

Walking the Talk? What Employers Say Versus What They Do

Devah Pager
Princeton University

Lincoln Quillian
Northwestern University
University of Wisconsin–Madison

This article considers the relationship between employers' attitudes toward hiring ex-offenders and their actual hiring behavior. Using data from an experimental audit study of entry-level jobs matched with a telephone survey of the same employers, the authors compare employers' willingness to hire black and white ex-offenders, as represented both by their self-reports and by their decisions in actual hiring situations. Employers who indicated a greater likelihood of hiring ex-offenders in the survey were no more likely to hire an ex-offender in practice. Furthermore, although the survey results indicated no difference in the likelihood of hiring black versus white ex-offenders, audit results show large differences by race. These comparisons suggest that employer surveys—even those using an experimental design to control for social desirability bias—may be insufficient for drawing conclusions about the actual level of hiring discrimination against stigmatized groups.

In 1930, Richard LaPiere, a Stanford professor, traveled twice across the country by car with a young Chinese student and his wife. The purpose of the trip, unbeknown to his travel companions, was to assess the reactions of hotel and restaurant proprietors to the presence of Chinese customers. During the course of 251 visits to hotels, auto camps, restaurants, and cafes, only once were they refused service. Six months later, LaPiere mailed a survey to each of the proprietors in which one of the questions

asked, "Will you accept members of the Chinese race as guests in your establishment?" More than 90 percent of the respondents indicated unequivocal refusal. The discrepancy between these proprietors' responses to the surveys and their actual behavior is indeed striking: Although nearly none of the proprietors expressed a willingness to accept the patronage of Chinese customers, virtually all of them did so when confronted with the situation (LaPiere 1934). If we were to make generalizations based on either the survey results or the field study alone, we would develop radically different views on the level of racial hostility toward the Chinese at that time in history.

LaPiere's study provides a much needed reality check for researchers who rely on expressed attitudes for insight into the nature and causes of discriminatory behavior. Unfortunately, there have been very few efforts to provide the kind of comparison offered in LaPiere's study. Measures from surveys often are accepted as an adequate proxy for behaviors, with little effort to validate this assumption.

The current article seeks to make headway in this discussion, following up on the insights provided by LaPiere more than 70 years ago. In this discussion, we compare the self-reported

Direct all correspondence to Devah Pager, Department of Sociology, Princeton University, Princeton, NJ 08544 (pager@princeton.edu). Support for this research includes grants from the National Science Foundation (SES-0101236), the National Institute of Justice (2002-IJ-CX-0002), the Joyce Foundation, and the Soros Foundation. The views expressed in this document are those of the authors and do not necessarily represent those of the granting agencies. The authors are grateful to Harry Holzer, Michael Stoll, and Steven Raphael for sharing their survey instrument with us. The authors received helpful comments and suggestions from Bruce Western, Harry Holzer, Jeff Manza, Paul DiMaggio, Susan Fiske, and from the *ASR* editor and anonymous *ASR* reviewers.

attitudes exhibited by a sample of Milwaukee employers with their actual behavior in real-life employment situations. By placing our analysis within the context of research on discrimination in contemporary labor markets we hope to underscore the degree to which method matters in our interrogation of the social world.

RACIAL ATTITUDES, DISCRIMINATION, AND CONTEMPORARY LABOR MARKET INEQUALITY

In the years since LaPiere's study, much has changed about race relations in the contemporary United States. In present times, it would be extremely rare to find respondents willing to state racial objections as candidly as those reported in LaPiere's survey. Indeed, trends in racial attitudes demonstrate steady movement toward the endorsement of equal treatment by race and the repudiation of direct discrimination. According to surveys conducted in the 1940s and 1950s, for example, fewer than half of whites believed that white students should go to school with black students or that black and white job applicants should have an equal chance at getting a job. In contrast, by the 1990s, more than 90 percent of white survey respondents endorsed the principle that white and black students and job applicants should be treated equally by schools and employers (Schuman et al. 2001).

Consistent with these trends, many indicators of social and economic status show that African-Americans have made great strides in approaching parity with whites. Blacks, for example, are now nearly equal to whites in rates of high school completion, and have become increasingly well-represented in occupational sectors previously dominated by whites (Farley 1997; Mare 1995; Wilson 1978). Likewise, in the decade after the Civil Rights Movement, and again during the 1990s, the wage gap between black and white workers was substantially reduced (Couch and Daly 2002; Harrison and Bennett 1995; but see Western and Pettit forthcoming). The rapid social mobility among blacks in the United States provides support for the notion that the progressive trends apparent in measures of racial attitudes reflect a real shift in the opportunities now available to African-Americans. In fact, these positive indi-

cators have led some prominent academics to proclaim the problem of discrimination solved. Economist James Heckman, for example, has asserted that "most of the disparity in earnings between blacks and whites in the labor market of the 1990s is due to the differences in skills they bring to the market, and not to discrimination within the labor market." He went on to refer to labor market discrimination as "the problem of an earlier era" (Heckman 1998:101–102). Indeed, for many observers of contemporary race relations, the barrier of discrimination appears to have withered away, leaving blacks the opportunity to pursue unfettered upward mobility.

And yet, despite the many signs of progress, there remain important forms of social and economic inequality that continue to differentiate the experiences of black and white Americans. According to many indicators, blacks, and black men in particular, continue to lag far behind their white counterparts. Some indicators show black men doing steadily worse. African-Americans, for example, experience roughly double the rate of unemployment experienced by whites, with very little sign of change over time. Likewise, rates of joblessness among young black men have been rising over time (Holzer, Offner, and Sorensen 2005).

As a further troubling indicator, many of these young black men, instead of making their way through school and into jobs, are instead increasingly finding themselves housed in an expanding number of correctional facilities. Approximately 1 in 3 black men will spend some time in prison over his lifetime, as compared with only 1 in 17 white men (Bonczar 2003). Among young black high school dropouts, this figure rises to nearly 60 percent. Rivaling other conventional social institutions—such as military service, employment, and marriage, incarceration has now become a typical event in the life course of young disadvantaged men (Pettit and Western 2004).

How can we explain the discrepancies between these varied measures? On the one hand, the progressive trends in racial attitudes may reflect a genuine openness among white Americans to racial integration and equality. In this case, the continuing difficulties facing segments of the black population may simply reflect the "bumpy road" on an otherwise steady trajectory toward racial parity (Gans 1992).

Further, white racial attitudes are not the only barrier to black mobility. Changes in the economic structure, family composition, and crime policy, among other factors, may each exert an exogenous influence on the black population in ways that inhibit mobility, independent of the racial openness of contemporary institutions. From this perspective, continuing black disadvantage could be explained by a reasonable lag between changing attitudes and outcomes, as well as by the multiple influences that shape patterns of racial inequality.

On the other hand, traditional survey measures of racial attitudes may not accurately reflect the degree to which race continues to shape the opportunities available to African-Americans. Indeed, a great deal of evidence suggests that racial stereotypes remain firmly embedded in the American consciousness, affecting perceptions of and interactions with racial minorities even among respondents who overtly endorse the principle of equal treatment (Devine and Elliot 1995). Substantial levels of discrimination have likewise been detected by experimental field studies, which find consistent evidence of racial bias against black applicants in housing, credit, and employment markets (Bertrand and Mullainathan 2004; Turner, Fix, and Struyk 1991; Yinger 1995). As a further reflection of lived experience, the large majority of blacks continue to perceive discrimination as routine in matters of jobs, income, and housing (Feagin and Sikes 1994).

Given the available information, it is difficult to evaluate the extent to which direct discrimination plays a role in shaping the opportunities available to blacks in contemporary society. Surveys of racial attitudes portray one optimistic picture, whereas indicators of economic and social inequality present more mixed results. It is only through direct comparisons of these differing measures that we can assess how and why they may project such divergent conclusions.

In this article, we focus on the specific issue of employment discrimination. Substantively, we are interested in assessing the degree to which employer preferences or biases influence the opportunities available to stigmatized workers. Methodologically, we seek to assess the degree to which choice of measurement strategy affects our understanding of these processes. In our analysis of survey data and behavioral out-

comes, we engage with LaPiere's central concern about the correspondence between measured attitudes and behaviors.

We begin with a review of the literature on the attitude-behavior relationship since LaPiere's study, focusing specifically on the case of attitudes toward and treatment of stigmatized groups. We then turn to concerns regarding the use of survey measures as proxies for measures of discrimination. Finally, we discuss the results from a matched field experiment and telephone survey that are the basis of our empirical results. Throughout this discussion, we seek to emphasize that investigations of important substantive concerns cannot be separated from a discussion of the methods by which these investigations take place.

ATTITUDES AND BEHAVIORS

UNDERSTANDINGS OF THE ATTITUDE-BEHAVIOR RELATIONSHIP

Attitude questions are frequently asked because they are believed to be illuminating about one or more behaviors of interest. One of the most common uses of attitude research is to assess prejudices, stereotypes, and other measures of social distance with the goal of gaining insight about the nature of discriminatory behavior (National Research Council 2004). Attitude questions have been widely used as tools to assist in understanding the basis of behaviors such as discrimination in employment (Bobo, Johnson, and Suh 2000), residential mobility related to white flight (Farley et al. 1994, Krysan 2002), and the influence of race on voting patterns (Sniderman and Piazza 1993).

Because of the difficulty of gathering data on discrimination in natural settings, many substantive sociological studies of discrimination rely on easier-to-gather survey or interview data in the place of behavioral measures. Some studies focus on attitudinal indicators alone, leaving the connection to behavior implicit. Others ask respondents about past behavior or anticipated behavior in response to hypothetical situations. A wide range of survey scales and specific survey techniques have been developed to measure specific forms of prejudice and discrimination (National Research Council, 2004, chapter 8).

As one important example, a survey technique that has become increasingly popular for

assessing situational discrimination involves use of the vignette question, which elicits reactions from respondents about fairly detailed hypothetical scenarios (Sniderman and Grob 1996). An influential early example of the vignette method was developed by Reynolds Farley and colleagues for a better understanding of the attitudinal sources of racial segregation (Farley, Bianchi, and Colosanto 1979; Farley et al. 1978; Farley et al. 1994). With Farley's approach, respondents are asked to express the level of discomfort they would experience living in hypothetical neighborhoods of varying racial compositions, and to estimate the likelihood that they would move out of such neighborhoods. Farley's innovative technique has become widely used in subsequent research, in part because it is easily combined with experimental survey techniques (discussed in the next section) (Emerson, Yancey, and Chai 2001; Krysan 2002).

A key assumption of vignette studies is that reported hypothetical behavior is an accurate proxy for the behavior that would be observed if the respondent actually encountered the situation. In the case of vignette studies that attempt to illuminate the process of white flight, for example, the assumption is that respondents who say they would feel discomfort or would move is highly related to the behavior of moving out (or not moving in) that would occur if the respondent actually lived in the hypothetical neighborhood. Although a perfect attitude-behavior correspondence is not required, these studies make the assumption that the two are related. An almost complete separation between attitudes and corresponding behaviors would undermine the rationale behind most attitudinal studies.

The expectation of attitude-behavior correspondence results naturally from the view that human action is the product of conscious mental states. Several psychologists, most notably Fishbein and Ajzen (1975), have formally modeled the relationship between these components to describe the formation of attitudes and their subsequent influence on behavior. In their model, feelings or beliefs about an object give rise to positive or negative evaluations of the object. These evaluations then influence behavioral intentions, which ultimately influence behavior (Ajzen 2001; Fishbein and Ajzen 1975). If attitudes can be measured successful-

ly by survey questions or interviews, then these should have at least some power to predict overall patterns of behavior toward the attitude object.

