S. Michael Gaddis
Editor

Audit Studies: Behind the Scenes with Theory, Method, and Nuance

Chapter 2
Making It Count: Discrimination Auditing and the Activist Scholar Tradition

Frances Cherry and Marc Bendick Jr.

Abstract Discrimination auditing can usefully be viewed as part of a tradition of social science activist scholarship since World War II. This perspective suggests that the single-minded pursuit of methodological rigor, especially when reflected in exclusive reliance on documents-based audits, often sacrifices other characteristics historically associated with auditing's unique contributions to societal and scientific advancement. This chapter advocates and illustrates a balanced research agenda in which the most rigorous auditing studies are paralleled by others more directly in the activist scholar tradition. The hallmarks of that tradition are: in-person testers, the lived experience of discrimination, researcher-community partnerships, and goals beyond academic ones.

Keywords Situation testing · Participatory action research · Prejudice · Employment · Community organizing

2.1 Introduction

> The next day, Dorothy parked her 1946 Plymouth on Palmerston Boulevard. As she walked with Langston up the steps to the house, Dorothy noticed the red and white For Rent sign still on the door….
>
> "Not a good sign," Langston said. He rang the bell. Watson opened the door and stepped out onto the porch….
>
> "Well, we're here," Dorothy said. "We'd like to sign the contract, pay you, and bring our things in from the car."
>
> Langston watched the man open his mouth, close it, stop, pause.… Langston instantly knew that they would not get the flat. The coming refusal was as certain as the sunset – but Langston sensed that it would come in a distinct way….

F. Cherry (✉)
Department of Psychology, Carleton University, Ottawa, Canada
e-mail: francherry@cunet.carleton.ca

M. Bendick Jr.
Bendick and Egan Economic Consultants, Inc., Alexandria, VA, USA
e-mail: marc@bendickegan.com

S. M. Gaddis (ed.), *Audit Studies: Behind the Scenes with Theory, Method, and Nuance*, Methodos Series 14, https://doi.org/10.1007/978-3-319-71153-9_2

> "I'm so sorry," Watson said, looking only at Dorothy, "and I hope you haven't been overly inconvenienced, but I have made other arrangements. A retired couple came by yesterday, after you left. They needed a quiet place, and they were prepared to take out a two-year lease, and I'm sorry, but I couldn't refuse them."
>
> "Yes, you could have," Dorothy shot back.
>
> Lawrence Hill, *Any Known Blood*

This passage by novelist Lawrence Hill fictionalizes his parents' experience as an interracial couple in post-World War II Toronto. His father, Daniel Hill, was completing his doctorate in sociology, after which he became the first full-time Director of the Ontario Human Rights Commission in 1962. His mother, Donna Bender Hill, worked for the Toronto Labour Committee for Human Rights documenting discrimination in employment, housing, and restaurants to promote anti-discrimination legislation. They thus simultaneously experienced and studied the segregated, racialized daily life of North America in the 1950s.

Any Known Blood goes on to describe auditing in the Hills' response to this encounter. They persuade a white couple to apply for the apartment they had been refused, and the landlord promptly offers the property to the new applicants. The landlord reassures the white couple that he knows of no black neighbors and that he would "draw the line there."

Writ large, such grassroots responses emerged as the American Civil Rights Movement of the 1950s and 1960s. Ordinary persons' lived experience of discrimination in real situations was a key resource mobilized by that movement to recruit activists, sway public opinion, secure anti-discrimination laws, and support their enforcement. Auditing was one of the alliances between civil rights advocates and professional researchers generating that resource.

This chapter examines the emergence, growth, and evolution of those efforts from the end of World War II through the present. It describes the work of "activist scholars" (Cherry 2004, 2008; Cherry and Borshuk 1998; Torre and Fine 2011; Torre et al. 2012) from multiple academic disciplines and their partnerships with a range of community members and advocacy organizations.

