers; yet, for responsiveness rates, it is the middle-class rates that do not vary by race. Broadly, therapists in the middle-class experimental condition who decide to call back anyone at all are more likely to give both black and white help seekers the courtesy of a callback than they are to go the extra step of offering appointments to both help seekers. It is possible that therapists were undecided when they called back black help seekers, intending to conduct additional screening (e.g., to determine articulateness) that would inform their decision.

Working-class help seekers were less likely to be extended the courtesy of a callback. Callbacks are valuable even when accompanied by rejection because they can facilitate the help seekers’ search by offering reassurance that they should in fact seek treatment or by offering useful information, such as a referral to another therapist. Moreover, regardless of the content of the message, the receipt of a callback could reduce distrust of mental health care providers or disillusionment with the mental health care system more broadly.

### Mechanisms

The audit study method enables precise estimates of discrimination with high internal validity. The method is less well suited for determining the types of reasoning that drive discrimination. I will explore two classes of possible mechanisms that may contribute to the observed discriminatory effects: (1) implicit (nonconscious) and (2) deliberate (conscious).

This study supports earlier findings that the implicit (nonconscious) out-group biases of clinicians are one mechanism through which disparities emerge. A variety of stereotypes about blacks’ low intelligence, high hostility, and reluctance to comply with treatment suggestions could lead to aversion and avoidance (Abreu 1999; Dovidio et al. 2008; van Ryn and Burke 2000). The private circumstances under which psychotherapists make their decisions about access may be particularly conducive to the emergence of nonconscious biases that lead to their discriminatory accessibility. Given the low overall callback rate—even middle-class

### Table 1. Odds Ratios for (a) Receiving Any Appointment Offer and (b) Receiving an Appointment within the Preferred Time Frame.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Model 1 (n = 640)</th>
<th>Model 2 (n = 640)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1a: Any Appt</td>
<td>1b: Preferred Appt</td>
</tr>
<tr>
<td>Help seeker</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>.61* (.12)</td>
<td>.54* (.13)</td>
</tr>
<tr>
<td>Working class</td>
<td>.29*** (.08)</td>
<td>.35*** (.13)</td>
</tr>
<tr>
<td>Female</td>
<td>1.40 (.37)</td>
<td>2.00* (.66)</td>
</tr>
<tr>
<td>Therapist</td>
<td></td>
<td></td>
</tr>
<tr>
<td>30+ years experience</td>
<td>.85 (.26)</td>
<td>.86 (.33)</td>
</tr>
<tr>
<td>PsyD degree</td>
<td>.62 (.17)</td>
<td>.60 (.20)</td>
</tr>
<tr>
<td>Manhattan</td>
<td>2.77*** (.00)</td>
<td>3.69* (.20)</td>
</tr>
<tr>
<td>Constant</td>
<td>.31*** (.13)</td>
<td>.13*** (.11)</td>
</tr>
</tbody>
</table>

**Note:** Exponentiated coefficients adjusted for clustering on therapist ID. Standard errors in parentheses. Appt = appointment.

*p < .05, **p < .01, ***p < .001; two-tailed test.
whites were called back only half of the time—one can assume that there are not strong professional norms dictating callbacks. Therefore, the effect of implicit bias is not mitigated by normative pressures to overcome discomfort in order to maintain an egalitarian self-image.

Similarly, deeply rooted stereotypes associated with the working class could lead to negative reactions to their requests for care. For example, although symptoms were held constant in this study, it is possible that therapists imputed different meanings to symptoms based on class stereotypes. For instance, being “unable to get going in the morning” may convey laziness if spoken by a working-class help seeker but not a middle-class help seeker.

A related interpretation of the results is implicit in-group favoritism. Perhaps the disparities in accessibility are driven not by negative bias toward blacks and working-class individuals but by a strong nonconscious preference for in-group members. One formulation of in-group favoritism is homophily, whereby people tend to form new ties with people like themselves (e.g., Kossinets and Watts 2009). Homophily may contribute to psychotherapists’ greater accessibility toward patients like themselves: predominantly white and middle class. Therapists in this study were not more likely to express preference for help seekers of their own gender, but it is possible that strong race and class homophily contributed to the discrimination.

