

## Situation Testing for Employment Discrimination In the United States of America

Marc Bendick, Jr.



**Marc Bendick, Jr.** received his Ph.D. in economics from the University of Wisconsin. For 30 years, he has been a Washington-based researcher and policy analyst on problems of poverty, under-employment, and discrimination in the USA. He is the author of more than 100 scholarly publications, a consultant on workforce diversity management to some of the nation's largest employers, and a frequent expert witness in large class action employment discrimination litigation. Since 1984, Marc has been a Principal in Bendick and Egan Economic Consultants, in Washington, DC USA.

### Summary

Over the past decade, more than 30 situation testing studies have been conducted in the USA on race, gender, age, and other forms of discrimination in the labor market. This research has typically documented discriminatory behavior, conscious or unconscious, by 20% to 40% of employers. These findings have influenced public opinion and government policy by publicizing the continued prevalence of discrimination in a particularly persuasive way. They have also been used to change employers' behavior through legal action, and they have added to our understanding of the psychological and organizational processes of discrimination. Expanded use of this investigative technique can contribute importantly toward addressing the serious problem of employment discrimination in industrial societies.

### Key words

minorities – race – women – hiring – stereotypes – affirmative action

## **Why the USA has conducted situation testing for employment discrimination**

In the United States of America (USA) and other industrial nations, the labor market is a “high stakes” institution. It is a major determinant of the income, social status, and psychological rewards experienced by individuals and families, and thus of personal well-being and social inclusion. Concurrently, the efficiency of the labor market is a major determinant of national economic growth and international competitiveness. Accordingly, ensuring the proper functioning of this socio-economic institution is a perennial concern of governments.

Economists define employment discrimination as valuation in the labor market of workers' characteristics which are not related to the workers' productivity (Arrow, 1998). This definition establishes a direct link between discrimination and loss of economic efficiency which, especially in a globally-competitive economy, leads nations to lose national income. However, such losses tend to be invisible. They accrue in small amounts through thousands of daily decisions on hiring, promotions, job assignments, or terminations of individual workers, where it is essentially impossible to measure the long-term productivity-reducing impact of each discriminatory decision. The cumulative effects are eventually reflected in lower national productivity figures, but intermingled with so many other simultaneous economic developments that productivity losses due to discrimination are virtually impossible to isolate.

Employment discrimination also creates in the minds of its victims a sense of inequity and disenfranchisement that threatens national social solidarity. This sense has tended to increase over time as nations have legally recognized rights to equality for an expanding range of personal characteristics<sup>1</sup> and as the populations of nations, such as those in North America and Europe, have become more demographically diverse<sup>2</sup>. In some cases, expressions of this sense of unfairness have been dramatic - most notably, in large-scale civil disorders which racked minority-dominated neighborhoods in large cities in the USA during the 1960s and in France more recently.

Studies conducted after these civil disorders in the USA clearly documented that a central source of resentment for many rioters was their work lives of high unemployment, low wages, and low status (Kerner, 1968). But these poor labor market outcomes reflect both discrimination in the labor market itself - that is, biased behavior by employers - and “pre-market” discrimination - that is, differences in

---

(1) In the USA, Title VII of the federal Civil Rights Act of 1964 made equal employment opportunity a legal right for racial/ethnic minorities, persons of non-US birth or ancestry, persons of all religions, and women. Similar provisions for other groups followed in the Age Discrimination in Employment Act of 1967 and the Americans with Disabilities Act of 1992. Legal protection for other personal characteristics is still developing. For example, gender orientation and marital/ parental status are legally protected in only some of the USA's 50 states.

(2) For instance, California, Texas, and several other states within the USA are “majority minority,” meaning that more than 50% of all residents belong to one or more racial/ethnic “minority” groups (Bendick, 1996, p. 2-6).

workers' access to quality schooling, cultural enrichment, health care, residential location, and other factors which equip job seekers with prerequisites for many better-paid, more prestigious jobs. In these circumstances, studies which show that African Americans, Hispanics, or women in the USA on average earn lower wages than white males do not readily measure discrimination by employers *per se* because of the complexities of separating the impact of employer behavior from broader social and economic discrimination (Neal and Johnson, 1996).

Partly because of this lack of definitive research about the prevalence and consequences of employment discrimination, public policy debates in the USA on this topic have often been dominated by ideology and rhetoric rather than facts and data. For the first decades after the USA's major civil rights laws were passed in the mid-1960s, this lack had little effect on the nation's action against employment discrimination. Initiatives were carried along by the moral force of America's "civil rights revolution" symbolized by the work of Dr. Martin Luther King, Jr. The general public and government officials were readily convinced that employment discrimination existed because they had personally observed it in its most obvious forms. On television, they had seen African Americans in the Southern states encountering workplace signs announcing, "No colored workers need apply." In newspaper job vacancy advertisements, they had seen employment opportunities divided into separate listings for "help wanted-males" and "help wanted-females."

As the 1960s and 1970s gave way to the 1980s, increasing numbers of employers abandoned such obvious, conscious forms of discrimination. However, race-ethnic minorities, women, and other groups in the American labor market increasingly encountered discrimination in more subtle forms. African Americans were no longer denied jobs simply because they were African American. They were still not being hired, but now it was because employers stereotypically assumed that all African Americans had such weak educational background that they were unqualified to be productive. Women were no longer denied promotions simply because they were women. They are still not being promoted, but now it is because when they were hired, they were assigned to jobs which gave them no experience preparing them for promotion (Moss and Tilly, 2002). Increasingly, unequal employment outcomes tended to be the product not of conscious racism or sexism but of more subtle, often unconscious, bias. They tended to be generated by psychological and organizational processes such as stereotyping, "micro-inequities," "social comfort," differential access to social networks, differential access to information, lowered career aspirations, "subjective decision-making," and occupational segregation (Bielby, 2000; Gaertner and Dovidio, 1986; Greenwald, McGeer and Schwartz, 1998; Hilton and von Hippel, 1996; Valian, 1998).

Reflecting this shift, more and more employers in the USA professed to be puzzled and troubled as minorities and women have remained under-represented in their workforces because these employers, being unconscious of these processes, are sure that they themselves are free of bias. More and more of the general public came to agree that employers were no longer discriminating. For example, one typical nationwide public opinion poll in 1989 reported that only 37% of white

respondents thought that an African American applicant who was as qualified as a European American would be less likely to be hired for a job they both wanted, and only 41% felt that the African American would be less likely to be promoted to a supervisory position (Bendick, 1999, p. 54).