This history provides important guidance for today. As other chapters in this volume document, discrimination auditing in the Twenty-First Century often embodies considerable methodological rigor, especially when documents-based audits are conducted rather than audits involving live testers (Crabtree 2018; Gaddis 2018; Lahey and Beasley 2018). The creative search for rigor has undoubtedly enhanced the method's credibility and power in some ways. However, single-minded pursuit of rigor risks sacrificing other considerations historically associated with auditing's unique contributions to both society and science. This chapter calls for a more balanced research agenda in which the most rigorous auditing studies are paralleled by others more directly in the tradition of their historical precedents. The hallmarks of that tradition are: in-person testers, the lived experience of discrimination, researcher-community partnerships, and goals beyond academic publication.

2.2 Auditing in the Era of Gradualist Persuasion

The history of scholar activism underlying this call begins around the end of World War II. Throughout the late 1940s and 1950s, literally hundreds of civic, religious and educational organizations sought to improve inter-racial and inter-religious relations in the United States (Giles and Van Til 1946; Watson 1947; Williams 1947). These groups included state and local race relations committees (e.g., mayors' unity committees and state fair housing commissions), national race relations organizations (e.g., the National Association for the Advancement of Colored People, NAACP, and the National Urban League), faith-based organizations (e.g., the American Friends Service Committee and the National Conference of Christians and Jews), and educational institutions (e.g., the Bureau of Intercultural Education). President Truman's Committee on Civil Rights (President's Committee on Civil Rights 1947) was a national-level instance of the same approach.

Typically lacking legal enforcement powers and even statutes making discrimination illegal, these organizations relied primarily on persuasion and voluntary cooperation to advance their objectives. Information on the prevalence of discrimination and its adverse consequences was often their primary resource (Biondi 2003; Gordon 2015; Jackson Jr 1998, 2001; Jackson 1990; Richards 1997).

Three organizations – the Society for the Psychological Study of Social Issues (SPSSI), the NAACP, and the Commission of Community Interrelations (CCI) of the American Jewish Congress – were particularly prominent in connecting researchers and activists to generate that information (Cherry and Borshuk 1998; Jackson Jr 2001; Richards 1997). Advocating such linkages, a 1947 monograph for the Social Science Research Council by Cornell sociologist Robin Williams stated that the "necessary inclusion of fact finding among techniques of action suggests that research must be seen as an integral part of inter-group relations. Scientific study is a form of social action" (Williams 1947, p. 25). That same year, in book sponsored by the American Jewish Congress, psychologist Goodwin Watson cited the mantra of prominent social psychologist Kurt Lewin "No action without research; no research without action." Watson proposed university "action research service bureaus," particularly to evaluate the effectiveness of different approaches to reducing racial and religious discrimination (Watson 1947, p. 151).

At that time, Fisk University was already a center of scholarship embodying this approach. Envisioning fact finding as the basis for educational and legislative efforts, renowned sociologist Charles Johnson worked with Fisk's Department of Race Relations, which had been established in 1942 by the American Missionary Society as an action arm of Fisk's Social Sciences Department. Over many years, Johnson and his colleagues developed community self-surveys in which in-person and telephone interviews as well as other techniques were used to document race relations in multiple localities across the United States (Gilpin and Gasman 2003). Communities themselves collected the data, which then were analyzed by Fisk and published in the

form of statistical reports. For example, Sanders (2001) describes a 2-year community self-survey in Burlington Iowa conducted from 1949 through 1951.

A second center of scholarship in this tradition was the Commission on Community Interrelations (CCI) of the American Jewish Congress. Established by Kurt Lewin at the Massachusetts Institute of Technology, CCI engaged in various types of action research from the late 1940s through the early 1950s. The CCI approach moved even further than Fisk's in emphasizing community control over researcher leadership (Cherry and Borshuk 1998).