Despite the clarity and intuitive appeal of this model, what is most striking about past investigations of the attitude-behavior relationship is the wide range of correlations reported across different studies. Both Deutscher (1966) and Wicker (1969), for example, review a number of studies that find virtually no relationship between attitudes and behaviors. Schuman and Johnson (1976) also discusses a number of notable studies in which a zero or negative correlation between attitudes and behaviors was found. In their review, however, they conclude that a majority of research on the attitude-behavior relationship finds a moderate positive relationship. With examples for each extreme, their article reports correlations close to zero among attitude-behavior assessments of racial bias and transient economic transactions, while demonstrating correlations exceeding .85 among studies of voting behavior. Most others are shown to fall somewhere in between (Schuman and Johnson 1976).

This literature supports the conclusion that no simple formula can describe the attitude-behavior relationship. Rather, tremendous variation exists in the measurement of attitudes and their associated behaviors, and assumptions about their correspondence should be reviewed with caution.

ATTITUDES AND BEHAVIOR TOWARD STIGMATIZED GROUPS

Despite the appeal of using attitudinal measures as proxies for behavior, particularly for hard-to-measure behaviors such as discrimination, a number of factors complicate the relationship between verbal expressions on surveys and actual behaviors directed toward members of stigmatized groups. First, social surveys have long been plagued by the problem of social desirability, or the phenomenon that respondents seek to give socially appropriate answers to questions, even if this involves distorting the truth (Bradburn 1983). In the contemporary United States, the norms of racial equality are so strong that survey respondents are unlikely to feel comfortable expressing negative opinions about members of other racial groups (Crandall

1994). When asked questions about race or other sensitive issues, respondents may be led by these pressures to shift their opinions subtly (or in some cases not so subtly) in the direction of answers they perceive to be the most socially acceptable. To the extent that real-world discrimination continues, this has the effect of biasing survey results in the direction of politically correct, nonprejudicial responses, and of weakening the relationship between measured attitudes and behavior.

Researchers have adopted creative techniques to minimize the problem of social desirability bias, using experimental survey designs to avoid direct group comparisons. These methods build on the split-ballot survey design, in which randomly chosen subsamples of a survey are primed with one of several variants of a survey question to assess responses to a particular group or condition (Sniderman and Grob 1996).¹ For instance, Schuman and Bobo (1988) used a split-ballot design in which half the sample was asked about objections to a Japanese family moving into their neighborhood, while the other half was asked about objections to a black family moving into their neighborhood. Had each respondent been asked about both a black and a Japanese family on the same survey, they may have biased their responses toward similar evaluations of the two groups, consistent with norms of equal treatment. Through statistical comparisons across the two groups, split-ballot studies are thought to produce valid population-level estimates of the importance of race for the question of interest while reducing concerns about social desirability bias that arise from direct racial comparisons. Experimental survey designs have clear advantages for the measurement of sensitive topics, and their results have indeed shown a greater incidence of prejudice than those from traditional survey designs (Schuman 1995; Schuman and Bobo 1988).

¹ As the name implies, split-ballot designs commonly use two experimental groups. More recent work has taken advantage of computer-assisted telephone interviewing to extend this design to include many more variations in survey questions, including variation across multiple dimensions (see Sniderman and Grob 1996 for a review).

We view social desirability bias, then, as a problem that has received substantial attention in the research literature, with some promising developments. Nevertheless, to our knowledge, no research has provided a behavioral validation of experimental survey results. We have little concrete evidence, therefore, with which to evaluate when and to what degree experimental survey measures are in fact accurate proxies for behavior.

A second problem in using attitudinal measures as proxy assessments for discriminatory behavior concerns the emphasis of this method on consciously held beliefs or feelings. With the use of such measures, subjects are typically prompted for their attitudes in ways that allow for some degree of conscious deliberation. However, a growing literature in psychology has documented the existence and influence of implicit attitudes toward stigmatized groups that may influence judgments and actions without conscious awareness (Devine 1989). The intrapsychic processes that promote discrimination are likely to be more strongly activated in the context of a live interaction than in the abstract context of a survey question (Fiske 2004). Discrimination resulting from these interaction-triggered implicit stereotypes would necessarily remain undetected in survey responses, even those using an experimental design.

Creating a similar problem, some measured forms of discrimination may be perceptible only in the context of direct interaction. Social psychological evidence suggests that whites commonly experience heightened levels of social discomfort in the presence of blacks, at times leading to behaviors that are in effect discriminatory despite (accurately reported) nonprejudicial attitudes (Poskocil 1977). For instance, Word, Zanna, and Cooper (1974) show that white subjects conducting mock interviews with trained black applicants make more speech errors, ask fewer questions, and terminate interviews more quickly than with similar white applicants (see also Dovidio, Kawakami, and Gaertner 2002). Again, because these forms of discomfort are activated only by direct social contact—not by questions about hypothetical situations or prejudicial attitudes—these alternate psychological sources of discriminatory behaviors are unlikely to be captured by questions on survey instruments.

Finally, discriminatory action often is strongly influenced by situational factors, further reducing the extent of attitude–behavior correspondence. Complex decisions about where discrimination may be expressed are influenced by a combination of prevailing social norms as well as context-specific considerations (Merton 1949). In hiring, for instance, employers must balance their need for employees, the applicant pool, and other situational contingencies together with their taste for applicants along several dimensions. These situational factors can sometimes overwhelm the influence of prejudice in discriminatory action, resulting in a low correspondence between the two.

Indeed, it is notable, for example, that LaPiere's (1934) study found higher levels of racial bias apparent in the survey responses than in the field situation. Similar studies by Kutner, Wilkins, and Yarrow (1952) and Saenger and Gilbert (1950)—but focused on discrimination against blacks rather than the Chinese—report similarly counterintuitive results. These findings are especially remarkable in light of the contemporary literature on social desirability bias, which overwhelmingly assumes that survey reports will tend to underestimate the amount of discrimination that will occur. We believe the direction of the discrepancy between self-reports and behavior in these three studies most likely results from the importance of situational factors (Ajzen 1991). In the context of these studies, open discrimination likely would have involved some direct interpersonal confrontation. Unlike the decision not to call someone back for a job interview (a relatively passive form of discrimination), the refusal of service, or other more active forms of discriminatory treatment, can impose significant social costs.² In LaPiere's investigation, for example, the discriminator risked creating an uncomfortable interpersonal situation, possibly resulting in a scene. In certain cases, then, behaviors may be constrained in ways that verbal expressions are

not, again leading to a lower correspondence between the two.

The historical evolution of strong norms against openly racist statements makes it less likely that contemporary field studies would find nearly as high a level of openly expressed prejudice as found in the aforementioned three studies (Schuman et al. 2001). And more recent, if indirect, comparisons of attitudes and behaviors usually have found stronger signs of racial discrimination in behaviors than in self-reports of behavioral intentions (Crosby, Bromley, and Saxe 1980). At the same time, it remains plausible that situational factors could still result in higher levels of stated than actual discrimination in certain cases, depending on the context and the attitudinal instrument.

The complexities involved in the relationship between attitudes and behaviors toward stigmatized groups emphasize the need for careful assessments of our measurement tools. The links between these measured attitudes and observed behaviors require systematic evaluation. Unfortunately, despite the frequent use of verbal expressions to draw conclusions about behaviors, very few studies directly calibrate survey responses with corresponding behavioral assessments.

EXPLICIT STUDIES OF PREJUDICE–DISCRIMINATION CORRESPONDENCE

Whereas the sociological literature on the attitude–behavior relationship is small, the recent literature on the specific attitude–behavior case of prejudice and discrimination (in sociology) is virtually nonexistent.³ Indeed, we turned instead to research in psychology for guidance in these matters. Social psychologist Susan Fiske (2004), in a recent, comprehensive meta-analysis, examine 54 studies containing empirical investigations of prejudice–discrimination correspondence. Consistent with the findings from the attitude–behavior literature more generally, Fiske finds wide variation in the relationship between prejudice and discrimination

² Laboratory studies have found that whites behave more aggressively toward blacks than toward other whites, but only when the consequences to the aggressor are low, such as when acting under conditions of anonymity, or in situations with limited possibility of retaliation or punishment (Crosby et al. 1980).

³ Prejudice refers to negative judgments or opinions about a group (attitudes). Discrimination refers to unfavorable treatment directed toward members of a group (behavior).

across studies, with correlations ranging from $-.38$ to $.69$, with a mean of $.26$. Her results thus support a general association between prejudice and discrimination, albeit at low average levels and with great variability across situations (see also Schutz and Six 1996).

The Fiske (2004) review, primarily featuring the work of psychologists, shows that sociologists have largely abandoned the study of prejudice–discrimination correspondence since the mid-1970s. Of the 10 articles in sociology journals included in the Fiske review, the most recent was published in 1973. This is not because sociologists have stopped using attitudinal measures and survey items to study discrimination against marginalized groups, as demonstrated by reviews such as those of Krysan (1999) and Schuman et al. (2001). Rather, sociologists have done little recent work to validate the assumption that these attitudinal measures are associated with discrimination. Krysan's (1999) review, for instance, notes the issue of attitude–behavior correspondence, but does not cite any recent studies on the topic. Instead, Krysan points to the similar trend directions for racial attitude items and corresponding behavioral indicators from unrelated samples and studies (Krysan 1999:139). Evidence of this sort does support a correspondence of attitude and behavior toward stigmatized groups, but only weakly so because a similar trend direction of indicators over time provides only very general evidence of meaningful correspondence.

In contrast to sociologists, among psychologists, the correspondence of attitudes and behavior toward stigmatized groups continues to be the subject of considerable research. Psychological research of this type has provided several important insights into the correspondence between different types of attitudes and behaviors, pointing to, for example, varying relationships between explicit/conscious attitudes, implicit/unconscious attitudes, and various forms of behavior (Dovidio et al. 2002). From a sociological standpoint, however, these studies have some important limitations, most notably those arising from a reliance on behavioral measures obtained in laboratory settings. For instance, the studies Fiske (2004) reviews use outcome behaviors such as ratings of perceived friendliness in interaction with a mock interviewer, subtle behavioral measures such as the number of blinks and the length of eye

contact, or the results of role-playing situations. These outcomes often are far removed from the actual decisions made in their social contexts—to hire, to rent, or to move, to name a few—that are most relevant to understanding the behavioral processes that produce disadvantage among members of stigmatized groups.

For our purposes, the most relevant studies comparing prejudice and discrimination are those that assess these factors in realistic social settings, focusing on forms of discrimination that produce meaningful social disparities. Unfortunately, the three studies that fit this description each were conducted more than 50 years ago (Kutner et al. 1952; LaPiere 1934; Saenger and Gilbert 1950). We have very few means by which to assess the correspondence between contemporary racial attitudes and the incidence of discrimination.

EMPLOYER ATTITUDES AND HIRING DECISIONS

The current study provides an opportunity to investigate these processes in a contemporary context. Bringing together a unique combination of data, we present a direct comparison of self-reported attitudes and corresponding behavior in the context of a real-world setting with important implications for inequality. The substantive focus of this study is on employers' willingness to hire blacks and/or ex-offenders for an entry-level position in their company. In both cases, the sensitive topics under investigation lead us to question the use of employer reports alone. By calibrating the estimates we received from surveys with behavioral measures from an experimental audit study, we are able to gain insight into the consistency between these two important indicators of group preference.