In the USA, where public opinion leads, government policy soon follows. During the 1980s, political support for vigorous action against employment discrimination began to falter. In 1981, the conservative Ronald Reagan was elected president, and under his administration, enforcement of anti-discrimination laws by the federal Equal Employment Opportunity Commission dramatically weakened (Leonard, 1985). During the 1980s, several decisions by the US Supreme Court made discrimination litigation against employers more difficult to pursue. In 1996, voters in California amended their state constitution to forbid “affirmative action” compensating for employment discrimination.

Anti-discrimination researchers and advocates had long felt frustrated by the limited empirical evidence they could muster to support their work. However, prior to these new political developments, the major impact of this lack of evidence had been simply to slow their efforts. Now their long-term frustration transformed into an urgent sense of political crisis. American society might be shifting from making progress against employment discrimination too slowly toward moving in the opposite direction - revoking public policies which had been generating that progress. More than anything else, this sense of crisis was what led researchers and advocates to turn to situation testing as a new investigative technique (Bendick, 1999, p. 55).

### **How situation testing has been implemented in the USA**

Situation testing has unique potential for studying the behavior of actual employers in real workplaces while maintaining the methodological rigor of a laboratory-like scientific experiment. It is therefore appropriate to define the technique in a way emphasizing its links to rigorous empirical research traditions in the social and behavioral sciences. In this spirit, we define situation testing<sup>3</sup> as a systematic research procedure for creating controlled experiments analyzing employers' candid responses to employees' personal characteristics.

As noted in the previous section, economists define employment discrimination as valuation in the labor market of workers' characteristics not related to productivity. In situation testing, pairs of research assistants are sent to apply for the same actual job vacancy. Within each pair, employee characteristics likely to be related to a worker's productivity on the job - such as education, work experience, professional certifications, and technical skills - are made equal by selecting, training, and

---

(3) Outside the USA, the methodology examined in this paper is usually referred to as “situation testing.” In the USA, synonyms such as employment testing, employment auditing, and paired-comparison testing are more common.

credentialing testers to appear equally qualified for the positions they seek. Simultaneously, personal characteristics unrelated to job performance are experimentally manipulated by pairing testers who differ in only one of these characteristics - for example, an American of European ancestry and an American of African ancestry, a male and a female, or a person age 32 and a person age 57. If testers within these pairs experience substantially different responses to their job-seeking efforts, few assumptions are required to interpret that difference as the employers' reaction to that one differing personal characteristic.

This interpretation is appropriate only if employers are presented with pairs of job candidates who truly appear equally qualified. This condition is relatively easy to achieve in situation testing studies involving only paper resumes, which are mailed, faxed or e-mailed to employers (Bertrand & Mullainathan, 2002; Bendick, Jackson & Romero, 1996). In these studies, testers' resumes describe equivalent education, work experience, and job skills while varying resume details and resume formats to avoid appearing obviously similar. The resumes communicate the demographic characteristics of the applicants through such means as gender-specific names, ethnically-related activities, or age-revealing dates of receiving degrees.

Resume-based tests can probe only the initial stages of the hiring process, up to an employer's decision to invite job candidates for in-person interviews. To study the complete hiring process, it is necessary to dispatch equally-qualified pairs of persons posing as job applicants. Maintaining a controlled experiment in that circumstance requires substantial care at each step in the research process (Bendick, 1989; Bovenkerk, 1992; Lodder, 1994).

The first step in maintaining the controlled experiment is to recruit testers who meet a daunting set of requirements: ability to play the job-seeker role convincingly while simultaneously making accurate observations about the hiring process; willingness to approach the study objectively; similarities between testing partners in general appearance and demeanor; and the differing demographic characteristics required by the study design. University students, professional actors, actual job seekers, and adult volunteers have all been recruited for the tester role. Recruiting individuals meeting all the requirements of the position is often a time-consuming and painstaking process. In one typical study, 93 potential testers were interviewed before 4 testers were selected to form 2 testing pairs (Nunes and Seligman, 1999, p. 6).

The second step in maintaining the controlled experiment is training to make pairs of testers equally credible job applicants. During training, testers develop and memorize their false resumes, receive coaching on effective job interviewing techniques, and rehearse similar answers to common interview questions. Concurrently, testers are trained to be "human tape recorders" by drilling them to notice and remember important details of their job application experiences and instilling the value of objective observation. In typical well-run testing studies, training requires about 3 days.

A third step in maintaining the controlled experiment involves carefully managing testers' actions throughout the job application process. The two testers within a pair usually present themselves to employers in random order, with the second tester applying a few minutes after the first. Each tester documents his or her experiences as soon as practical after the event and prior to being told the experiences of her/his testing partner. Testers typically record their data in writing using pre-structured questionnaires, and they are constantly reminded to focus on observable facts rather than to make judgments or interpretations about what they observe. Such careful management requires continuous, hands-on monitoring of each test by a trained "Test Coordinator," who usually can adequately supervise no more than 3 testing teams concurrently.

The final step in maintaining a controlled experiment is to repeat the job application experiment for dozens or hundreds of job vacancies, to "average out" random circumstances which may affect the outcome in any single test. In analyzing test outcomes, the basic statistical measure is the "net rate of discrimination" - the proportion of tests in which testers with the characteristic hypothesized to be disfavored (e.g., African American) is successful minus the proportion of tests in which testers with the characteristic hypothesized to be favored (e.g., European American) is successful.<sup>4</sup> Here, "successful" is typically defined as reaching an identifiable milestone in the hiring process, such as being offered a job interview or being offered a job.

Given the net rates of discrimination observed in typical testing studies, statistically significant estimates of the main effect studied in the experiment - e.g., whether the net rate of discrimination for females paired with equally-qualified males is above zero - have been obtained with as few as about 50 completed tests. Samples of about 100 tests have sometimes proved sufficient to observe statistically significant effects of interacting circumstances ("mediating factors") on net rates of discrimination - for example, in estimating whether the net rate of discrimination against women is higher in occupations offering higher earnings. However, when testing results are analyzed using more complex statistical techniques, larger numbers of completed tests are typically required (Kenney and Wissoker, 1994).

The following are examples of employment outcomes which situation testing studies have reported as discrimination:

---

(4) In testing studies in the USA, tests in which both testers are unsuccessful at the beginning of their job application (e.g., both are told that the job has already been filled) are usually included in computing the net rate of discrimination. In studies conducted under the auspices of the International Labour Organization (ILO), in contrast, such tests are normally excluded. The former procedure is particularly appropriate if even such early, apparently-innocent outcomes actually reflect employers' decisions - for example, if, although the advertised vacancy has been filled, the employer is still open to hiring applicants they consider particularly promising for other, unadvertised vacancies. The latter procedure is appropriate if both applicants were really not considered by the employer - for example, where hundred of applications have already been received for one job vacancy, and that is the only position the employer is filling. The latter approach generates higher estimated rates of discrimination than the former. For example, in the study reported in the first column of Table 1 later in this paper, the net rate of discrimination against African American job seekers estimated under the "USA" procedure is 13.0%, while under the "ILO" procedure, it is 22.3% (Bendick, 1999, p. 17-20).