In particular, CCI researcher Claire Selltiz, working with housing activist Margot Haas Wormser, promoted researcher-community partnerships in which "citizens of a community are responsible for and participate in every phase of the investigation" (Wormser and Selltiz 1951). Their approach to community self-surveys reflected a concept enunciated by Kurt Lewin that local residents know best what would work for their community and how best to produce required changes.

Additionally, social psychologists at that time were beginning to argue that personal contact among individuals from different ethnic groups working together toward a common goal would itself importantly transform individual attitudes and behavior (Allport 1954). In particular, through collaborative efforts of Blacks and Whites studying their own communities, Whites would gain an understanding of the perspective of Black persons from whom they were normally segregated (Sanders 2001).

Auditing was one of several fact-gathering techniques promoted by Wormser and Selltiz as a logical outgrowth of this belief in the personally transformative experiences of people of different backgrounds working together. Structured audits in community self-surveys during the 1940s and 1950s can in part be thought of as formalizing the comparisons emerging naturally when members of different ethnic or religious groups examine prejudice and discrimination while sitting side by side.

In 1951, Wormser and Selltiz, produced a manual titled *How to Conduct a Self-Survey of Civil Rights* discussing participatory self-studies of housing, education, and public facilities and services (Wormser and Selltiz 1951). The manual's stated goal was to empower community groups to gather credible information in their own localities with only limited assistance from outside consultants. The manual included "test cases" – paired-comparison audits – as one method of data gathering.

Wormser and Selltiz played that consulting role in a pilot project in a small, highly segregated New Jersey town anonymized as "Northtown." They recruited local community organizations, both minority-based and not, to form a sponsoring committee. CCI staff provided technical advice on survey methods, trained volunteer interviewers, and participated in data interpretation. However, the local committee determined the effort's scope and style.

The Northtown sponsoring committee agreed to many of CCI's proposed approaches including random sampling, parallel interviews with both minority and non-minority individuals, and publication of findings with an action program based on those findings. However, it declined to implement other elements in the manual, including "test cases." The committee could not agree whether fact-finding should include observational procedures or only surveys and interviews; whether data should be used only in education and persuasion or put to more legal and confron-

tational uses; and the extent to which local employers should be publicly embarrassed. Moreover, during the McCarthyite anti-communist hysteria of the 1950s (Schrecker 1986), community activism was generally suspect, and rumors circulated that the Northtown study was the work of Communists. Within an already-controversial undertaking, "test cases" did not command consensus support.

Selltiz had more success in including "test cases" in a project of the Committee on Civil Rights of East Manhattan (CCREM), a group organized to address potential mistreatment of diplomats of color assigned to the new United Nations headquarters (Selltiz 1955; Biondi 2003). Along with a number of colleagues formerly at MIT, Selltiz was by then affiliated with the Research Center for Human Relations at New York University.

CCREM's first project was an audit study of Manhattan restaurants, for which Selltiz developed a study design, trained testers, and analyzed data. Pilot field work was conducted as thesis research by two Columbia University social work students (Landa and Littman 1950), and after methodological refinements based on that pilot, a full-scale study was conducted on 62 Manhattan restaurants in June of 1950. The study found no instances in which testers of color were refused service. However, differential treatment was documented in seating them in undesirable locations and providing poorer service (Schuman et al. 1983).

These findings were then moved into action, primarily in ways Selltiz (1955) described as "educational" and "persuasive" rather than "militant." CCREM representatives met with associations of restaurant owners and unions of restaurant employees seeking their pledge of equal treatment for all patrons. Letters requesting the same pledge were sent to the owners and managers of restaurants throughout the neighborhoods from which the audited restaurants had been sampled. Although no individual restaurant was publicly identified as having discriminated, private meetings were held with managers of some of those establishments. The Committee also issued a press release citing its findings and the restaurant industry's pledge. Members of the Committee were interviewed on radio, and 10,000 copies of a pamphlet, "Have You Heard What's Cooking?" were distributed. A follow-up audit study in 1952 found that discrimination had decreased significantly. However, it was not clear if the work of CCREM was the cause, as just before the new study, New York State legislation instituted more effective enforcement against discrimination in restaurants and other public accommodations (Selltiz, p. 25).