Measures of attitudes come in many forms, ranging from abstract statements of feelings (e.g., "I don't like members of group X") to more concrete statements of intended action (e.g., "I would not hire members of group X"). The latter, referred to as behavioral intentions, are considered the form of attitude that should most closely correspond to observed behavior, because of their conceptualization in terms of specific measurable action (Fishbein 1967; Fishbein and Ajzen 1975; Schuman and Johnson 1976). Thus a weak relation between behavioral intentions and behavior suggests an even weaker relation between the behavior and more

general attitudinal measures. In the current study, we rely on the behavioral intentions expressed by employers as an indicator of their attitudes about blacks and ex-offenders. Comparing what employers said they would do in a hypothetical hiring situation with what we observed them doing in a real hiring situation forms the basis of our current investigation.

METHODS

In the first stage of the study, employers' responses to job applicants were measured in real employment settings using an experimental audit methodology. Between June and December of 2001, matched pairs of young men (testers) were sent to apply for a total of 350 entry-level job openings in the Milwaukee metropolitan area.⁴ The two white testers (one with a fictional criminal record and one without) applied for one set of randomly selected jobs ($n = 150$), and the two black testers (using profiles identical to those of the white pair) applied for a second set of jobs ($n = 200$).⁵ The preferences of employers were measured based on the number of call-backs to each of the applicants, as registered by four independent voice mail boxes. Additional voice mail boxes were set up for calls to references listed on the testers' resumes.

⁴ Jobs were randomly selected from ads placed in the Sunday classified section of the *Milwaukee Journal Sentinel* and on Jobnet, a state-sponsored Internet job site. Entry-level jobs are defined as those requiring no more than a high school degree and limited work experience. Testers were 23-year-old college students from Milwaukee chosen for their effective styles of self-presentation and for their comparability in terms of physical and interpersonal attributes.

⁵ The tester pair rotated which member presented himself as the ex-offender for each successive week of employment searches, such that each tester served in the criminal record condition for an equal number of cases. By varying which member of the pair presented himself as having a criminal record, unobserved differences within the pairs of applicants were effectively controlled. The criminal record used in all cases was a felony drug conviction. Although the more general term "ex-offenders" is used in reference to this group, it is important to note that the relationship between attitudes and behaviors may differ for individuals convicted of different crimes.

For a more detailed discussion of the research design, see Pager (2003).

The findings of the audit showed large and significant effects of both race and criminal record on employment opportunities. Call-backs were received by 34 percent of whites with no criminal record, 17 percent of whites with criminal records, 14 percent of blacks without criminal records, and 5 percent of blacks with criminal records (Pager 2003). Thus, overall, blacks and ex-offenders were one-half to one-third as likely to be considered by employers, with black ex-offenders suffering the greatest disadvantage.

The second stage of the study provided employers with the opportunity to express their hiring preferences verbally in the context of a telephone survey. Several months after completion of the audit study, each of the 350 employers was called by interviewers from the Michigan State Survey Research Center and asked to participate in a telephone survey about employers' hiring preferences and practices (see Pager [2002] for more detailed discussion of the survey instrument and results). Calls were directed to the person in charge of hiring for each establishment. The final survey sample included 199 respondents, representing a 58 percent response rate (Appendix A).

During the course of this survey, employers were read a vignette describing a job applicant with characteristics designed to match closely the profile of the testers in the audit study. Employers who had been audited by white testers were read a vignette in which the hypothetical applicant was white, and employers who had been audited by black testers were read a vignette in which the applicant was black. In this way, the survey design mirrored the split-ballot procedures used by Sniderman and Piazza (1993) and Schuman and Bobo (1988), avoiding direct racial comparisons within the same survey.

The wording of the vignette was as follows:

Chad is a 23-year-old [black/white] male. He finished high school and has steady work experience in entry-level jobs. He has good references and interacts well with people. About a year ago, Chad was convicted of a drug felony and served 12 months in prison. Chad was released last month and is now looking for a job. How likely would you be to hire Chad for an entry-level opening in your company?

Employers were asked to rate their likelihood of hiring this applicant with the following range of responses: very likely, somewhat likely, somewhat unlikely, and very unlikely.

The vignette presented in the survey was designed to correspond closely to the profile of the testers in the audit study. Chad, the hypothetical applicant, was presented with levels of education, experience, and personal qualifications similar to those on the resumes presented by the testers. The type of crime was identical, although the prison sentence in the vignette (12 months) was shorter than that reported in the audit study (18 months).⁶ Thus the vignette aimed to measure employers' self-reports concerning how they would respond to such an applicant, whereas the audit measured how they actually did respond to an applicant with almost identical characteristics. The parallel scenarios of the vignette and audit should maximize the correspondence between the two measures (Schuman and Johnson 1976).

In the current study, the primary outcome of interest represents the employers' willingness to hire an applicant depending on his race and criminal background. As described earlier, in the survey, employers were asked to report how likely they would be to hire the applicant described in the vignette. In the actual employment situations, by contrast, we measured the number of employers who responded positively to testers after they had submitted their application. In most cases, this simply involved the employer inviting the tester to come in for an interview, although in a few cases, the applicant was offered the job on the spot. As we later discuss, the behavioral indicator should thus provide a highly inclusive measure of "willingness to hire," given that a call-back represents only an initial step in the hiring process.

RESULTS

In the following section, we examine the relationship between the survey results and the audit study. Initially, we compare the *level* of willingness to hire blacks and ex-offenders indicated by the audit results and the survey. We then examine the *association* between the two meas-

ures, considering whether employers who indicated high willingness to hire ex-offenders in the survey called back testers in the criminal record condition at higher rates than those who indicated low willingness to hire in the survey. In each of these comparisons, we seek to assess the degree to which what employers say is accurately reflected in what they do.

Figure 1 presents the key results from both data sources. The first two columns represent the percentage of employers who reported that they would be "very likely" or "somewhat likely" to hire the hypothetical applicant, depending on whether he was presented as white or black. We include the "somewhat likely" group here to correspond to our behavioral measure, which is a call-back rather than an actual hire (see discussion below).

The second two columns represent results from the audit study, illustrating the percentage of call-backs received by each group. In the audit study, call-backs also can be considered a measure of "willingness to hire," given that this represents a first cut in the hiring process. Although a call-back is by no means a guarantee of employment, given that employers typically call back several applicants before selecting their preferred hire, it does indicate a favorable initial review.

The results of the two outcomes, however, are anything but comparable. As Figure 1 shows, employers reported a far greater willingness to hire drug offenders in the survey than was found in the audit. In the survey, more than 60 percent of the employers said they were somewhat or very likely to hire a drug offender irrespective of the applicant's race. In the audit, by contrast, only 17 percent of white and 5 percent of black applicants with drug felonies actually received a call-back.⁷

The disparities apparent in these results are extremely consequential for our understanding of the social world. In the survey data, employ-

⁶ The length of sentence was varied moderately between survey and audit to avoid arousing suspicion.

⁷ Call-back rates include all the employers from the audit study, even those who did not complete the survey. Including only respondents captured in both samples produces even more disparate results. Call-back rates for white and black ex-offenders in the overlapping sample are 14 and 3 percent, respectively, demonstrating an even greater distance from the survey results.

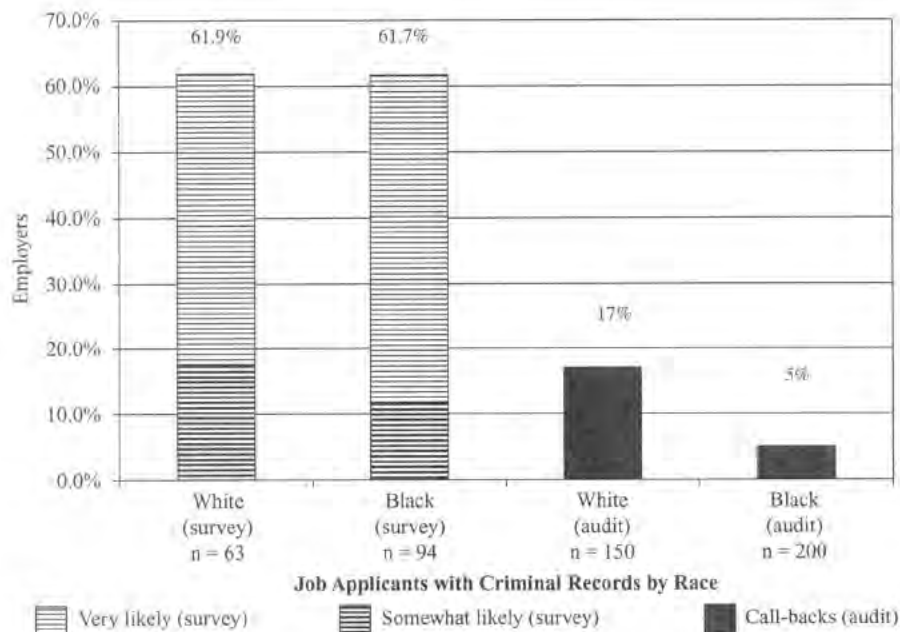


Figure 1. Expressed Willingness to Hire a Drug Offender According to Employer Survey and Audit

Note: Survey results include employers who said they were “very likely” or “somewhat likely” to hire the hypothetical applicant (with “very” at bottom of columns). Audit results represent the percentage of call-backs for each group. Differences between within-race comparisons of survey and audit results are significant on the basis of a two-sample test of proportions ($p < .05$)

ers’ responses present a view of openness to blacks or applicants with drug felonies that is far greater than the reality measured in actual hiring situations. Accepting the survey results as an accurate indicator of the opportunities available to blacks and ex-offenders would grossly understate the barriers to employment they face.

Although the results of this initial comparison are compelling, there remain several possible objections to equating the findings from our survey measure with those from the audit study. First, collapsing the categories of “very” and “somewhat” into one category may artificially exaggerate the distance between survey and audit results. Indeed, if we look only at the “very likely” category, the discrepancy is far less striking. There is a literature on the meaning of vague quantifiers that attempts to offer greater precision to our understanding of these terms (Pace and Friedlander, 1982; Schaeffer, 1991). Lichtenstein and Newman (1967), for example, report that respondents assigned a mean probability of .87 to the phrase “very likely” and a mean probability of .59 to the phrase “somewhat likely.” Whatever the exact probabilities to

which these terms correspond, this literature indicates that such phrases imply a greater likelihood of hire than not. Remember that employers with greater reservations about the applicant also had the option of “somewhat unlikely” to indicate their ambivalence.

Whereas the survey asked employers to rate their likelihood of *hiring* the applicant, the audit merely measured whether the applicant was invited back for an interview or not. Although a call-back may represent a necessary condition to the decision to hire, it is by no means sufficient. In fact, according to the survey results, employers reported interviewing an average of eight applicants for the last noncollege job they had filled. Furthermore, employers on average reported interviewing 55 percent of the applicants that applied (Pager, 2002). Although these self-reported estimates may be inflated, they provide some evidence that the interview stage is far from synonymous with a hire. Rather, a call-back may in fact represent a fairly low bar of approval.

Thus, despite the different metrics on which our measures are based, we believe they provide roughly comparable indications of interest in the

applicant, corresponding to a moderately favorable review. In the results presented later, we provide an analysis of individual-level correlations that should be unaffected by these concerns.