- A large-circulation newspaper carried an advertisement for a supervisor at a restaurant in an affluent neighborhood. The African American tester who presented himself at the restaurant was told that he would be called if the restaurant wished to pursue his application. Minutes later, a European American tester whose resume showed the same level of education and restaurant experience followed the same procedure. He was called later that day to schedule an interview, interviewed the day after that, and subsequently offered the position. The African American tester made four follow-up calls to reiterate his interest, including one shortly after the European American tester refused the job offer, with no response (Bendick, Jackson, and Reinoso, 1994, p. 33).
- A vacancy for a receptionist in an optometrist's office was advertised in a local newspaper in an affluent neighborhood. When a tester with a Latina name and slight accent telephoned the next day to apply, she was put on hold, called Carmen when she had given her name as Juanita, and told that the office was not taking any further applications. When her testing partner with an Anglo name and no accent called 13 minutes later, she was given an appointment for an interview the following morning (Bendick, Jackson, Reinoso, and Hodges, 1991, p. 475).
- An employment agency advertised for an "account representative" to do executive recruiting. Two white males, whose resumes and appearance portrayed them as age 32 and 57 respectively, responded by telephone and were granted interviews. The older tester's interview lasted 48 minutes, during which the tester was cautioned against making a precipitous career change and instructed to call back if he was still interested after reading books on sales techniques. The younger tester's interview lasted 85 minutes, during which the interviewer discussed a variety of work and non-work topics in a friendly manner and commented enthusiastically on the tester's questions and responses. This tester was invited back for a second interview, after which he was offered a job (Bendick, Brown, and Wall, 1999, p. 12-13).
- An automobile services shop advertised in a newspaper for a service technician to lubricate and repair automobiles. When a female applicant whose resume showed experience in physically hard jobs applied for the position, the manager who interviewed her told her that "the auto lube job is hard for a woman," said that he liked her smile, and offered her an alternative, lower-paying position serving coffee to customer while they waited for their cars. When her male testing partner applied for the position several hours later, he was interviewed for the advertised service position (Nunes and Seligman, 2000, p. 10).

### **Testing-based estimates of the prevalence of discrimination**

Occasional employment studies in the spirit of situation testing were completed in the USA as long ago as the 1950's. However, studies which describe the contemporary American labor market and embody the substantial samples and methodological rigor described earlier in this paper began

around 1990, based initially on a design developed in 1989 (Bendick, 1989). Since then, at least two dozen well-documented studies have appeared in the USA, conducted by a number of different researchers (Table 1)<sup>5</sup>. Additional testing studies currently being conducted involve, for example, posting resumes for women and racial minorities on internet job boards, sending racial/ethnic minorities with and without accents to apply for server positions in upscale restaurants, and sending persons in wheelchairs to apply for office positions.

Race/ethnicity/color/national origin is the personal characteristic most commonly examined in this research, accounting for 16 of the 22 studies summarized in the table. Among these studies considered credible<sup>6</sup> and covering all stages of the hiring process, the estimated net rate of discrimination for African Americans compared to European Americans or Hispanics/Latinos compared to Anglos, averaged about 20%. Employment discrimination based on gender was examined in 3 studies in the table, with an average net rate of discrimination of 25%. Discrimination based on workers' age was measured in three studies, with an average net rate of discrimination for the two studies covering all stages of the application process of 29%<sup>7</sup>.

---

(5) Several dozen studies have also been completed in industrialized nations other than the USA (Riach and Rich, 2002). For instance, discrimination has been studied for immigrants and race/ethnic minorities such as West Indians in Canada (Henry and Ginzberg, 1985), Turks in Germany (Goldberg, Mourinho, and Kulke, 1996), and Moroccans in Spain (de Prada et al., 1996). Other studies have investigated discrimination against women in Austria (Weichselbaumer, 2000), older workers in Australia (Gringart and Helmes, 2001), and persons with disabilities in Great Britain (Frye, 1986). Like the studies in Table 1, these studies nearly unanimously conclude that such groups encounter substantial discrimination when seeking employment in their respective labor markets.

(6) Only one study in the USA found that minority testers were treated better than their non-minority counterparts - an estimated -10% net rate of discrimination for Latinos in Denver (James & DelCastillo, 1992). This study is excluded in computing the average reported in this paragraph because its methodology and rigor have been seriously questioned (Fix and Struyk, 1993, p. 175-177).

(7) No situation testing studies of discrimination based on disabilities have been published in the USA. Among four studies conducted in European labor markets, the estimated net rate of discrimination averaged 36% (Ravaud, Madiot, and Ville, 1992; Gras, et al., 1996; Fry, 1986; and Graham, Jordan, and Lamb, 1990).



**Table 1 - Selected situation testing studies of employment discrimination in the USA**

<b>Demographic Contrast</b>	<b>African American/European American</b>									
<b>Author(s)</b>	Turner, Fix & Struyk (1991)	James & DelCastillo (1992)	Bendick, Jackson & Reinoso (1994)	Nunes and Seligman (1999)	Bertrand & Mullainathan (2002)	Pager (2003)	Lodder, McFarland & White (2003)	Lodder, McFarland & White (2003)	Busey & Trasvina (2003)	Pager and Western (2005)
<b>Tests completed</b>	476	145	149	45	130	350	80	169	109	252
<b>Location of jobs</b>	Chicago, Washington	Denver	Washington	San Francisco	Boston, Chicago	Milwaukee	Chicago	Chicago	San Francisco, Los Angeles	New York
<b>Source of job sample</b>	newspaper ads	newspaper ads	newspaper ads, industry lists, walk-ins	industry lists	newspaper ads	newspaper ads, internet lists	newspaper ads, internet ads, walk-ins	newspaper ads, internet ads	industry lists	newspaper ads, job fairs
<b>Method of application</b>	in-person	in-person	in-person	in-person	mailed resumes	in-person	in-person	mailed resumes	in-person	in-person
<b>Education in resumes</b>	completed secondary school	completed secondary school	2 years of university	2 years of university	incomplete secondary school to university graduate	completed secondary school	completed secondary school	completed secondary school	2 years of university	completed secondary school
<b>Industry or occupation tested</b>	mixed entry-level	mixed entry-level	mixed entry-level	Employment agencies	mixed entry-level	mixed entry-level	suburban retail	suburban retail	employment agencies	mixed entry-level
<b>Interacting circumstances</b>	none	none	none	none	qualifications, residence	criminal record	customer treatment	skill level	none	criminal record
<b>Net rate of discrimination</b>	<b>13% *</b>	<b>2%</b>	<b>24%</b>	<b>38%</b>	<b>3% *</b>	<b>17%</b>	<b>12%</b>	<b>5% *</b>	<b>31%</b>	<b>10%</b>