The relationships that communities formed with activist scholars at Fisk and CCI were not unique (Lambert and Cohen 1949; Torre and Fine 2011; Greenberg 1997). However, many other partnerships were short-lived and not well documented. Nevertheless, recognizing such activities, community self-studies and "test cases" were allocated a full chapter in the first edition of *Research Methods in Social Relations,* a textbook widely adopted in social science classes throughout the 1950s. That chapter described "test cases" as "the most direct method of getting information about possible discriminatory practices" and producing "evidence so clear-cut that it cannot be doubted." However, the chapter also noted that committees sponsoring community self-studies "represent a cross section of the community and are likely to be rather cautious" about approving "staged tests" that may

be seen as "dishonest" (Jahoda et al. 1951, pp. 621–622; see also Cherry 1995, and Torre and Fine 2011).

Towards the end of the 1950s, scholars' participation in action research waned. The anti-communist chill of cold war politics continued to reach onto college campuses to render academics' involvement in progressive political causes suspect (Schrecker 1986). Concurrently, some of the most prominent academic advocates of scholar-community partnerships were no longer active; for example, Kurt Lewin died in 1947, and Marie Jahoda moved to England.

Most importantly, academically prestigious, "cutting edge" attention in multiple social science disciplines was turning in other directions. In psychology, social psychologists increasingly defined their field as an experimental science marked by separation of research and application, experimenter and subject, and laboratories and real communities (Cherry 2009; Collier et al. 1991). In economics, the "institutionalist" tradition of researchers' involvement in practical issues such as development of the Social Security Act became less prestigious than analyses of economic behavior through mathematical models (Hodgson 2003). In sociology, academic interest in concrete interactions of individuals in real world situations was largely displaced by more abstract modeling of social structures and their functions (Fine 1995). These shifts were reflected in the evolution of *Research Methods in Social Relations*. By the 1959 edition of this textbook (Selltiz et al. 1959), the chapter on community self-surveys, including "test cases," was no longer included.

2.3 Auditing in the Era of Civil Rights Laws

During the 1960s, the Civil Rights Movement resulted in major federal legislation, prominently including the Equal Pay Act of 1963, Civil Rights Act of 1964, the Voting Right Act of 1965, and the Fair Housing Act of 1968. In concert with major Supreme Court decisions including *Brown v. Board of Education*, these developments collectively decimated *de jure* segregation across the southern states (Branch 1988).

With their counterpart statutes in many states and localities, these federal laws also prohibited *de facto* discrimination embodied in social custom rather than law and prevailing nation-wide rather than primarily in the South. These discriminatory practices were commonly embedded in the routine behavior of non-minority individuals and institutions, and the need to document that behavior sparked renewed interest in auditing. Housing was the first policy area in which auditing became central.

Fair housing committees of concerned citizens had existed in many localities across the nation since the 1940s or earlier, and many of them had applied auditing in investigating individual complaints against property owners, rental agents, real estate developers, mortgage lenders, and others (Yinger 1995, p. 28). In 1977, the U.S. Department of Housing and Urban Development (HUD) dramatically scaled up and systematized these local efforts by sponsoring a Housing Market Practices

Survey estimating the prevalence of these problems nation-wide, primarily with respect to African Americans. That project's 3,264 tests in 40 metropolitan areas documented widespread discrimination in both rental and owner-occupied housing (Wienk et al. 1979).

During the 1980s, auditing of housing discrimination matured into a sustainable activist scholar practice. In 1982, the U.S. Supreme Court's decision in *Havens Realty Corp. v. Coleman* unanimously upheld the standing of testers and the fair housing organizations employing them to bring litigation under the 1968 Fair Housing Act (Boggs et al. 1993, p. 346). A National Fair Housing Alliance was formed in 1988 and soon acquired 90 non-profit member organizations. At least 72 studies were conducted in individual cities, and in 1889, HUD sponsored a second nation-wide study, this time documenting the experiences of African Americans and Hispanics in 25 metropolitan areas (Yinger 1986, 1998). Findings from that HUD-sponsored study were credited with a major role in shaping the 1988 Amendments to the Fair Housing Act (Yinger 1998, p. 28).