A second possible objection to this comparison is that the very framing of the vignette item may artificially exaggerate the difference between survey and audit results. When considering a hypothetical applicant, employers do not have to take into account alternative possibilities among the applicant pool. Thus the hypothetical applicant may exceed the minimum threshold for acceptability even if in actuality there tend to be other applicants who are better qualified. By contrast, the tester in the audit study is competing with a pool of real applicants of varying quality. To the extent that real applicants provide better qualifications than does the tester's profile, the tester will receive few call-backs for reasons unrelated to race or criminal record.

An alternative way of presenting the information that addresses this concern is to calculate the likelihood that a tester with a criminal record will receive a call-back *relative to* a white tester without a criminal record. White testers without criminal records in this case represent a kind of baseline, presenting a given set of qualifications common among all testers, but without the handicaps of minority status or a criminal record. Employers who made call-backs to white testers without criminal records signaled that this level of education and experience was sufficiently desirable to make the first cut. Relative to this baseline, we can assess the proportion of blacks and whites with criminal records who received call-backs, thereby reducing the effect of employer nonresponses attributable to extraneous factors.⁸

⁸ Similarly, we can consider the proportion of employers who reported that they were likely to hire an ex-offender (61.9 or 61.7 percent) as relative to an implicit baseline of 100 percent for a hypothetical applicant similar to the one described in the vignette, but without a criminal record. To the extent that some employers would report it not likely that they would hire such an applicant (if, for example, they hire only applicants with college experience), the ratio of the self-reported likelihoods of hiring an ex-offender relative to a nonoffender would be even larger, thus

Figure 2 displays the results of this procedure, comparing the likelihood of hire based on the survey and audit results with audit results recalculated as a ratio of the percentage of testers in the offender condition who received call-backs to the percentage of white testers with identical qualifications but no criminal background who received call-backs. Overall, 34 percent of white applicants with no criminal records, and with the given set of human capital characteristics presented by all testers, received call-backs. This group serves as our baseline (denominator) for calculating the relative call-back rates for the other groups. Only 17 percent of white testers with identical characteristics plus a criminal record received call-backs, indicating that white testers with a criminal record were 50 percent as likely to receive call-backs as those without a criminal record (Figure 2). Black ex-offenders were the least likely to continue in the employment process—only 5 percent received call-backs—indicating that they were just less than 15 percent as likely to receive a call-back as a similar white tester without a criminal record.

The differences between self-reports and behaviors in this comparison, although smaller, remain consistent when call-back frequency is judged relative to that for white non-offenders. In the case of white ex-offenders, the distance between the survey and audit results has narrowed substantially, although it remains marginally significant statistically. The case for black applicants, on the other hand, maintains a clear and dramatic difference. Even relative to contemporaneous call-back rates for white testers, the call-back rate for black ex-offenders (14.7) remains far short of the survey estimates of hiring likelihoods (61.7). For black ex-offenders, the survey and audit measures provide dramatically different indications of willingness to hire.

Whatever measure is used, two main findings remain clear: First, whereas the survey responses present a rather benign view of the employment barriers facing ex-offenders, the audit results tell a very different story. Employers indicate a high level of willingness to hire drug offenders, but in actual employment situations,

rendering the contrast between the self-reports and behavioral outcomes even greater.

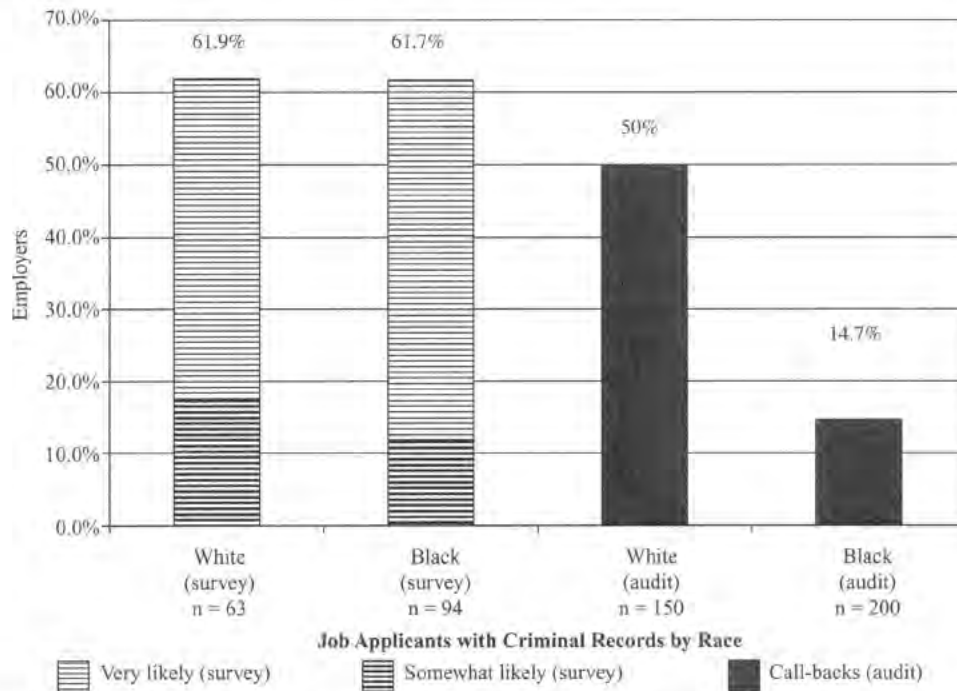


Figure 2. Expressed Willingness to Hire a Drug Offender According to Employer Survey and Recalibrated Audit

Note: Survey results include employers who said they were “very likely” or “somewhat likely” to hire the hypothetical applicant (with “very” at bottom of columns). Audit results represent the ratio of the percentage of call-backs for each group to the percentage of call-backs for white nonoffenders. Differences in within-race comparisons of survey and audit results are marginally significant for white applicants ($p < .06$) and significant for black applicants ($p < .05$) on the basis of a two-sample test of proportions.

they are less than half as likely even to call back such applicants relative to those without criminal records. This result underscores the importance of using great caution in relying on employers’ self-reports as an accurate reflection of behavior.

Second, the degree to which race is a factor in hiring decisions is virtually undetectable in the survey results, in sharp contrast to what we find in the audit study. Table 1 shows the relative risk of receiving a call-back for white and black applicants in the survey and audit. In the survey, although separate employers were asked the vignette in which the hypothetical applicant was white or black, the estimates of hiring likelihoods for both applicants were virtually identical.⁹ By contrast, actual behavioral meas-

ures in the audit show that white ex-offenders are more than three times as likely to receive consideration from employers as black ex-offenders.¹⁰ These results suggest that employer surveys, even those with split-ballot designs, do not always provide an effective way to gauge the degree to which sensitive characteristics such as race affect actual employment opportunities. Later, we discuss the methodological and theoretical processes that might account for these differences.

Finally, we turn to the issue of individual-level consistency between survey reports and audit results. Even if the *levels* of openness to hiring ex-offenders are inconsistent between survey and audit, it remains possible that a *correlation* exists between the two: Employers who indicate

⁹ Even if we were to restrict our attention to the “very likely” category, the black–white ratio (1.5) still would vastly understate the degree of racial disparity apparent in the audit.

¹⁰ This result is calculated as a ratio of the call-back rate for white drug offenders (25/150) relative to the call-back rate for black drug offenders (10/200).

WHAT EMPLOYERS SAY VERSUS WHAT THEY DO 367

Table 1. Comparison of the Influence of Race on Hiring: Audit versus Survey

Relative Willingness to Hire	Survey	Audit
White vs. Black (White/Black)	1.00	3.33
95% CI	(.78, 1.29)	(1.65, 6.73)

Note: The difference in white/black ratios between the survey and audit is statistically significant at $p < .001$ (Mantel-Haenszel test, $\chi^2 = 12.01$, 1 df).

Table 2. Individual-level Consistency between Employers' Self-reports and Behavioral Outcomes

Survey Results	Audit Results (for Testers Presenting Drug Felony)	
	No Call-Back	Call-Back
Likely to Hire Drug Offender		
No	56 (93.3 %)	4 (6.7 %)
Yes	89 (92.7 %)	7 (7.3 %)
Difference of Percentages (95% CI)		0.6% (-8.6%, 8.8%)
Correlation (Kendall's Tau-b) (95% CI)		.012 (-.143, .167)

Note: This table includes all employers who responded to the survey. Call-backs in the right column above represent calls to the tester in the criminal record condition only.

willingness to hire ex-offenders may be more likely to hire an ex-offender than employers who do not indicate such willingness, even if the overall openness to hiring ex-offenders is strongly overstated in the survey results. This final analysis allows us to compare the survey responses with the audit outcomes at an individual rather than an aggregate level. The results of this cross-tabulation are presented in Table 2.¹¹ Consistent with the results reported earlier, we find that the survey responses have very little connection to the actual behaviors exhibited by these employers.

¹¹ The percentage of call-backs is lower than for the aggregate comparisons because of survey non-response. A full breakdown of survey responses (including all four survey response categories) by audit results is presented in Appendix B. Given the small sample sizes in this final comparison, a further breakdown by race of the tester would be impossible. Analyses, therefore, include all call-backs to testers in the criminal record condition regardless of race.

Among those who reported a favorable likelihood of hiring an applicant with a prior drug conviction in the survey, 7.3 percent made calls to the tester with the criminal record in the audit study, relative to 6.7 percent of those expressing an unfavorable likelihood. This difference is in the expected direction, but is only slightly greater than zero (0.6 percent), and far too small to reach statistical significance. Likewise, correlational measures for ordinal data, such as Kendall's Tau-b, show nearly zero association between survey and audit.¹² Considering the possibility that our relatively small sample size

¹² The individual-level comparison further allows us to reconsider the practice of combining "very" and "somewhat" likely responses into a single category. To investigate this question, we recalculated Kendall's Tau-b between the audit call-backs and the vignette question using the four survey answers, with separate "very" and "somewhat" categories. Instead of producing stronger results, this coefficient is not statistically significant, and the negative sign is opposite that expected (Tau-b = -.0391).

limits the reliability of these estimates, we calculate confidence intervals allowing us to assess the potential relationship that would obtain if we had used a larger sample. A 95 percent confidence interval for the difference in percentages includes a range from -8.6 to 8.8 percent, indicating that we are 95 percent confident that employers who indicate "yes" are no more than 8.8 percentage points more likely to make a call-back than employers who indicate "no."¹³ A difference of 8.8 percentage points for making a call-back still is a fairly low level of correlational consistency.¹⁴ We can thus be fairly confident that, given this pattern of results, even a much larger sample would be unlikely to produce a substantial relationship between survey and audit results.

These results cast strong doubt on the accuracy of survey data for indicating relative likelihoods of hiring. Individuals who report a higher likelihood of hiring an ex-offender are only trivially more likely to do so. Confirming the aggregate findings described earlier, the individual-level associations presented here appear to be no better at establishing a relationship between attitudes and behaviors.

Nevertheless, several limitations of this analysis must temper its conclusions. In the following discussion, we consider possible threats to the validity of our findings caused by measurement error or study design. In the first case, we consider the possibility for error in the survey or audit results, either of which could lead to a weakened correlation between the two.