**Table 1 - Selected situation testing studies of employment discrimination in the USA (cont'd.)**

Demographic Contrast	Hispanic / Anglo						Females/Males			Older / Younger		
<b>Author(s)</b>	Cross, Kennedy, Mell & Zimmerman (1990)	Bendick, Jackson, Reinoso & Hodges (1991)	Bendick, Jackson, Reinoso & Hodges (1991)	James & DelCastillo (1992)	Firestone, Yanoff & Montenegro (2002)	Pager and Western (2005)	Neumark (1996)	Nunes & Seligman (2000)	Discrimination Research Center (2004)	MCAD (1994)	Bendick, Jackson & Romero (1996)	Bendick, Brown & Wall (1999)
<b>Tests completed</b>	360	282	186	140	122	252	65	40	24	49	79	102
<b>Location of jobs</b>	Chicago, San Diego	Washington	Washington	Denver	Washington	New York	Philadelphia	San Francisco	San Francisco	Boston	nationwide	Washington
<b>Source of job sample</b>	newspaper ads	newspaper ads	newspaper ads, industry lists	newspaper ads	newspaper ads, industry lists	newspaper ads, job fairs	industry list	newspaper ads, industry lists	industry list	Newspaper ads	lists	Newspaper ads
<b>Method of application</b>	in-person	telephone	mailed resumes	in-person	telephone	in-person	in-person	in-person	in-person	in-person	mailed resumes	in-person
<b>Education in resumes</b>	completed secondary school	some university	university graduate	completed secondary school	some university	completed secondary school	completed secondary school	completed secondary school	completed secondary school	not documented	university graduate	university graduate
<b>Industry or occupation tested</b>	mixed entry-level	mixed entry-level	mixed entry-level, employment agencies	mixed entry-level	mixed entry-level, employment agencies	mixed entry-level	restaurants	auto service shops	construction trades	entry-level office and retail	mixed sales, office, and professional	sales and managerial
<b>Interacting circumstances</b>	none	none	none	none	none	criminal record	restaurant price range	none	none	none	cover letter strategy	none
<b>Net rate of discrimination</b>	<b>20%</b>	<b>22%*</b>	<b>&gt; 12%*</b>	<b>-10%</b>	<b>25%</b>	<b>4%</b>	<b>40%</b>	<b>27%</b>	<b>8%</b>	<b>17%</b>	<b>27%*</b>	<b>41%</b>

Notes: The net rate of discrimination is the success rate for the presumed disfavored group minus the success rate for the presumed favored group. Negative numbers mean that the presumed disfavored group had greater success than the presumed favored one. Rates are based on all tests initiated, including those where both applicants received a negative response at the first stage of applications. Rates marked \* are incomplete measures based on only the first stages of the application process.

Source: adapted and updated from Bendick (1999), p. 56.

Collectively, these findings clearly support the conclusion that discrimination remains an important force in the contemporary labor market in USA<sup>8</sup>. Beyond that general conclusion, however, it is not usually informative to discuss net rates of discrimination without considering precisely what demographic groups were tested, what jobs were tested, in what locations the tests were conducted, and other important circumstances of the tests. This attention to detail is necessary because the studies themselves demonstrate that the prevalence of discrimination is importantly affected by such contextual factors. For example:

- Job seeking processes. Job applicants commonly seek jobs simultaneously through multiple channels, ranging from open processes (e.g., “help wanted” advertisements in widely-circulated newspapers) to closed ones (e.g., “spreading the word” through current employees’ personal contact networks of families and friends). Closed processes are presumed to be adopted in part for discriminatory reasons - to restrict knowledge about the job vacancy to channels where few minority job candidates are likely to hear about it. However, for practical reasons, situation testing studies have been conducted primarily on publicly advertised vacancies. This circumstance implies that the studies are likely to under-estimate the rate of discrimination in the overall labor market. Consistent with this hypothesis, for example, one situation testing study which estimated a net rate of discrimination against African Americans of 14.7% for publicly-advertised positions found a rate of 34.3% for positions which were not widely advertised and 66.7% for positions listed with private employment agencies (Bendick, Jackson, and Reinoso, 1994, p. 36).
- Job-seeking strategies. Some testing studies have incorporated experimental manipulation of job-seeking strategies. For example, in one study of age discrimination using mailed resumes, different cover letters were randomly assigned to accompany the older workers’ resume. In one letter, the older applicant described herself/himself as career-committed, energetic, and technologically up-to-date - positive attributes stereotypically associated with younger workers; an alternative letter described the older applicant as experienced, mature and stable - positive attributes stereotypically associated with older workers; and a third letter contained neither statement. The first cover letter generated a substantially higher rate of favorable employer responses to the older applicant than the other two (Bendick, Jackson, and Romero, 1996, p. 39-41).
- Employers’ stated policies. Some employers publicly advertise themselves as “equal opportunity employers.” However, one situation testing study found no difference between these employers and employers making no such claims in the probability that situation testing would discover discriminatory behavior (Bendick, Jackson, and Reinoso, 1994, p. 38).

---

(8) For a conservative’s dissent from this conclusion incorporating methodological criticisms of situation testing, see Heckman (1998). Heckman’s criticism rests heavily on speculative, ideological assumptions. Specifically, he assumes that all employers are forced by a competitive labor market to consider only productivity-related characteristics of job applicants, so that differences in employment outcomes between paired testers must signal failure by the testing process to match the testers on some subtle but important productivity-related characteristics.

## **Have these findings influenced public opinion and government policy?**

As was discussed earlier in this paper, the principal initial goal for situation testing studies in the USA was to document the continued presence of discrimination in ways which could convince public opinion and government policy-makers that employment discrimination remains a major problem in American society.

In terms of its likely impact on public opinion, situation testing provides information well suited to this goal. Psychological research has concluded that statistical evidence has longer-lasting persuasive impact on the receiver of information than do narrative individual stories, whereas individual stories tend to be more powerful than statistical evidence in attracting attention and eliciting emotional responses (Kopfman, Smith, Yun, and Hodges, 1998). Testing studies are usually reported in the same two-part manner used earlier in the present paper. That is, overall findings are presented statistically, as a “shocking” net rate of discrimination to summarize the overall findings and perhaps to fit easily into news headlines. Narrative descriptions of individual discrimination incidents are then presented in ways likely to be repeated in the text of news articles. These narratives put a human face on the abstract concept of discrimination, appealing to readers’ intuitive sense of justice and engaging readers’ personal sympathy.