Like most research, the present study cannot distinguish between out-group bias and in-group favoritism. Doing so would require a neutral point of response (Greenwald and Pettigrew 2014). It would also involve extending the experiment to include help seekers of racial-ethnic groups that are culturally different from American whites but that do not invoke the same prejudice or negative stereotypes.

Second, conscious (deliberate) reasoning may contribute to the discriminatory outcomes. Financial considerations are undoubtedly part of psychotherapists’ conscious decision making about accessibility, particularly because they do not have office staff to make those judgments. In this study, measures were taken to minimize potential financial incentives that promote conscious bias against lower-/working-class help seekers, thereby isolating implicit biases. Specifically, all therapists selected were in-network providers for the same private insurance plan that covered all help seekers. Nonetheless, it is possible that mental health care providers view working-class patients as a financial risk—relative to those with the same insurance but higher income—because they cannot be relied upon to keep appointments or pay insurance copayments at the time of treatment. One cannot know if these stereotypes influenced therapists’ decisions. But it is very likely that if insurance coverage had not been held constant in this study, conscious economic inferences based on racial and class stereotypes would have significantly augmented discrimination even beyond that which was observed.

In addition, therapists may explicitly doubt their own cultural competence. That is, the observed discrimination could be a consequence of misguided awareness of the need to consider cultural
differences during treatment. Consistent with this explanation, Teasdale and Hill (2006) note that therapists in training may be concerned about their ability to help patients with different backgrounds from themselves. Perhaps the therapists in the present study were concerned that they would be ill equipped to understand the experiences of black or working-class help seekers. If so, on the therapists’ voice messages for help seekers, one might expect to hear them express doubts about being a good match; this was not the case.

The explanations presented here are not mutually exclusive. It is possible that combinations of factors influence therapists’ decisions about accessibility. It is also worth noting that the race-based and class-based disparities in this study may be driven by different mechanisms. The coefficients in bivariate regressions on race change very little when class is added to the models. Similarly, the coefficients of simple regressions on class are barely moved by the addition of race. This lack of correlation, combined with the finding that the black–white difference exists only among middle-class callers, suggest that different mechanisms are at play.

**Limitations**

One shortcoming of this study is its potentially limited geographic generalizability. Patient selection processes in New York City may differ from locations with different racial or socioeconomic compositions. It may also differ from locations with different levels of provider supply and help-seeker demand for therapy. Other regionally specific variables could also be of import, such as the frequency with which black and working-class individuals seek care.

In addition, the sample of subjects is limited to those with a doctoral degree, so as to more efficiently target solo practitioners. Future research in this area should investigate whether the patterns of accessibility observed for PhD and PsyD therapists can be generalized to populations of therapists with different counseling degrees (e.g., clinical social work) and in different clinical settings.

Furthermore, it is challenging to interpret the intentions of therapists who do not address the help seeker’s request for an appointment in their return message. They may be reserving judgment about the desirability of caller as a patient until they have an opportunity to speak with him or her directly. “Phone tag” is not uncommon in scheduling appointments with psychotherapists, particularly in the absence of office staff. A therapist’s discretion over how long this phone tag continues provides an opportunity to indirectly express favoritism for some prospective patients over others. Moreover, the conversation represents another point in the selection process when disparities in accessibility could expand. Alternatively, disparities in this subset might be smaller because the decision not to specify an appointment time may be an indication of a desire to accommodate the help seeker’s schedule to the extent possible. Regardless of the intent, this process delays the help seeker’s entrance into treatment.

Last, the inability to conduct a true manipulation check is a notable limitation of all audit studies. The use of four scripts introduces variation into the manipulation, which reduces the strength of experimental control, relative to laboratory experiments. Pretesting the recordings of the messages through Mechanical Turk was an important step toward ensuring that the intended race and class were identifiable across all scripts, but it is an imperfect solution. In this study, manipulation check results varied across trials, with race identification consistently above 85%; class identification rates were lower and less reliable. Respondents were least confident about the social class of the black middle-class woman, with correct class identification ranging from 64% to 75%. It may be that perception of the middle-class voices was driven more strongly by race cues (primarily name and pronunciation) than class cues (primarily vocabulary and
grammar). The conflation of perceived race and class presents an obstacle to most sociological research. In this study it is evidenced by the lower rates of class agreement for black middle-class voices and white working-class voices. That is, uncertainty arose where there was discordance between the expectation that blacks are working (or lower) class and the expectation that whites are middle (or upper) class.