¹³ This confidence interval is calculated using the "plus 4" method of Agresti and Caffo (2000). The Agresti and Caffo method has the advantage of providing accurate (and slightly conservative) intervals even when the count of successes or failures is very small. By contrast, the methods in standard introductory statistics books usually require at least five successes and failures for each group.

¹⁴ Similarly, the upper limit of the 95 percent confidence interval for Kendall's Tau-b suggests that the degree of correlational consistency in a large sample probably would be quite low (below .167). The confidence interval for Kendall's Tau-b is asymptotic, and thus should be regarded with more caution than the confidence interval for the proportion, which is computed by a method with good small-sample properties.

In the case of the survey, the most plausible source of measurement error arises from those cases in which the survey respondent was different from the individual who reviewed the testers' applications. To the extent that hiring practices vary within firms depending on the individual manager or human resource officer, the consistency between survey and audit results will be attenuated. Although recent evidence suggests that labor market discrimination typically operates at the level of the firm rather than the level of individual discriminatory actors,¹⁵ this possibility remains a potential source of measurement error. Nevertheless, although within-firm heterogeneity may indeed affect the individual-level consistency in measures of attitudes and behavior, there should be no effect on the average level of support for hiring ex-offenders in the aggregate (as presented in Figures 1 and 2). In cases wherein respondents differ, there is little reason to believe that the hiring agent would be systematically more or less likely to consider hiring ex-offenders than would the survey respondent.¹⁶

In the audit study, measurement error also may pose a problem. Because each employer was visited only once, we have only one data point with which to assess their hiring tendencies. Given the many factors at play in any given hiring situation (e.g., the urgency with which the position must be filled, the number and quali-

¹⁵ Recent studies have found that firm-level variables such as the presence of a human resource apparatus, the use of applicant tests, and affirmative action policies have far more influence on the racial composition of a company than the individual characteristics of hiring managers or owners (Holzer 1996; Holzer and Neumark 2000).

¹⁶ One might also question whether possible changes in the economic climate at the time of each measurement may be responsible for some of the disparity. The unemployment rate in Milwaukee averaged 4.8 percent during the time of the audit, whereas in the 2 months during which the survey was administered, it had risen to nearly 6 percent (Bureau of Labor Statistics 2002). Given that employers' openness to less desirable workers increases in the context of tight labor markets (Freeman and Rodgers 1999), we would rather expect more favorable responses from employers during the period of the audit study relative to the period of survey data collection.

ty of other applicants), any single data point representing an employer's treatment of an ex-offender may be subject to the measurement error of circumstance.¹⁷ Nevertheless, the almost complete absence of association between survey and audit measures leads us to question random error as a sufficient explanation. If the hiring process tends to be so complex as to defy any straightforward relationship between the abstract intentions of employers and their ultimate decisions, this would imply more than a simple problem of measurement. Indeed, we later discuss the complexity of hiring decisions as one of our hypothesized explanations for the discrepancy between outcomes.

More systematic forms of error can creep into an audit design if the experimental procedures are somehow compromised. Of primary concern are the potential biases that can be caused by the performance of testers, either because of poor matching or because of testers' self-fulfilling expectations (Heckman and Seligman 1993). Fortunately, a rather direct test of this hypothesis is possible. One would expect that if differences in testers' personalities or behavior shaped the outcomes (above and beyond any effects of race or criminal condition), we should see stronger results among those cases in which the testers had the opportunity to interact with employers. Applications submitted with little or no personal contact should be far less affected by such concerns. Analyzing the outcomes for these two kinds of tests, we find no evidence of tester bias. In fact, the main effects of both race and criminal background are substantially attenuated among those who did have personal contact with the employer (Pager 2003). This suggests that instead of exaggerating negative stereotypes, the appealing characteristics of these testers actually worked to reduce the measured effects, thus biasing the results in a direction consistent with the survey responses. It is unlikely then that these various sources of error can account for the significant disconnect between the survey and audit results.

¹⁷ As in the case of the survey, measurement error in the audit results will attenuate individual-level associations, as presented in Table 2, but should not affect the aggregate comparisons presented in Figures 1 and 2.

A final limitation of the comparison provided earlier is its reliance on a single survey item. Flaws in the survey design or peculiarities of the question wording could lead to anomalous patterns of responses. Fortunately, an additional item was included in the survey that allows for similar comparisons to be drawn. Whereas the vignette most strongly paralleled the audit situation—including a match of the applicant's race—the second item also probed employers' willingness to consider hiring ex-offenders, in this case focusing on a generic applicant with a criminal record. The exact wording of this item was "Next, I am going to list several types of applicants. Please tell me if you would accept each type for the [most recent noncollege] position . . . an applicant who has a criminal record?" As described earlier, we compare those who answered that they "definitely" or "probably" will hire with those who answered "definitely" or "probably" not in relation to the audit outcomes.¹⁸ The correlation between this measure and the audit results is again very small. The correlation coefficient from Kendall's Tau-b is .0003 (95 percent confidence interval, $-.154$ to $.155$). Again, our tests for significance in this case cannot reject the hypothesis of no relationship.

The various aforementioned limitations must certainly temper our conclusions. Nevertheless, the almost total lack of correlation between the survey and audit results is troubling. If these findings are an accurate assessment of the level of consistency for these and related measures, then studies that use similar survey items to draw conclusions about characteristics or circumstances associated with discrimination may come to strongly misleading conclusions.

On the basis of several methods for assessing the attitude-behavior relationship, all comparisons tell a similar story: it is difficult to get

¹⁸ In the original survey, 25 percent of the respondents gave the response "it depends." As a conservative estimate, these respondents were treated as "willing to hire," producing a stronger correlation than when they are excluded from the analysis. In fact, the association of survey and audit becomes slightly negative when this category is excluded. (By contrast, the correlation coefficient from Kendall's Tau-b for this survey item and the original vignette item was .55).

an accurate picture of actual hiring outcomes based on responses to the employer survey used in this study. Employers generally express a greater likelihood of hiring applicants with criminal records, and a far greater likelihood in the case of black applicants, than we see in actuality. Furthermore, employers who indicate greater willingness to hire an ex-offender in response to a survey question seem to be only slightly more likely actually to offer an interview to such an applicant. Both in terms of making aggregate- and individual-level predictions, our evidence points to weak correspondence between survey results and actual hiring outcomes.

ATTITUDES AND BEHAVIORS: WHY DO THEY DIFFER?

Why might employers' survey responses present results so discrepant from their actual behavior? Several theoretical explanations could be used to account for this incongruity. In this section, we provide a discussion of these explanations, considering the range of underlying processes that may give rise to differing outcomes.

SOCIAL DESIRABILITY AND COMPENSATORY ESTIMATION

As discussed earlier, efforts to measure attitudes about sensitive topics are complicated by the problems of social desirability bias. According to this perspective, respondents may conceal their true feelings about blacks or ex-offenders in answering survey questions. If this is the case, the discrepancy between self-reports and behaviors should be viewed as the difference between false and true measures of a respondent's attitudes. Although social desirability pressures certainly result in some distortion of survey results, we do not believe that this can account fully for the differences between expressed willingness to hire ex-offenders and the actual employment outcomes based on applicants' criminal record. In fact, at other points in the survey, respondents expressed strong opposition to considering applicants with criminal records other than drug felonies: nearly 70 percent of employers expressed an unwillingness to hire an applicant who had been convicted of a property crime, and more than 75

percent were self-described as unwilling to hire an applicant who had been convicted of a violent crime (Pager, 2002; see also Holzer, Raphael, and Stoll 2003).¹⁹ It therefore seems unlikely that social (or legal) pressures to accept ex-offenders whitewashed employer responses. High levels of acceptance were reported only for the applicant described as a drug felon.

Social desirability bias may be a greater concern in the measurement of racial preferences from the survey results, which is where we find the largest disparities between expressed attitudes and observed behaviors. To preempt this concern in the current study, we used a split-ballot format in which each employer responded to only one hypothetical (black or white) candidate. It remains possible, however, that social desirability bias is a problem if, even in the absence of direct comparisons by race, employers are aware that the race of the hypothetical applicant has been specified and therefore make conscious or unconscious efforts to compensate verbally for any negative reactions they may have to a black applicant. If respondents do in fact suppress negative reactions to race-specific targets, even when no racial comparison is provided, this calls into question the effectiveness of split-ballot survey designs as a strategy for measuring underlying racial prejudice. Any self-reported attitude toward a black target may in fact be distorted by the respondent's own compensatory estimation procedure.

ABSTRACT VERSUS SPECIFIC CONSIDERATIONS

A second possible explanation for the discrepancy between measured attitudes and behaviors in this study relates to differences in the criteria used to evaluate a hypothetical versus an actual job candidate. It is plausible that the affirmative responses of employers considering the acceptability of a hypothetical applicant

¹⁹ Questions about specific crime types followed the presentation of the initial vignette described earlier. Employers were asked to report the likelihood of hiring Chad if, instead of a drug crime, he had been convicted of a property crime, such as burglary, or of a violent crime, such as assault. Response categories for property/violent offenders were "very likely" (10/7 percent), "somewhat likely" (21/17 percent), "somewhat unlikely" (32/29 percent), and "very unlikely" (37/47 percent).

indicate their genuine willingness to consider hiring an applicant with a criminal record *in the abstract*. In these general terms, apart from the minority of employers who categorically reject all applicants with criminal records, a prior conviction is not typically grounds for immediate disqualification. Rather, if the applicant's overall characteristics exceed a minimum threshold of employability, the respondent is likely to indicate a willingness to hire.

By contrast, in actual employment situations, the applicant's characteristics are judged not only according to some minimum threshold, but also relative to the pool of available applicants, and to the specific requirements of a job. In this case, many more contingencies are at play, and the presence of a criminal record may become a salient criterion by which to weed out less-qualified applicants.²⁰ Even if the employer genuinely believes that she or he would hire the applicant described in the abstract vignette, when confronted with the situation in real life, the contingencies of the hiring process may render hypothetical scenarios irrelevant.

Recognizing this potential disconnect, we made efforts to calibrate the behavioral responses to a concrete indicator of employability at that place and time. In this case, the white nonoffender, our baseline in Figure 2, serves this function by providing assurance that this level of qualification was sufficient to elicit a call-back during that particular hiring process. Despite this adjustment, the willingness to hire expressed in the survey appears to be much higher than in the audit.

Even if differences between the exact vignette and the audit situations can explain some of the discrepancies between survey and audit measures of overall willingness to hire, this

explanation cannot account for the considerable difference in race effects detected by these comparisons. In the survey, black and white applicants appear equally likely to receive offers, whereas in the audit, there is a large gap in favor of white applicants. An investigator using these survey data alone would be strongly misled about the role of race in shaping actual hiring decisions.

THE INTENSITY OF PRIMING

A third perspective on the discrepancy between self-reports and behaviors proposes that the priming of characteristics during a phone interview may not elicit the same intensity of response as the in-person presentation of the same characteristics. Hearing a description of a hypothetical black ex-offender is quite different from seeing a young black man approach one's business in search of employment. The live interaction may trigger feelings of fear, anxiety, or threat in ways that a recited vignette does not (Poskocil 1977). These feelings may then influence employment decisions in ways that cannot be fully replicated in hypothetical scenarios in surveys.