Concurrently, auditing continued to be applied sporadically to discrimination in other aspects of daily life. Within housing, the initial focus on landlords' and owners' willingness to rent or sell individual properties broadened to examine "redlining" of neighborhoods in mortgage lending and homeowner insurance (Galster 1993). In 1988, social scientists from Howard University joined with a coalition of churches in Washington, DC to audit discrimination by taxicabs against African American riders and riders going to predominantly-African American neighborhoods. This project resulted in successful litigation against three cab companies and the important precedent of holding cab companies liable for discriminatory acts by individual drivers (Boggs et al. 1993, p. 348). A scholar at the Yale Law School used auditing to document race and gender discrimination in retail car sales (Ayres 1991). Psychologists used auditing procedures to measure the extent to which random samples of shoppers, motorists, and subway riders would assist strangers of different races (Cosby et al. 1980).

Such developments caught the attention of James Gibson, then head of the Equal Opportunity Program at the Rockefeller Foundation. Over the 1980s, Gibson had become increasingly concerned about erosion of public concern about racial discrimination throughout American society. Federal enforcement of anti-discrimination laws had substantially weakened since the 1981 advent of the conservative Reagan Administration (Clark 1989). Several Supreme Court decisions – notably *Regents of the University of California v. Bakke* in 1978 and *City of Richmond v. Croson* in 1989 – signaled increasing judicial skepticism of race-based affirmative action in education, employment, and government contracting. Perhaps most troubling, public opinion polls were reporting that increasing proportions of the non-minority U.S. population considered discrimination merely a problem of the past (Bendick 1999, p. 54). Gibson reasoned that auditing's success in housing might be expanded to provide fresh momentum to the flagging anti-discrimination cause.

Gibson translated this aspiration into substantial, multi-year grants to two non-profit organizations, The Urban Institute and the Washington Lawyers' Committee for Civil Rights and Urban Affairs. With respect to employment, these seminal grants

subsequently generated: a design for applying auditing to hiring (Bendick 1989); a study of employers' treatment of Hispanics under the federal Immigration Reform and Control Act of 1986 (Cross 1990); studies of hiring discrimination against African Americans (Turner et al. 1991; Bendick et al. 1994), Hispanics (Bendick et al. 1991), and older workers (Bendick et al. 1999); two successful testing-based lawsuits, one based on race and the other on gender (Boggs et al. 1993, pp. 362–363); a workshop training 72 academics and advocates on employment auditing (FEC 1993); and Congressional and state legislative testimony on the continued prevalence of discrimination and the continuing need for affirmative action (e.g., Bendick 1995). Promoting applications to other fields, the grants also supported two books (Fix and Struyk 1993; Fix and Turner 1999) offering creative ideas for auditing in retail sales, business lending, government contracting, and health care.

2.4 Audits Combining Rigor and Relevance

As other chapters in this volume reflect, over the two decades since these events, auditing research has expanded steadily, and its methodological sophistication has increased markedly (e.g., Edelman et al. 2017). These developments undoubtedly contribute to the method's credibility and influence today. But concurrently, some characteristics historically associated with auditing's unique contributions to societal and scientific advancement have become de-emphasized. In particular, four characteristics prominent in auditing prior to 2000 are relatively neglected today.