Consistent with the prediction that situation testing results should have high impact, when individual testing studies are released in the USA, they tend to be prominently reported in the mass media. When presented to Congressional committees, state legislatures, and other governmental fact-finding processes, they tend to raise considerable excitement and to be quoted by political leaders (Bendick, 1995). Recognizing their newsworthiness, television networks in the USA have conducted their own situation tests on employment discrimination using hidden microphones and miniature cameras, broadcasting the results nationwide (Whipple, 1990).

Despite these successes, situation testing, either by itself or in conjunction with other kinds of research, has not completely reversed the political developments which Section 1 described as the major motivator for developing the technique. To be sure, in 1991, the US Congress passed a major new civil rights law over-riding the Supreme Court decisions which had threatened to undermine enforcement of anti-discrimination laws. On the other hand, not only did voters in California amend their state constitution in 1996 to outlaw affirmative action as a remedy for employment discrimination; state voters adopted the same provisions in Washington in 2000 and Michigan in 2006. In all three states, situation testing results had been prominently quoted by advocates seeking to persuade the voters not to do so (Bendick, 1995).

On balance, it is reasonable to credit situation testing with making some contributions to increasing public understanding of employment discrimination. Public opinion polls over the past decade show that a majority of the American public remains skeptical that “employment discrimination” is currently widespread. Yet concurrently, considerable national consensus has emerged that employers in the USA have not been dealing adequately with their increasingly diverse work forces. This consensus tends to avoid labels such as “employment discrimination” in favor of such alternative terms as “workforce diversity” or “workplace inclusion.” However, it tends to recognize the more subtle, often unconscious processes of discrimination which, as this paper discusses, tend to predominate today (Kochan, et al., 2003; Egan and Bendick, 2003). By helping to keep these issues before the public in particularly visible, memorable ways, situation testing has undoubtedly contributed to this new consensus.

### **Situation testing’s role in litigation and other efforts to change individual employers**

When employers in the USA violate the equal employment provisions of federal or state laws, they can be sued either by government agencies (such as the federal Equal Employment Opportunity Commission, EEOC) or in private litigation brought by the victims of discrimination. If the plaintiffs are victorious, the employer may be liable for substantial damage payments (in some cases, more than \$100 million), as well as mandatory, court-supervised changes in its employment practices. Since the mid-1960s, such litigation, or employers’ desire to avoid it, has been a major force reducing employment discrimination in the USA (Leonard, 1985).

To date, situation testing has played little role in this enforcement activity<sup>9</sup>. However, American law gives it considerable potential to do so. If an employer is sued based on evidence other than situation testing, documentation of an employer’s discriminatory behavior generated through situation testing can be used to corroborate and reinforce that other evidence. In addition, individual testers and non-profit, non-governmental organizations employing testers have “standing” to become plaintiffs in anti-discrimination litigation based solely on the testing evidence itself (Boggs, Seller and Bendick, 1993, p. 361-362; Landever, 1993; Oh, 1993; Yelmolinsky, 1993).

This latter provision gives testing the potential to close an important enforcement gap in the US legal system. Historically, only about 15% of complaints to the EEOC about employment discrimination have been about discrimination in hiring. Knowledgeable persons assume that the actual rate of such discrimination is considerably higher, but that it is particularly difficult for job applicant to sense or to prove that they have been discriminated against (compared, for example, to a person already employed by a firm who, if denied a promotion, knows who was promoted instead).

---

(9) In contrast, situation testing has been extensively applied in litigation against discrimination in housing (Boggs, Sellers, and Bendick, 1993, p. 345-349).

Despite this potential, only a handful of enforcement actions have been brought based on situation testing. For example, in 1990, a civil lawsuit, *Fair Employment Council v. BMC Marketing* was filed in the District of Columbia. In it, the plaintiffs were two African American university students employed as testers and the non-governmental, non-profit anti-discrimination organization which employed them. The defendant was a local office of a major nationwide job placement agency. The agency had interviewed, coached, and placed in jobs two European American university students who were the testing partners of the two African American plaintiffs, while it failed to do so for the African American testers. The litigation was subsequently settled with payment of damages to the non-profit organization and promises by the employment agency to discontinue its discriminatory practices (Boggs, Seller, and Bendick, 1993, p. 362-363).

Litigation is the most adversarial approach within a range of ways situation testing can be used to reduce discriminatory behavior by employers. For example, the results of situation tests on specific employers might be released to the general public, news media, insurers, investors, trade unions, and other “stakeholders” who have influence over corporations in the US, providing information by which these stakeholders can pressure discriminating employers to change their behavior (Egan, Mauleon, Wolff, and Bendick, 2007). Employers themselves might employ testers posing as job applicants to monitor whether their employees (such as managers interviewing job applicants) are following company policies against discrimination (Boggs, Sellers, and Bendick, 1993, p. 356-357; Wymer and Sudbury, 1992).

### **Using situation testing to deepen our understanding of discrimination**

As Section 2 of this paper discussed, many of the situation testing studies in Table 1 were conducted by policy analysts with pragmatic, short-term public policy advocacy goals. In consequence, analysis of testing results have emphasized simple summary numbers such as the net rate of discrimination. However, these same studies can be analyzed to reveal more detailed, complex information concerning how discrimination operates in the contemporary labor market. This information, in turn, can assist to reduce employment discrimination in the long run through shaping improved employment practices by employers and enforcement actions by government.

Table 2 provides some examples of such information, as reported in one situation testing study of race (Bendick, Jackson, and Reinoso, 1994). Many of these findings document aspects of the job seeking process not previously measured or analyzed in non-testing research.

**Table 2 : Testing-based information on the mechanisms of discrimination**

Indicator		Furthest Stage Reached	European American	African American	Ratio of African American / European American
Treatment	Met with a person authorized to make a hiring decision	Application Interview	26,3%	24,6%	,94
		Job Offer	73,3%	79,2%	1,08
		Job Offer	89,5%	83,3%	,93
	Minutes of contact	Application Interview	15,7	16,1	1,02
		Job Offer	17,2	13,3	,77
	Number of topics discussed	Application Interview	22,2	23,3	1,05
		Job Offer	0,8	0,82	1,02
	% moving to the next stage	Application Interview	2,98	3,02	1,01
Job Offer		3,33	2,5	,75	
Judgments expressed by employer	Number of comments made by the employer	Application Interview	94,6%	91,9%	,97
		Job Offer	71,0%	67,9%	,96
		Job Offer	0,22	0,13	,59
	Ratio of positive comments to negative comments	Application Interview	0,83	0,51	,61
		Job Offer	1,75	1,5	,86
		Job Offer	0,5	0,4	,80
Employment outcomes	Received a job offer	Application Interview	2,9	0,5	,17
	Average hourly wage offered	Job Offer	35,5	0,0	,00
	Offered the advertised job	Job Offer	46,9%	11,3%	,24
		Job Offer	\$5,45	\$5,30	,97
		Job Offer	98,0%	94,6%	,96

Source: adapted from Bendick, Jackson, and Reinoso, 1994, p.40.