Similarly, social psychological evidence suggests that racial stereotypes exert many of their effects indirectly, by coloring the evaluation of ambiguous information (Darley and Gross 1983). When employers are evaluating applicants, for example, an energetic, outgoing, young white applicant may be viewed as motivated and eager to work, whereas a similarly energetic, outgoing, young black male may be seen as a hustler or a "player." Even relatively straightforward cues can be interpreted in vastly different ways, depending on the context of the situation or the characteristics of the actor (Sagar and Schofield 1980). Again, these sorts of distortions would most likely operate during in-person evaluations. The vignette, by contrast, leaves less room for distorted interpretations, because according to the explicit description, the hypothetical candidate "has good references and interacts well with people." This class of explanations suggests that discriminatory behavior in the employment of ex-offenders, especially African-Americans, may have a basis in sources other than consciously believed attitudes toward these groups.

²⁰ Indeed, in response to the second more general survey question discussed earlier, when employers were asked about their willingness to hire a generic applicant with a criminal record (with no additional information provided), a large fraction of employers refused the forced-choice response categories, insisting instead that "it depends"—on the crime, on the length of time since the conviction, on the type of job, and numerous other considerations. In this case, then, the employers themselves acknowledged that any estimation given in the abstract would have very little bearing on how they might respond in making an actual hiring decision.

In the article discussed earlier, LaPiere (1934) reinforces the view that surveys may elicit a different set of considerations than do concrete experiences. According to LaPiere, survey responses constitute “verbalized reaction[s] to a symbolic situation,” or reactions to a highly abstracted representation of reality (p. 231). According to this viewpoint, survey responses do tell us something meaningful about the attitudes of respondents, but we have no way of anticipating the degree to which these expressed attitudes will be reflected in any particular set of behaviors. Certainly, it is difficult to anticipate how any individual, including oneself, may react to a situation previously encountered only in hypothetical terms. In the case of hiring decisions about individuals with stigmatizing characteristics, our results suggest that very little can be implied from these self-reports of employers for the accurate prediction of employment outcomes.

It is not possible using the current data to demonstrate conclusively which underlying process may have generated the observed discrepancies. In fact, it is highly plausible that more than one process may have been at work simultaneously. What these results do demonstrate, however, is that employers’ expressed willingness in the survey taps into a set of processes very different from those measured through our behavioral study. Although these processes may be related to a common underlying disposition, the correspondence between the two can be quite weak. It is important that researchers recognize these limitations before drawing inferences about behavior from the self-reports of survey respondents.

RETHINKING THE ROLE OF ATTITUDES

What can we conclude from these results regarding the usefulness of data on attitudes? Should we disregard all employers’ self-reports? Certainly, it would be premature to advise such a radical stance. In fact, despite the large discrepancies between the self-reports and behaviors measured in the current study, we believe that survey results remain useful, even if they cannot be viewed as an alternative procedure to the measurement of actual discrimination.

Even in cases in which expressed attitudes have little relationship to measured discrimination, survey data can nevertheless tell us

something useful about how employers *think* about important hiring issues. Responses to the survey suggest, for instance, that many employers who discriminate against blacks do not necessarily do so because of a principled belief that black employees should not be hired. In fact, we think it likely that many employers genuinely believe their own responses to surveys, professing the value of equal opportunity, while simultaneously justifying their behavior in hiring situations on grounds other than race (e.g., assumptions about the family/social/educational backgrounds of black applicants; see Kirschenman and Neckerman 1991). In this case, the difference between employers’ self-reports and their actual behavior represents a meaningful discrepancy between two legitimate realities. The resolution of these differences represents an important focus of sociological investigation in its own right. Although low correlations between attitudes and associated behaviors often are viewed as a purely methodological test of survey questions, in many cases, these discrepancies actually may provide clues toward a better substantive understanding of the cognitive-emotional basis for action.

Furthermore, it remains possible that survey research may provide a better proxy for behavior in situations that are less complex and subject to fewer contextual influences than hiring. Action in any real social situation is the result of many factors other than the actor’s attitude toward the object, including norms, perceived consequences of the action, and implicit or unconscious attitudes toward the object. The many complex influences on hiring decisions make these situations exactly the sort for which survey measures are least likely to be an effective substitute. Indeed, the three “classic” studies that found very weak associations between expressed behavioral intentions and behaviors all were studies of discrimination in social situations (Kutner et al. 1952; LaPiere 1934; Saenger and Gilbert 1950). We believe it possible that survey responses may provide a much more effective proxy for behavior in other contexts, such as those that involve voting (Traugott and Katosh 1979), signing of a petition (Brannon et al. 1973), or patterns of consumer behavior (Day et al. 1991), in which the link between behavioral intentions and actual behavior is less subject to contextual influences apart from the respondent’s attitude or intention.

Finally, we have focused on only a few of the many survey techniques that have been developed to measure prejudice and discrimination. Though our measure of behavioral intentions was designed to offer the closest match to the audit context, it remains possible that other more abstracted measures of racial bias may in fact correlate more closely with measures of discrimination. There is an extensive literature that attempts to investigate modern or subtle forms of racial attitudes using survey questions (National Research Council 2004, chapter 8), and certain of these alternative approaches could prove more effective at capturing behavioral outcomes than what we found in this study.

Three sociological approaches that we believe to be especially promising place their respective emphasis on stereotype measurement, past behavior, and in-depth interviews. The first of these, group stereotype measurement, has a long history in the social sciences, with research demonstrating a persistence of racialized attributions across numerous dimensions (Devine and Elliot 1995). As a recent example, researchers have developed a series of scales measuring respondents' images of different racial groups along a wide range of social and psychological characteristics (Bobo and Kluegel 1997; Smith 1991). Survey techniques such as these have shown respondents to rate blacks as worse or inferior relative to other groups on dimensions such as criminality and intelligence, suggesting that traditional measures of racial attitudes may be missing a great deal of underlying racial bias.

A second approach, used extensively by Harry Holzer and colleagues, asks employers to focus on the last worker hired, thereby grounding responses in a concrete recent experience (Holzer 1996; Holzer, Raphael, and Stoll 2003). By focusing on a completed action, Holzer is able to avoid the ambiguities of hypotheticals or general statements; and by focusing on actual outcomes, he is able to document "revealed preferences" rather than expressed ones. Likewise, Holzer's format calls for race to be assessed only as one incidental characteristic in a larger series of questions concerning the recent employee, thereby reducing the social desirability bias often triggered when the subject of race is highlighted. Whereas recall or motivational biases emerge as concerns in the reporting of prior experiences (Bradburn, Rips, and

Shevell 1987), this particular approach focuses on a well-defined event that leaves less room for error-prone estimation.

Finally, some data suggest that in-depth in-person interviews may be more effective in eliciting candid discussions about sensitive hiring issues than other modes of interviews (Kirschenman and Neckerman 1991; Moss and Tilly 2003; Wilson 1996). In-depth interviews offer the opportunity for respondents to discuss the complexities and, at times, the inconsistencies in their views of various groups, thereby going beyond the more generalized assessments expressed in traditional survey items. Likewise, the opportunity for rapport-building in the in-person interview context may reduce social desirability pressures, making respondents feel at greater ease in expressing counter-normative beliefs.²¹

Although each of these represents a promising approach, our results caution against the unreflective assumption that the results of any method are necessarily good proxies for behavior. None of these techniques has yet been subjected to direct assessments by comparison of their responses with corresponding behavioral measures, a step we view as necessary for understanding the relation of these measures to behavioral outcomes. LaPiere's (1934) warning, that hypothetical scenarios often cannot convey the experience of concrete situations, deserves to be taken more seriously by current generations of survey and interview researchers.

CONCLUSION

LaPiere (1934) showed a striking inconsistency in the way hotel and restaurant proprietors reacted to Chinese customers in person, as compared with how they responded in surveys. The current study notes a similar discrepancy between employers' self-reported likelihood of hiring a particular applicant and their actual hiring behaviors when faced with a nearly identical candidate. We found an especially large and robust disparity between the reported likelihood of employers hiring black ex-offenders and actual rates of hiring. The low correlation between expressed and observed hiring out-

²¹ For a critique, see National Research Council (2004), p. 175.

comes presents an epistemological worry: our assessments of the degree of disadvantage faced by black ex-offenders would be substantially underestimated on the basis of the survey results alone. Moreover, we found little correlation between greater expressed likelihood of hiring ex-offenders in the survey and actual increased rates of call-backs for ex-offenders in real hiring situations. Given that most research on hiring preferences and practices comes from the self-reports of employers themselves (Downing 1984; Holzer 1996; Husley 1990; Jensen and Giegold 1976; Waldinger and Lichter 2003; Wilson 1996), these results indeed have serious implications.

In terms of the methods used to measure discrimination, these findings suggest that sociologists may need to reevaluate what is learned from studies that use vignettes of hypothetical situations to study behaviors toward stigmatized groups. Although we believe that these vignette studies often do tell us about respondents' abstract beliefs, in some cases these beliefs may have relatively little influence on the behavior of interest. Feelings and evaluations in a concrete social situation may be very different from those in the abstracted situation of the survey, but the two often are treated as nearly identical. An important next step in evaluating the contribution of survey measures for understanding behaviors of interest is to relate these items to actual behavior.

More broadly, these results suggest the limits of survey questions alone for understanding the changing nature of racial inequality. Survey questions indicating a liberalizing of racial attitudes among white Americans have been cited widely as evidence supporting the declining significance of race in American society. But if the items analyzed in this study have any bearing on survey responses more generally, we have reason to question that changing public opinion on matters of race has any necessary correspondence to the incidence of discrimination. Rather, our results support the perspective that there has been a growing gap between the principled statements and beliefs of white Americans in favor of racial equality and their concrete actions.²² Survey questions provide

one important perspective on American race relations, but they must be combined with other information for a complete picture.

Fortunately, methods to improve our understanding of the prejudice–discrimination relationship are readily available and feasible for even small groups of researchers. The comparison of survey measures and behavioral indicators does not require an unprecedented level of resources. Even for pedagogical purposes, sociology teachers could readily incorporate the dual design within a two-semester timetable. Whereas audit studies of labor markets can be quite involved, numerous other everyday social settings provide countless opportunities for implementing small-scale experiments: searching for an apartment, shopping, hailing a taxi, passing security checkpoints, and the like (for an example of a classroom application of the audit methodology, see Massey and Lundy 2001). Moreover, the investigation of prejudice and discrimination could be usefully applied to many other groups: Asians, Latinos, Muslims, women with criminal records, gays and lesbians, to name just a few.

For creative sociology teachers, then, a single class could readily achieve a paired audit study and telephone survey with sufficient sample sizes for meaningful comparisons. Both substantively and methodologically, the pairing of survey and audit research can yield important insights for the study of contemporary discrimination.

It is not the case that employers' thoughts and beliefs can tell us nothing about important employment issues. In fact, in many cases, surveys and other methods of eliciting employer opinions can provide useful information about attitudes and beliefs. In other cases, surveys may provide a very close reflection of actual employer behaviors. What this research emphasizes, however, is the importance of testing one's assumptions and providing external validation of key results. In the case of employers' behavior with respect to hiring black ex-offenders, the survey results presented here are far off base. The correspondence between self-reports and

²² This perspective has also found support from certain well-designed surveys that manage to capture

respondents' contradictory or competing beliefs (Bobo, Kluegel, and Smith 1997; Kinder and Sears 1981).

behaviors with respect to other important hiring outcomes has yet to be established.