Although this is an important issue, it did not have a substantial effect on the findings presented here. The large main effect of social class, in particular, demonstrates the robustness of the results. Indeed, among working-class callers, the study showed equal rates of appointment offers between white and black callers; if perceived race were causing class misidentification by therapists, then one would instead expect to see lower appointment offers for black working-class callers. If anything, the true race differences within the middle class may be slightly smaller than observed, and the class differences among blacks may be slightly larger than observed. Ultimately, the sizeable and statistically significant effects support the conclusion that there is a true disadvantage to black middle-class help seekers and all working-class help seekers, relative to middle-class whites.

CONCLUSION

This study employed a field experimental method to answer a previously unaddressed question: whether psychotherapists—highly educated professionals in an underscrutinized helping profession—engage in race-, class-, or gender-based discrimination. The investigation provided a window into an otherwise private exchange between psychological help seekers and providers during the initial request for care. The results exposed a subtle avenue through which providers discriminate against a vulnerable population (i.e., those in need of mental health care) who already suffer from the disadvantages of being black and working class in American society.

The behavior of providers in a professional setting where discrimination is invisible is an important component in the reproduction of inequality in health care. Individuals who experience discrimination during their help-seeking process may view reaching out to psychotherapists as a fruitless activity or develop negative attitudes toward a class of professionals already regarded with skepticism. For those who do persist in their search for care, every instance of blocked access means additional time and effort spent placing numerous phone calls to identify a psychotherapist willing to respond and accommodate their schedules. This is time and effort that those suffering from mental illness—especially those of low socioeconomic status—do not have to spare.

SUPPLEMENTAL MATERIAL

The online appendix is available at http://jhsb.sagepub.com/supplemental.

FUNDING

This research was funded by the Center for Health and Wellbeing, Princeton University.

NOTES

1. Some scholars suggest that the relationship between socioeconomic status and unmet need for care is less clear after controlling for race (see Roy-Byrne et al. 2009).
2. Over 5,000 psychologists lived in New York City at the time they were licensed.
3. Depression and anxiety were selected because they are the most common mental disorders and the most investigated (Roy-Byrne et al. 2009). Due to the frequency of requests for help with such disorders, they were very unlikely to alarm the therapists in this study. The weekday evening time slot was specified because (1) it is in high demand and (2) it reinforces that all callers are employed.
4. Consistent with Bisgaier and Rhodes (2011), calls were separated by one month. They were placed in early November and early December 2013.
5. The voice used for the black working-class condition focuses on the segmental pronunciation features associated with African American Vernacular (e.g., deletion of the postvocalic [r]), rather than the grammatical features, in order to ensure that the grammatical variation would not produce any confusion for the therapists about the content of the message.
6. For quality assurance, each message was coded by two research assistants, and discrepancies were adjudicated by the primary investigator.
7. The vast majority of callbacks were received within 48 hours.
8. This number of return messages excludes multiple callbacks to the same help seeker by the same therapist.
9. This accounts for the fact that each therapist was exposed to two race conditions.
10. Sample size limitations prevent the examination of interaction effects involving therapist characteristics.

REFERENCES

Impression and Diagnostic Ratings by Therapists.”
Lutfey, Karen E., Kevin W. Eva, Eric Gerstenberger, Carol L. Link, and John B. McKinlay. 2010.


Teasdale, Anthony C., and Clara E. Hill. 2006. “Preferences of Therapists-in-training for Client


**AUTHOR BIOGRAPHY**

Heather Kugelmass is a PhD candidate in the Department of Sociology at Princeton University and a Woodrow Wilson Scholar Fellow. Her areas of interest include social inequality, health care, and experimental methods. Her research focuses on disparities in access to mental health care providers.