Human Testers One of these increasingly rare characteristics is audits employing live human testers rather than "correspondence studies" or "document studies" in which "testers" are presented only through written or electronic documents such as job resumes or mortgage applications. Because human testers are time-consuming and expensive to recruit and field, live tester studies tend to have samples on the order of 100 or fewer completed tests. In contrast, document audits, especially those in which application documents are computer generated, can afford samples of thousands. Larger samples increase the probability of observing statistically significant results and allow variations within the study design to examine multiple hypotheses and complex interactions. In addition, documents can be more rigorously controlled than the individual personality and appearance of live testers permit. Document studies thereby sidestep the inevitable skeptical questions about whether the tester within each pair who was treated less favorably was actually less qualified in some subtle, undocumented way (e.g., Heckman 1998).

These characteristics tend to make document studies easier to publish in scholarly outlets and more prestigious by conventional academic standards. However, their narrowness may limit the real-world applicability of their findings. Social psychologists have written extensively about the trade-off of experimental control and relevance since the late 1960s, when laboratory experimentation clearly became the preferred methodology in their field (Aronson and Carlsmith 1968;

Elms 1975; Cherry 2009). Parallel discussions can be found in other social science disciplines as well.

Have rigorously-controlled, document-based audit studies rendered live human audits obsolete? Are document-based audits the preferred approach when many selection processes today – such as job applications, college admissions, and loan applications – are commonly conducted at least partly on-line? The history of auditing suggests otherwise.

The most obvious reason for conducting in-person audits is that document studies typically cover only the initial stage of a selection process – for example, an employer's decision concerning which job applicants to invite to face-to-face interviews. Studies in which live testers have pursued the selection process all the way to the end often document discriminatory behavior appearing only at late stages – for example, where employers feel constrained by legal or social pressure to interview a racially-diverse slate of job candidates but then offer positions only to non-minority applicants (e.g., Bendick et al. 1994). In such circumstances, document audits of only the initial stages systematically under-estimate the overall prevalence of discrimination.

Correct measurement of outcomes is not the only benefit of deploying human testers. Contemporary concepts such as "implicit bias" and "micro-inequities" make clear that much discrimination today is unconscious and subtle (Jones et al. 2014). Especially when audio or video recordings provide word-for-word transcripts, in-person audits can illuminate the details of screening processes where such problems often lurk – for example, by documenting the influence of stereotypes on interviewers' judgments about job seekers' qualifications or the influence of in-group bias on interviewers' informal provision of assistance and encouragement to job seekers (Bendick and Nunes 2012).

The Lived Experience of Discrimination A second characteristic of historic auditing that is relatively rare today is efforts to communicate how discrimination feels to those experiencing it. Mainstream social and behavioral science research has tended to focus on the attitudes and behavior of the perpetrators of prejudice and discrimination rather than their targets. To be sure, such research is valuable in developing procedures for reducing harmful behavior. However, studying the other side of interactions between discriminators and their targets is also important for understanding the ways in which the targets – and the broader society – are harmed (e.g., Steeler 1997; Swim and Stangor 1998; Bendick and Nunes 2012).

The increasing visibility and power of community-based social movements – in forms such as marches and demonstrations advocating rights for women, visible minorities, person with disabilities, or based on sexual orientation or gender identity – often give a loud but unsystematic voice to persons experiencing discrimination in their daily lives. The community self-studies described earlier in this chapter began the process of systematically studying these perspectives by conducting extensive community interviews prior to developing their formal surveys (Wormser and Selltiz 1951). Live tester auditing studies tend to make these experiences even more central by collecting and disseminating detailed narratives of testers' actual experiences.

These narratives are particularly powerful in influencing the attitudes and behavior of individuals who have not personally experienced discrimination. As was discussed earlier in this chapter, when "test case" auditing started as part of community self-surveys, personal exposure to discrimination was intended to build understanding and empathy among minority and non-minority testers, members of local committees, and local residents. When auditing is part of lawsuits enforcing anti-discrimination law, the personal testimony of testers is often crucial in convincing judges and juries. And when audit findings have been presented to public policymakers – for example, in testimony to state or federal legislators – vivid anecdotes of testers' personal experiences tend to catch legislators' and media attention; as skilled public speakers know, human interest stories tend to be influential in ways that statistics alone are not (Bendick and Nunes 2012). Because auditing at its best provides *both stories and statistics*, it can uniquely retain the accuracy of the latter while mobilizing the persuasive power of the former.