Devah Pager is an Assistant Professor of Sociology at Princeton University. Her research focuses on institutions affecting racial stratification, including education, labor markets, and the criminal justice system. Continuing her work using experimental field methods, she currently is investigating discrimination in low-wage labor markets in New York City (with Bruce Western). Her recent publications include "The Mark of a Criminal Record," published in the American Journal of Sociology (March 2003), and "The Structure of Disadvantage: Individual and Occupational Determinants of the Black-White Wage Gap," published in the American Sociological Review (August 2001) (coauthored with Eric Grodsky).

Lincoln Quillian is Associate Professor of Sociology at Northwestern University and the University of Wisconsin-Madison. His research focuses on inequality, racial attitudes, and urban population distribution. His current research projects include studies investigating the role of stereotypically biased evaluations in the creation of racial inequality and the consequences of spatial income segregation for social stratification.

APPENDIX A

SURVEY DESIGN AND IMPLEMENTATION

The baseline survey instrument was developed by Harry Holzer and his colleagues.²³ It includes questions about the company such as size, industry, employee turnover, and racial composition; questions about hiring procedures such as the use of interviews, personality or aptitude tests, and background checks; questions about the last worker hired for a position not requiring a college degree including age, race, and sex of the worker, recruitment method, wage, and promotion opportunities; and questions concerning the employer's attitudes about various kinds of applicants including welfare recipients, applicants with long spells out of the labor market, unstable work histories, or criminal records. In

addition, several survey items were added to mirror the audit study more closely (as described earlier).

The survey was administered by the Michigan State Survey Center. The final survey sample included 199 respondents, representing a 58 percent response rate. Response rates were calculated according to the basic formula $(I + P) / (I + P + R)$, where I equals the number of completed interviews, P equals the number of partial interviews, and R represents the number of refused eligible numbers (Groves and Lyberg, 1988). Between the time of the audit and that of the survey, two companies had declared bankruptcy, and an additional two had nonfunctioning numbers. These firms were dropped from the survey sample and excluded from the denominator for calculations of response rates.

Typical response rates for academic telephone surveys range from 50 to 80 percent. The current survey falls toward the lower end of the range of acceptable response rates as the result of several possible factors. Response rates for surveys of top management and organizational representatives typically lag behind those of employees or the general population (Baruch 1999). Likewise, there has been increasing resistance of businesses to participation in surveys, given the proliferation of market research firms and academics seeking employer participation for the growing number of studies involving businesses (Remington 1992). There has been a notable downward trajectory in the response rates from business surveys over the past 25 years (Baruch 1999; Cox et al. 1995), with increasing numbers of refusals citing that participation was against company policy (Fenton-O'Creevy 1996, cited in Baruch 1999). Even among the general population, Curtin, Presser, and Singer (2000) reported that the number of calls required to complete an average interview and the proportion of interviews requiring refusal conversion doubled between 1979 and 1996. The inundation of telemarketers (and, to a lesser extent, survey research) matched by the technological advances of caller-ID and privacy managers has made it increasingly difficult to recruit survey respondents for academic research (Remington 1992).

To assess the possible bias that may result from selective participation, two comparison tests were made. The first test compared basic characteristics of employers who responded to

²³ The first version of this survey was developed for the Multi-City Study of Urban Inequality Employer Survey (Holzer 1996). Holzer, Stoll, and Raphael (2002) later modified the initial instrument to focus more closely on applicants with criminal records. The instrument used for the current study was further modified to reflect the priorities of this research project.

the survey with the characteristics of those eligible for participation but refused. On the basis of industry, location, and call-back rates, the two groups were very similar, although some differences in occupational distribution were apparent: employers for restaurant jobs were the most likely to respond to the survey, whereas employers for laborer or service positions were the least likely. This difference probably has to do with the accessibility of employers in locally run restaurants, as compared with those in decentralized factories, warehouses, or companies. In an effort to account for this overrepresentation, key outcomes are recalculated using weights to achieve the sample distribution of the audit study (available upon request). A reweighting of the survey sample to match the distribution of the audit sample produced only a slight change in the mean likelihood (from .62 to .60). It is unlikely, therefore, that differential response rates of employers across industries have any

effect on the survey outcomes or on the differences between survey responses and observed behavior.

Even without these adjustments, however, the distribution of responses on key attitude items closely match that for a previous sample of Milwaukee employers. In a second test of sample bias, basic employer characteristics from the current sample were compared with an identical set of questions asked of a more representative sample of Milwaukee employers in 1999 (Holzer and Stoll 2001). Although the earlier Milwaukee survey included a broader geographic area and oversampled large firms, the general attitudes expressed by employers in both samples were strikingly similar (Table A1). The consistency of these findings provides some reassurance that the current sample can serve as a useful gauge for the priorities and concerns of employers in the broader Milwaukee metropolitan area.

Table A1. Comparison of Employer Attitudes and Characteristics across Two Milwaukee Surveys

Variable	Pager 2002	Holzer and Stoll 2001
Employees (n)	66.95	180.47
Vacancies (n)	4.48	7.79
Minority-owned Companies (%)	8.40	8.41
Unionized Employees (%)	9.30	15.19
Industry		
Manufacturing (%)	12.43	20.00
Retail trade (%)	49.72	21.00
Services (%)	21.47	39.00
Other industry (%)	16.38	20.00
Hire Welfare Recipient		
Definitely/probably would (%)	97.40	96.62
Definitely/probably not (%)	2.60	3.37
Hire Applicant with GED		
Definitely/probably would (%)	98.80	97.23
Definitely/probably not (%)	1.20	2.77
Hire Applicant with Criminal Record		
Definitely/probably would (%)	49.20	49.20
Definitely/probably not (%)	50.80	50.80
Hire Applicant Unemployed >1 year		
Definitely/probably would (%)	70.90	80.15
Definitely/probably not (%)	29.00	19.86
Hire Applicant with Unstable Work History		
Definitely/probably would (%)	60.50	67.49
Definitely/probably not (%)	39.50	32.51

Note: GED = General Education Diploma.

APPENDIX B

Table B1. Comparison of Employers' Self-Reports and Behavioral Outcomes for Overlapping Sample

Survey Results	Audit Results	
	No Call-Back	Call-Back
Likely to Hire Drug Offender		
Very Unlikely	31 (96.9 %)	1 (3.1 %)
Somewhat Unlikely	25 (88.3 %)	3 (10.7 %)
Somewhat Likely	69 (93.2 %)	5 (6.8 %)
Very Likely	20 (90.9 %)	2 (9.1 %)

Note: N = 199 respondents. Data shown as number of respondents with percent in parentheses.

REFERENCES

- Agresti, Alan and Brian Caffo. 2000. "Simple and Effective Confidence Intervals for Proportions and Differences of Proportions Result from Adding Two Successes and Two Failures." *American Statistician* 54:280-88.
- Ajzen, Icek. 1991. "The Theory of Planned Behavior." *Organizational Behavior and Human Decision Processes* 50:179-211.
- . 2001. "The Nature and Operation of Attitudes." *Annual Review of Psychology* 52:27-58.
- Baruch, Yehuda. 1999. "Response Rate in Academic Studies—A Comparative Analysis." *Human Relations* 52:421-38.
- Bertrand, Marianne and Sendhil Mullainathan. 2004. "Are Emily and Greg More Employable than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination." *American Economic Review* 94:991-1013.
- Bobo, Lawrence, Devon Johnson, and Susan A. Suh. 2000. "Racial Attitudes and Power in the Workplace: Do the Haves Differ from the Have-Nots?" Pp. 491-522 in *Prismatic Metropolis: Inequality in Los Angeles*, edited by Lawrence Bobo, Melvin Oliver, James Johnson Jr., and Abel Valenzuela Jr. New York: Russell Sage Foundation.
- Bobo, Lawrence and James Kluegel. 1997. "Status, Ideology, and Dimensions of Whites' Racial Beliefs and Attitudes: Progress and Stagnation." Pp. 93-120 in *Racial Attitudes in the 1990s: Continuity and Change*, edited by Steven A. Tuch and Jack K. Martin. Westport, CT: Praeger.
- Bobo, L., J. R. Kluegel, and R. A. Smith. 1997. "Laissez Faire Racism: The Crystallization of a 'Kinder, Gentler' Anti-Black Ideology." Pp. 15-42 in *Racial Attitudes in the 1990s: Continuity and Change*, edited by S. Tuch and J. Martin. Westport, CT: Praeger.
- Bonczar, Thomas P. 2003. "Prevalence of Imprisonment in the U.S. Population, 1974-2001." *Bureau of Justice Statistics Special Report*. Washington, DC: U.S. Department of Justice, NCJ 197976.
- Bradburn, N. M. 1983. "Response Effects." Pp. 289-328 in *Handbook of Survey Research*, edited by P. Rossi, J. Wright, and A. Anderson. New York: Academic Press.
- Bradburn, Norman M., Lance J. Rips, and Steven K. Shevell. 1987. "Answering Autobiographical Questions: The Impact of Memory and Inference on Surveys." *Science* 236:157-61.
- Brannon, Robert, Gary Cyphers, Sharlene Hesse, Susan Hesselbart, Roberta Keane, Howard Schuman, Tomas Viccaro, and Diana Wright. 1973. "Attitude and Action: A Field Experiment Jointed to a General Population Survey." *American Sociological Review* 38:625-36.
- Bureau of Labor Statistics. 2002. Local Area Unemployment Statistics. (Online Public Data Query). Retrieved March 1, 2003 (<http://www.bls.gov/lan/home.htm>).
- Couch, Kenneth and Mary C. Daly. 2002. "Black-White Wage Inequality in the 1990s: A Decade of Progress." *Economic Inquiry* 40:31-41.
- Cox, Brenda G., David A. Binder, B. Nanjamma Chinnappa, Anders Christianson, Michael J. Colledge, and Phillip S. Kott. 1995. *Business Survey Methods*. New York: John Wiley & Sons.
- Crandall, C. S. 1994. "Prejudice against Fat People: Ideology and Self-Interest." *Journal of Personality and Social Psychology* 66:882-94.
- Crosby, Faye, Stephanie Bromley, and Leonard Saxe.