One way in which situation testing can yield more such information is to equip testers with hidden tape recorders when they interact with employers, for example in job interviews. This practice has been common in the USA for a number of years in situation testing for housing discrimination, and it is starting to be experimented with for employment. This is a very promising direction for increasing the sophistication of situation testing, especially when the recorded material is analyzed using state-of-the-art research methods from psycholinguistic and communications theory (Bendick and Nunes, 2005).

Progress in deepening our understanding of employment discrimination will also be hastened if future situation testing studies move beyond their roots in short term advocacy to incorporate theoretical concepts from the social, behavioral, and managerial sciences. For example (Bendick and Nunes, 2005):

- The way in which test outcomes are reported need to be re-thought. The “net rate of discrimination,” emphasized Table 1, corresponds to the probability that a job seeker encounters discrimination in applying for one job vacancy. But in a typical job-seeking campaign in the USA, a job-seeker may apply for dozens of jobs. It would therefore be incorrect to interpret a testing-based net rate of discrimination of, for example, 20% to mean that only 20% of job seekers are adversely affected by discrimination. A more realistic measure would be the probability of encountering discrimination at least once in the course of a multi-application job campaign. Suppose we adopted as a convention for reporting discrimination rates the

probability that a job seeker would encounter discrimination at least once in 10 job applications. For a 20% probability of encountering discrimination in each application, the corresponding figure in this new measure would be 86.6%. This figure more accurately conveys how the labor market treats a typical job applicant than the 20% figure from which it is computed.

- Testing results are typically reported in a way which fails to communicate the long-term significance of the modest differences in treatment which testers often report. Some critics dismiss such differences as petty carping about minor indignities and insignificant inconvenience, not important economic losses. In particular, these critics minimize the importance of being tuned down for an entry-level job because the applicant can probably obtain similar employment from another employer after only a small delay (Heckman, 1998). By focusing solely on differences as testers report them, situation testing studies play into the hands of these critics.

Suppose, for example, that a tester is offered a sales job in a location with projected earnings of \$35,000 per year, and her/his testing partner is offered a different location whose earning potential is only 1% greater. The 1% figure corresponds to only \$350 in the first year, a mere \$.13 per hour. However, if the 2 testers are assumed to stay in that job for 30 years and each receive a 5% annual raise, by the end of 30 years, the difference in their earnings would total \$24,766. This amount, which is 70% of the one-year earnings the lower-paid tester was offered, would probably strike observers as a more serious matter than \$.13 per hour. Moreover, it would be untypical for two career paths which start at different levels to parallel each other, as this computation assumes. Differences in initial job offers, as well as the employer attitudes and perceptions underlying those differences, tend to put newly-hired employees on different paths for earnings increases and promotions. Accordingly, suppose that the tester who was offered 1% lower salary in the first year also received a 1% lower raise each subsequent year for the next 30 years. In the 30th year, that tester's annual earnings would be 34.2% lower than the other tester's, and the cumulative difference in their 30-year earnings would total \$521,000 - a measure of the impact of discrimination which is both more dramatic and more realistic<sup>10</sup>.

- The same personal characteristics associated with discriminatory treatment in the labor market tend to trigger discrimination other aspects of workers' daily lives - in renting or purchasing housing and obtaining home mortgages; obtaining services from public and private agencies, including health care providers and schools; interacting with law enforcement agencies; and receiving courteous treatment, fair prices, and quality merchandise when shopping (Fix and Turner, 1999; Cose, 1993, Feagin and Sikes, 1994). It is reasonable to hypothesize that discrimination in one venue will be more psychologically harmful to individuals and more destructive to social solidarity when the same individual has parallel experiences in other aspects of life. Situation testing offers unique opportunities to explore such effects through studies in

---

(10) Small differences in treatment also cumulate at the organizational level. For example, consider a workplace with 8 levels in its hierarchy, a system of "promotion from within," and 50% men and 50% women employees at the lowest level. If the promotional process at each level has a bias in favor of men of only 1 percentage point, then male representation at the top level in the organization will be 65% (Martel, Lane and Emrich, 1996). Thus, small inequities such as are documented by individuals in situation tests become amplified into broad patterns of great social concern, such as "glass ceilings" in the representation of women and minorities at higher levels of management.



which tests in employment are conducted simultaneously with parallel testing in other aspects of daily life. In this spirit, a proposal has been put forth in the USA to produce an annual “national report card” on discrimination gathering data on employment, housing, and consumer services simultaneously, using situation testing as a standardized methodology in each domain (Fix and Turner, 1999).

- Testing studies to date have implicitly embodied the model that employers who demonstrate discriminatory behavior in the hiring process are expressing hostility toward a specific demographic group - for example, African Americans. However, state-of-the-art theories of racism suggests that, in many cases, a different process may be at work - psychological and organizational processes in which employers disfavor persons who differ in any way from their “ideal” employee. Race might be one of those ways, but so might gender, age, disability, social class, personal appearance, former employer, or many other personal characteristics (Mazur, 1985). In such circumstances, it is a misleading to report the results of, say, a situation test pairing Hispanics and Anglos testers as measuring discrimination which is specifically anti-Hispanic.

Before adopting this new interpretation of testing results, however, it would be necessary to conduct situation testing to measure the extent to which the same employer equally disfavors a broad range of personal characteristics. Historically, testing studies have not been designed to test this hypothesis because each study addresses only one demographic characteristic (e.g., race). Instead, the same employers need to be tested for responses to testers differing on a range of demographic characteristics.

## **The future of situation testing for employment discrimination**

In all three venues in which it has been applied in the USA - public opinion and government policymaking, litigation, and research - situation testing for employment discrimination has proved to be a feasible and insight-generating investigative method. The modest number of testing studies which have been completed to date exploit only a small fraction of its potential. The primary challenge facing the situation testing community in the USA is to “scale up” from a “demonstration” level of activity to widespread, routine use.