The dominance of document-based auditing today tends to deprive audit studies of some of their most potentially influential findings. In addition, many audit studies are conducted with only academic peers as their target audience and incorporate few efforts to disseminate the findings more broadly. Career incentives in academia typically provide little credit for participating in community meetings, drafting pamphlets for distribution to consumers, or engaging with local news media. The earlier generations of activist scholars made such activities integral to their auditing research.

Community Partners A third characteristic of earlier auditing that is relatively rare today is partnerships between scholars and community groups.

In the early days of discrimination auditing, scholar-community partnerships were not formed merely for pragmatic reasons such community groups' need for trained researchers to analyze data or researchers' need to recruit community members as testers. Instead, both parties saw auditing as an important process of personal and organizational growth for the community groups, a strengthener of activist alliances through newly-shared perceptions and the team-building experience of working together. Promoting these processes was as integral an objective of the activity as were published reports.

Concurrently, working with a non-academic partner inevitably influences researchers as well. Interacting with individuals personally affected by discrimination often provides researchers with new insights into how institutions operate and new hypotheses to be studied. In addition, as the history in this chapter illustrates, community partners often shape studies in directions that researchers themselves would not necessarily have thought to initiate.

Decisions about paired versus unpaired audits provide an example of differences in priorities between researchers and community partners. The design of many audit studies today involves sending applications for the protected group tester and the control tester to different recipients, such as different employers or different mortgage lenders. The resulting response to the two groups of applications are then compared statistically to calculate overall rates of differences in treatment and to analyze the correlates of those differences. Researchers typically find this procedure

attractively efficient. However, when applications are presented to different recipients, it is not possible to identify individual decision-makers, such as employers or mortgage lenders, who have discriminated. The studies therefore document an abstract evil attributable only to the overall population from which the audit sample was drawn. Essentially, unpaired audits describe a *villainy without villains*. Community groups of adversely-affected individuals often want more specificity than that to facilitate concrete actions toward amelioration and often would not support such studies.

Objectives Beyond the Academic If properly conducted, unpaired auditing studies of the sort just described would be acceptable for scholarly publication. But that fact highlights an over-arching difference between many of today's auditing studies and auditing in the earlier activist scholar tradition. When Claire Selltiz and CCRM joined together to audit how restaurants treated customers of color, their goal was not to publish a study. Their goal was to reduce discriminatory behavior, with a study serving as an intermediate step. As was recounted earlier in this chapter, once the audit was completed, its findings were mobilized in multiple ways to promote behavioral change by restaurant owners, staff, and customers.

Activist scholars might or might not personally participate in such follow-up activities, but they must ensure that their studies are structured to support them. Invariably, this requirement means that audit studies take more time and resources. It often imposes study designs that might not be ideal from a pure research point of view, including use of live testers, having pairs of testers apply to the same company, and collecting narratives on testers' experiences in more detail than is statistically analyzable. In addition, researchers need to be prepared for the ego-crushing fact that the community groups see the audit as only a small piece of their long-term strategy.

2.5 Auditing Scholar-Activism Today

This section briefly sketches three examples of contemporary auditing scholar activism embodying most or all of the four characteristics just discussed.

Public Policy The most obvious example involves bringing audit evidence to bear on significant public policy issues.

In social policy today, few topics are more hotly debated than the complex relationships between race and the criminal justice system (Alexander 2012). One issue at the forefront of these discussions is so-called "ban the box" laws which limit employers' consideration of job applicants' criminal records in making hiring decisions. These laws have been adopted in 24 states and more than 150 localities, and have been considered in numerous additional jurisdictions (Rodriguez and Avery 2016).