1980. "Recent Unobtrusive Studies of Black and White Discrimination and Prejudice: A Literature Review." *Psychological Bulletin* 87:546-63.
- Curtin, Richard, Stanley Presser, and Eleanor Singer. 2000. "The Effects of Response Rate Changes on the Index of Consumer Sentiment." *Public Opinion Quarterly* 64:413-28.
- Darley, J. M. and P. H. Gross. 1983. "A Hypothesis-Confirming Bias in Labeling Effects." *Journal of Personality and Social Psychology* 44:20-33.
- Day, Dianne, Boon Gan, Philip Gendall, and Don Esslemont. 1991. "Predicting Purchase Behaviour." *Marketing Bulletin* 2:18-30.
- Deutscher, Irwin. 1966. "Words and Deeds: Social Science and Social Policy." *Social Problems* 13:235-54.
- Devine, Patricia. 1989. "Stereotypes and Prejudice: Their Automatic and Controlled Components." *Journal of Personality and Social Psychology* 56:5-18.
- Devine, P. G. and A. J. Elliot. 1995. "Are Racial Stereotypes Really Fading? The Princeton Trilogy Revisited." *Personality and Social Psychology Bulletin* 21: 1139-50.
- Dovidio, John, Kerry Kawakami, and Samuel Gaertner. 2002. "Implicit and Explicit Prejudice and Interracial Interaction." *Journal of Personality and Social Psychology* 82:62-68.
- Downing, David. 1984. "Employer Bias toward the Hiring and Placement of Male Ex-Offenders." Ph.D. dissertation, Department of Education, Southern Illinois University, Carbondale, IL.
- Emerson, Michael, George Yancey, and Karen Chai. 2001. "Does Race Matter in Residential Segregation? Exploring the Preferences of White Americans." *American Sociological Review* 66:922-35.
- Farley, Reynolds. 1997. "Racial Trends and Differences in the United States 30 Years After the Civil Rights Decade." *Social Science Research* 26:235-62.
- Farley, Reynolds, Suzanne Bianchi, and Diane Colasanto. 1979. "Barriers to the Racial Integration of Neighborhoods: The Detroit Case." *Annals of the American Association of Political and Social Sciences* 441:97-113.
- Farley, Reynolds, Howard Schuman, Suzanne Bianchi, Diane Colasanto, and Shirley Hatchett. 1978. "Chocolate City, Vanilla Suburbs: Will the Trend toward Racially Separate Communities Continue?" *Social Science Research* 7:319-44.
- Farley, Reynolds, Charlotte Steeh, Maria Krysan, Tara Jackson, and Keith Reeves. 1994. "Stereotypes and Segregation: Neighborhoods in the Detroit Area." *American Journal of Sociology* 100:750-80.
- Feagin, J. and M. Sykes. 1994. *Living with Racism, the Black Middle Class Experience*. Boston, MA: Beacon Press.
- Fishbein, M. 1967. "Attitude and the Prediction of Behavior." Pp. 477-492 in *Readings in Attitude Theory and Measurement*, edited by M. Fishbein. New York: Wiley.
- Fishbein, M. and I. Ajzen. 1975. *Belief, Attitude, Intention, and Behavior*. Reading, MA: Addison-Wesley.
- Fiske, Susan. 2004. "Predicting Discrimination: A Meta-Analysis of the Racial Attitude-Behavior Literature." Working Paper. Department of Psychology, Princeton University.
- Freeman, Richard B., and William M. Rodgers III. 1999. "Area Economic Conditions and the Labor Market Outcomes of Young Men in the 1990s Expansion." Working Paper No. 7073, Cambridge, MA: NBER.
- Gans, Herbert. 1992. "Second-Generation Decline: Scenarios for the Economic and Ethnic Futures of the Post-1965 American Immigrants." *Ethnic and Racial Studies* 15: 173-192.
- Groves, Robert M. and Lars E. Lyberg. 1988. "An Overview of Nonresponse Issues in Telephone Surveys." Pp. 191-212 in *Telephone Survey Methodology*, edited by Robert M. Groves, Paul P. Biemer, Lars E. Lyberg, James T. Massey, William L. Nicholls II, and Joseph Waksberg. New York: Wiley.
- Harrison, R. and C. Bennett. 1995. "Racial and Ethnic Diversity." Pp. 141-210 in *State of the Union, America in the 1990s*, vol. 2, *Social Trends 1990 Monograph Series*, edited by R. Farley. New York: Russell Sage Foundation.
- Heckman, James. 1998. "Detecting Discrimination." *The Journal of Economic Perspectives* 12: 101-116.
- Heckman, James and Peter Seligman. 1993. "The Urban Institute Audit Studies: Their Methods and Findings." Pp. 187-258 in *Clear and Convincing Evidence: Measurement of Discrimination in America*, edited by Michael Fix and Raymond J. Struyk. Washington, DC: Urban Institute Press.
- Holzer, Harry. 1996. *What Employers Want: Job Prospects for Less-Educated Workers*. New York: Russell Sage Foundation.
- Holzer, Harry and David Neumark. 2000. "What Does Affirmative Action Do?" *Industrial and Labor Relations Review* 53: 240-71.
- Holzer, Harry, Steven Raphael, and Michael Stoll. 2003. "Employer Demand for Ex-Offenders: Recent Evidence from Los Angeles." Discussion Paper No. 1268-03, Institute for Research on Poverty, University of Wisconsin, Madison, WI.
- Holzer, Harry, Paul Offner and Elaine Sorensen. 2005. "What Explains the Continuing Decline in Labor Force Activity among Young Black Men?" *Labor History* 46: 37-55.
- Holzer, Harry and Michael Stoll. 2001. *Employers and Welfare Recipients: The Effects of Welfare*

- Reform in the Workplace*. San Francisco, CA: Public Policy Institute of California.
- Husley, Lonnie Freeman. 1990. "Attitudes of Employers with Respect to Hiring Released Prisoners." Ph.D. dissertation, Department of Industrial Education, Texas A&M University, TX.
- Jensen, W. and W. C. Giegold. 1976. "Finding Jobs for Ex-Offenders: A Study of Employers' Attitudes." *American Business Law Journal* 14:195-225.
- Kinder, D. and D. O. Sears. 1981. "Prejudice and Politics: Symbolic Racism Versus Racial Threats to the Good Life." *Journal of Personality and Social Psychology* 40:414-31.
- Kirschenman, Joleen and Katherine Neckerman. 1991. "'We'd Love to Hire Them, but . . .': The Meaning of Race for Employers." Pp. 203-232 in *The Urban Underclass*, edited by Christopher Jencks and P. Peterson. Washington, DC: Brookings Institute.
- Krysan, Maria. 1999. "Prejudice, Politics, and Public Opinion: Understanding the Sources of Racial Policy Attitudes." *Annual Review of Sociology* 26:135-68.
- . 2002. "Whites Who Say They'd Flee: Who Are They, and Why Would They Leave?" *Demography* 39:675-96.
- Kutner, Bernard, Carol Wilkins, and Penny Rechtman Yarrow. 1952. "Verbal Attitudes and Overt Behavior Involving Racial Prejudice." *Journal of Abnormal Social Psychology* 47:649-52.
- LaPiere, Richard T. 1934. "Attitudes vs Actions." *Social Forces* 13:230-37.
- Lichtenstein, Sarah and J. Robert Newman. 1967. "Empirical Scaling of Common Verbal Phrases Associated with Numerical Probabilities." *Psychonomic Science* 9:563-64.
- Mare, R. D. 1995. "Changes in Educational Attainment and School Enrollment." Pp.155-213 in *State of the Union: America in the 1990s*, vol. I, *Economic Trends*, edited by R. Farley. New York: Russell Sage Foundation.
- Massey, Douglas and Garvey Lundy. 2001. "Use of Black English and Racial Discrimination in Urban Housing Markets: New Methods and Findings." *Urban Affairs Review* 36:452-69.
- Merton, Robert K. 1949. "Discrimination and the American Creed." Pp. 99-126 in *Discrimination and the National Welfare*, edited by Robert M. MacIver. New York: Harper and Row.
- Moss, Philip and Chris Tilly. 2003. *Stories Employers Tell: Race, Skill, and Hiring in America*. New York: Russell Sage Foundation.
- National Research Council. 2004. *Measuring Racial Discrimination*. Panel on Methods for Assessing Discrimination, Committee on National Statistics, Division of Behavior and Social Sciences and Education, edited by Rebecca M. Blank, Marilyn Dabady, and Constance F. Citro. Washington, DC: The National Academies Press.
- Pace, C. Robert and Jack Friedlander. 1982. "The Meaning of Response Categories: How Often Is 'Occasionally,' 'Often,' and 'Very Often'?" *Research in Higher Education* 17:267-81.
- Pager, Devah. 2003. "The Mark of a Criminal Record." *American Journal of Sociology* 108:937-75.
- . 2002. "The Mark of a Criminal Record." Doctoral dissertation. Department of Sociology, University of Wisconsin-Madison.
- Pettit, Becky and Bruce Western. 2004. "Mass Imprisonment and the Life Course: Race and Class Inequality in U.S. Incarceration." *American Sociological Review* 69:151-69.
- Poskocil, A. 1977. "Encounters Between Blacks and White Liberals: The Collision of Stereotypes." *Social Forces* 55:715-27.
- Remington, Todd D. 1992. "Telemarketing and Declining Survey Response Rates." *Journal of Advertising Research* 32:6-7.
- Saenger, Gerhart and Emily Gilbert. 1950. "Custom Reactions to the Integration of Negro Sales Personnel." *International Journal of Opinion and Attitude Research* 4:57-76.
- Sagar, H. and J. W. Schofield. 1980. "Racial and Behavioral Cues in Black and White Children's Perception fo Ambiguously Aggressive Acts." *Journal of Personality and Social Psychology* 39:590-98.
- Schaeffer, Nora Cate. 1991. "Hardly Ever or Constantly? Group Comparisons Using Vague Quantifiers." *Public Opinion Quarterly* 55:395-423.
- Schuman, Howard. 1995. "Attitudes." Pp. 68-89 in *Sociological Perspectives on Social Psychology*, edited by Karen Cook, Gary Fine, and James House. Boston, MA: Allyn & Bacon.
- Schuman, Howard and Lawrence Bobo. 1988. "Survey-Based Experiments on White Racial Attitudes toward Residential Integration." *American Journal of Sociology* 94:273-99.
- Schuman, Howard and Michael P. Johnson. 1976. "Attitudes and Behaviors." *Annual Review of Sociology* 2:161-207.
- Schuman, Howard, Charlotte Steeh, Lawrence Bobo, and Maria Krysan. 2001. *Racial Attitudes in America: Trends and Interpretations*. Rev. ed. Cambridge, MA: Harvard University Press.
- Schutz, Heidi and Bernd Six. 1996. "How strong is the Relationship Between Prejudice and Discrimination? A Meta-Analytic Answer." *International Journal of Intercultural Relations* 20: 441-62.
- Smith, Tom W. 1991. *Ethnic Images: General Social Survey Technical Report, 19*. Chicago, IL: National Opinion Research Center, University of Chicago.
- Sniderman, Paul M. and Douglas B. Grob. 1996.

- "Innovation in Experimental Design in Attitude Surveys." *Annual Review of Sociology* 22:377-99.
- Sniderman, Paul M. and Thomas Piazza. 1993. *The Scar of Race*. Cambridge, MA: The Belknap Press of Harvard University Press.
- Traugott, Michael W. and John P. Katosh. 1979. "Response Validity in Surveys of Voting Behavior." *Public Opinion Quarterly* 43:359-77.
- Turner, Margery, Michael Fix, and Raymond Struyk. 1991. *Opportunities Denied, Opportunities Diminished: Racial Discrimination in Hiring*. Washington, DC: Urban Institute Press.
- Waldinger, Roger and Michael Lichter. 2003. *How the Other Half Works: Immigration and the Social Organization of Labor*. Berkeley, CA: The University of California Press.
- Western, Bruce and Becky Pettit. Forthcoming. "Black-White Wage Inequality, Employment Rates, and Incarceration." *American Journal of Sociology*.
- Wicker, A. W. 1969. "Attitudes versus Actions: The Relationship of Verbal and Overt Behavioral Responses to Attitude Objects." *Journal of Social Issues* 25:41-78.
- Wilson, William Julius. 1978. *The Declining Significance of Race*. Chicago, IL: University of Chicago Press.
- . 1996. *When Work Disappears: The World of the New Urban Poor*. New York: Vintage Books.
- Word, C. O., M. P. Zanna, and J. Cooper. 1974. "The Nonverbal Mediation of Self-Fulfilling Prophecies in Interracial Interactions." *Journal of Experimental Social Psychology* 10:109-20.
- Yinger, John. 1995. *Closed Doors, Opportunities Lost*. New York: Russell Sage Foundation.