This paper has argued that more sophisticated study designs incorporating concepts from the social, behavioral and organizational sciences would increase the insights and impacts to be gleaned from such expanded activity. However, additional studies simply following the precedents of those in Table 1 would make substantial contributions. In particular, it would be useful to conduct a systematic, coordinated series of studies to “map out” patterns of discrimination in today’s labor market. Particularly useful studies might address the following three questions<sup>11</sup>:

---

(11) A fourth suggestion is often put forth - conducting studies in cities where situation testing has not been previously conducted. Testing studies to date suggest that differences among cities in net rates of discrimination tend to be modest, suggesting that in terms of generating new knowledge, such repetitions should not receive

- How prevalent is discrimination in different parts of the labor market? Non-testing research has identified certain industries or occupations in which discrimination appears to be particularly prevalent<sup>12</sup>, and situation testing targeted to these “trouble spots” could contribute by directing anti-discrimination enforcement efforts there.
- Is the prevalence of discrimination changing over time? Research using methodologies other than situation testing generally suggests a slow downward trend over time in the extent of employment discrimination in the USA. To date, no situation testing studies have been used to confirm or contradict this generalization by repeating studies which are parallel in design but separated in time. However, as mentioned earlier in this paper, a proposal has been set forth to create a “national report card on discrimination,” in which a standardized situation testing study would be repeated annually to produce that time series of findings (Bendick, 1999).
- How prevalent is discrimination against demographic groups for which few or no studies have yet been conducted? In the USA, these groups include persons with disabilities; persons of Arab or “Middle Eastern” appearance or ancestry; persons with accents; gay, lesbian, bisexual, and transgendered persons; persons with appearance stigmas such as obesity; persons with different shades of skin color; and persons with differing marital status or child-rearing responsibilities.

## Conclusion

By addressing these and other important questions, experience in the USA suggests that situation testing can make substantial contributions in addressing employment discrimination in industrial nations outside the USA. Of course, specific techniques for conducting tests have to be adapted to the labor market conditions in each nation. However, the experience of the ILO in its series of testing studies involving migrant workers in multiple countries (Bovenkerk, 1992) demonstrates that the adaptations are easily managed, and the resulting data are rich and provocative. Such studies may be particularly useful in nations where, unlike the USA, legal provisions present recording information such as race in employment records or national censuses, thereby limiting many non-testing methods for measuring employment discrimination.

---

high priority. However, a countervailing practical concern is that studies often receive particularly extensive attention from the news media, employers, and political leaders in the locality where they were conducted. Thus, such studies might represent a useful investment if the primary objective is to influence public opinion and local policy-makers.

(12) For example, based on non-testing research, Blumrosen, Bendick, Miller, and Blumrosen (1999), p. 22, identified grocery retailers, restaurants, department stores, and hospitals as high priority targets for anti-discrimination enforcement based on race.

A motto taught for decades to university students of business management in the USA is, "Measurement is the cornerstone of management" (Drucker, 1973). For nations grappling with employment discrimination, situation testing offers a cornerstone for measurement.

## References

- Arrow, K. (1998). "What has economics to say about racial discrimination?", *Journal of Economic Perspectives*, 12, p. 91-100.
- Bendick, Jr., M. (1999). "Adding testing to the nation's portfolio of information on employment discrimination". In M. Fix and M. Turner (Eds.) *A national report card on discrimination: The role of testing*. Washington: The Urban Institute, p. 47-68.
- Bendick, Jr., M. (1989). *Auditing race discrimination in hiring: A research design*. Washington: Bendick and Egan Economic Consultants, Inc. for The Urban Institute.
- Bendick, Jr., M. (1996). *Discrimination against racial/ethnic minorities in access to employment in the United States, Empirical results from situation testing*. International Migration Paper 12. Geneva: International Labour Office.
- Bendick, Jr., M. (1995). "Research evidence on racial/ethnic discrimination and affirmative action in employment". In *Discrimination and affirmative action, Are there any facts out there?* Sacramento: Assembly Judiciary Committee, California State Legislature, p. A58-A78.
- Bendick, Jr., M., and Nunes, A. (2005) "Behavioral Science Perspectives on Discrimination and their Implications for Testing", *Presentation to the Monash University Conference on Field Experiments on Discrimination in Markets*, Prato, Italy.
- Bendick, Marc, Jr., Brown, L., and Wall, K. (1999). "No foot in the door, An experimental study of employment discrimination against older workers", *Journal of Aging and Social Policy*, 10, p. 5-23.
- Bendick, Jr., M., Jackson, C., and Reinoso, V. (1994). "Measuring employment discrimination through controlled experiments", *Review of Black Political Economy*, 23, p. 25-48.
- Bendick, Jr., Marc, Jr., Jackson, C., Reinoso, V., and Hodges, L. (1991). "Discrimination against Latino job applicants: A controlled experiment", *Human Resource Management*, 30, p. 469-484.
- Bendick, Jr., M., Jackson, C., and Romero, J. (1996). "Employment discrimination against older workers: An experimental study of hiring practices", *Journal of Aging and Social Policy*, 8, p. 25-46.
- Bertrand, M., and Mullainathan, S. (2004). "Are Emily and Brendan more employable than Lakisha and Jamal? A field experiment on labor market discrimination", *American Economic Review* 113, p. 991-1011.
- Bielby, W. (2000). "Minimizing workplace gender and racial bias", *Contemporary Sociology* 29, p. 120-129.
- Blumrosen, A., Bendick, Jr., M., Miller, J., and Blumrosen, R. (1999). *Employment discrimination against women and minorities in Georgia*. Rutgers, NJ: Rutgers University Law School.

- Boggs, R., Sellers, J., and Bendick, Jr., M. (1993). "Use of testing in civil rights enforcement". In Fix, M. and Struyk, R. (Eds.), *Clear and convincing evidence: Measurement of discrimination in America*. Washington: Urban Institute Press, p. 345-376.
- Bovenkerk, F. (1992). *Testing discrimination in natural experiments, A manual for international comparative research on discrimination on the grounds of "race" and ethnic origin*. Geneva: International Labour Office.
- Cose, Ellis. *The Rage of a privileged Class*. (1993). New York: Harper Perennial.
- Cross, H. G., Kennedy, J., Mell, J., and W. Zimmerman (1990). *Employer hiring practices: Differential treatment of Hispanic and Anglo job applicants*. Washington: The Urban Institute.
- De Prada, M., Actis, W., Pereda, C., and Perez Molina, R. (1996). *Labour market discrimination against migrant workers in Spain*. International Migration Papers 9. Geneva. International Labour Office.
- Drucker, P. (1973). "Managing the public service institution", *The Public Interest* 33, p. 43-60.
- Egan, M., and Bendick, M., Jr. (2003). "Workforce diversity initiatives of U.S. multinational corporations in Europe", *Thunderbird International Business Review* 45, p. 701- 727.
- Egan, M., Mauleon, F., Wolff, D., and Bendick, Jr., M. (2007). "France's 'triple bottom line' reporting, An informational regulation approach to sustainable development", *Presentation to the Second International Conference of the International Center for Corporate Accountability, New York*.
- Feagin, J. and Sikes, M. 1994. *Living with racism, The Black middle class experience*. Boston: Beacon.
- Fix, M., and Struyk, M. (eds). (1993). *Clear and convincing evidence, Measurement of discrimination in America*. Washington: Urban Institute Press.
- Fix, M. and Turner, M. (Eds.) (1999). *A national report card on discrimination: The role of testing*. Washington: Urban Institute Press.
- Fry, E. (1986). *An equal chance for disabled people, A study of discrimination in employment*. London: The Spastics Society.
- Gaertner, S., and Dovidio, J. (1986). "The aversive form of racism". In Dovidio, J., and Gaertner, S. (Eds.), *Prejudice, discrimination, and racism*. San Diego, CA: Academic Press, p. 61-89.
- Goldberg, A., Mourinho, D., & Kulke, U. (1996). *Labour market discrimination against foreign workers in Germany*. International Migration Papers 7. Geneva. International Labour Office.
- Graham, P., Jordan, A., and Lamb, B. (1990). *An equal chance? Or no chance? A study of discrimination against disabled people in the labour market*. London. The Spastics Society.
- Greenwald, A., McGeer, D., and Schwartz, J. (1998). "Measuring individual differences in implicit cognition, The implicit association test," *Journal of Personality and Social Psychology*, 74, p. 1464 – 1480.
- Gringart, E. and Helmes, E. (2001). "Age discrimination in hiring practices against older adults in Western Australia, The case of accounting assistants", *Australasian Journal of Ageing*, 20, 23-8.
- Heckman, J. (1998). "Detecting discrimination", *Journal of Economic Perspectives*, 12, p. 101-116.
- Henry, F., and Ginzberg, E. (1985). *Who gets the work? A test of racial discrimination in employment*. Toronto. The Urban Alliance on Race Relations.
- Hilton, J.L., and von Hippel, W. (1996). "Stereotypes", *Annual Review of Psychology*, 47, p. 237-71.

James, F., and DelCastillo, S. (1992). "Measuring job discrimination, Hopeful evidence from recent audits", *Harvard Journal of African American Public Policy*, 1, p. 35-53.

Kenney, G., and Wissoker, D. (1994). "An analysis of the correlates of discrimination facing young Hispanic job-seekers", *American Economic Review*, 84, p. 674-83.

Kerner, O. (1968). *Report of the national advisory commission on civil disorders*. Washington: US Government Printing Office.

Kochan, T. Bezrukova, K., Ely, R., Jackson, S., Joshi, A., Jehn, K., Leonard, J., Levine, D., & Thomas, D. (2003). "The effects of diversity on business performance: Report of the diversity research network", *Human Resource Management*, 42, p. 3-21.

Kopfman, J., Smith, S., Yun, J., and Hodges, A. (1998). "Affective and cognitive reactions to narrative versus statistical evidence on organ donation", *Journal of Applied Communication Research*, 26, p. 279-300.

Landever, M. (1993). "Tester standing in employment discrimination cases under 42 USC 1981". *Cleveland State University Law Review*, 41, p. 381.

Leonard, J. (1985). "What promises are worth, The impact of affirmative action goals", *Journal of Human Resources*, 20, p. 3-20.

Lodder, L. (1994). *Employment testing for civil rights enforcement, An operations manual*. Chicago: Legal Assistance Foundation of Chicago.

Lodder, L., McFarland, S., & White, D. (2003). *Racial preferences and suburban employment opportunities*. Chicago: Legal Assistance Foundation of Metropolitan Chicago.

Martel, R., Lane, D., & Emrich, C. 1996. "Male-Female Differences, A computer simulation", *American Psychologist*, 51, p. 157-158.

Mazur, A. (1985). "A bio-social model of status in face-to-face primate groups," *Social Forces*, 64, p. 377-402.

MCAD (1994). *Age testing report*. Boston: Commonwealth of Massachusetts Commission Against Discrimination.

Moss, P., and Tilly, C. (2002). *Stories employers tell, Race, skill, and hiring in America*. New York: Russell Sage Foundation.

Neal, D. and Johnson, W. (1996). "The Role of Pre-market Factors in Black White Wage Differentials", *Journal of Political Economy*, 104, p. 869-895.

Neumark, D. 1996. "Sex discrimination in restaurant hiring: An audit study", *Quarterly Journal of Economics*, 20, p. 915-940.

Nunes, A., and Seligman, B. (2000). *A study of the treatment of female and male applicants by San Francisco Bay Area auto service shops*. Berkeley, CA: Discrimination Research Center of The Impact Fund.

Nunes, A. & Seligman, B. (1999). *Treatment of Caucasian and African-American applicants by San Francisco Bay Area employment agencies: Results of a study utilizing "testers."* Berkeley, CA: Discrimination Research Center of the Impact Fund.

Oh, A. Y.K. (1993). "Using employment testers to detect discrimination: An ethical and legal analysis", *Georgetown Journal of Legal Ethics* 7, p. 473.

Pager, D. (2003). "The mark of a criminal record", *American Journal of Sociology* 108, p. 937-975.

Pager, D., and Western, B. (2005), *Discrimination in low-wage labor markets, Results from an experimental audit study in New York City*, Presentation to the American Sociological Association Annual Meeting, Philadelphia.

Ravaud, J-F., Madiot, B., & Ville, I. (1992). "Discrimination towards disabled people seeking employment", *Social Science and Medicine*, 35, p. 951-958.

Riach, P., and Rich, J. (2002). "Field experiments in discrimination in the market place", *The Economic Journal*, 112, p. F480-F518.

Turner, M. A., Fix, M., and Struyk, R. (1991). *Opportunities Denied, Opportunities Diminished: Racial Discrimination in Hiring*. Washington: The Urban Institute.

Valian, V. (1998). *Why so slow? The advancement of women*. Cambridge, MA: MIT Press.

Whipple, C. (1990). "All American", *CBS News 60 Minutes*, broadcast February 11.

Weichelbaumer, D. (2000). "Is it sex or personality? The impact of sex-stereotypes on discrimination in applicant selection", *Eastern Economic Journal* 30, p. 159-186.

Wymer, J., and Sudbury, D. (1992). "Employment discrimination 'testers': Will your hiring practices 'pass'?", *Employee Relations Law Journal*, 17, 623-633.

Yelmolinsky, M. (1993). "Filling an enforcement void, Using testers to uncover and remedy discrimination in hiring for lower-skilled, entry-level jobs", *University of Michigan Journal of Law Reform*, 26, 404-459.