When this issue is debated, the discussions almost inevitably cite the live-tester audit research of Harvard professor Devah Pager which examined the effect of criminal records on job applicants' chance of being hired and the interaction between

those effects and applicants' race (Pager 2007). A google search in November 2016 using the combined terms "ban the box" and "Devah Pager" produced more than 150 citations outside of her academic field of sociology, including by researchers in applied fields such as criminology, by news media, by advocacy organizations, and in public documents including hearings of the U.S. Congress and state legislatures and official EEOC guidelines for employers.

Pager herself has participated in these debates by, for example, serving on a government advisory panel of the National Academy of Sciences, writing opinion pieces in the news media, speaking at non-academic conferences of civil rights and criminal rights advocates, and testifying before public agencies such as the EEOC and the New York City Council (Pager 2016). However, the more fundamental way that her work reflects the activist scholar tradition is that the design of her audit studies provided information, both statistical and anecdotal, directly relevant to the "ban the box" issue. Had she not provided this information, more than a decade of policy debates would have been far less empirically grounded.

Legal Enforcement A second example involves using audit evidence in legal enforcement.

Make the Road New York (MTRNY) is a membership-based organization representing people of color in the New York City boroughs of Brooklyn, Queens, and Staten Island. After several of its transgender members complained about being turned down for jobs at fast food restaurants and a survey of their transgender members found that 59% of them reported similar experiences in a range of industries, the organization joined with economist Marc Bendick, Jr. to document the problem more systematically. They selected and trained two tester teams, one pairing a transgender and a cisgender woman and the other pairing a transgender and a cisgender man. Their resumes showed education and experience making them equally qualified for these positions. During the spring and summer of 2009, these teams applied for entry-level sales positions at 24 clothing retail stores in Manhattan. They found that, while transgender job applicants were often treated as politely as their testing partners, in some cases they were not, and the net rate of discrimination against the transgender testers was 42%.

The MTRNY report documenting this study (Bendick and Madar 2009) named the companies where discrimination had been encountered and included testers' "personal testimony" such as the following:

> When I went to apply at J Crew, I spoke to the manager, who said she was busy. I then spoke to a sales associate who gave me an application, but was vague about whether they were hiring. I filled out the application and submitted it to the manager then and there. She said she would give it to the hiring manager, and when I asked if they were currently hiring, she didn't say yes or no and said they would call me in for an interview. Twenty minutes later, my cismale partner went in, and…he ended up getting hired….I called twice over the next two weeks and they said they were still looking over applications and would call me. They never did.

> I was interviewed at a few of the stores that we tested. At some point during the interview, I would tell the employer that I was transgender and that my preferred pronoun was "he." In one interview, at DSW, I asked the manager whether I would feel comfortable working

> in the store as a transgender person and they said "that's up to you." I was also continually referred to as "she" despite my stated preference for the pronoun "he." Facing these kinds of experiences over and over again was humiliating. This process took an emotional toll on me.... Although this was a controlled research study, this experience mirrors my real life.

MTRNY then contacted the Civil Rights Bureau of the New York State Attorney General seeking relief under the anti-discrimination laws of New York City and New York State. It presented detailed documentation from audits of two employers, J. Crew and American Eagle Outfitters, each of which had been tested by both teams. The Attorney General found the evidence insufficient to proceed against J. Crew but opened an investigation of American Eagle. In May 2010, American Eagle agreed to a legal settlement that included adding gender identity and gender expression as a protected category in the company's anti-discrimination policy, training employees on transgender issues, training employees on how to file discrimination complaints, and revising the company dress code to no longer forbid men to wear women's clothing and men to wear women's clothing. In announcing this victory, MTRNY predicted that the settlement by such a prominent firm would encourage retailers nation-wide to rethink their policies and practices on the same issues (Taylor 2010).

Community Organizing A third example involves audit studies conducted as part of community organizing.

The Restaurant Opportunity Center United (ROC-U) is a non-profit worker center seeking to improve wages and working conditions for low-wage restaurant employees, including many people of color and recent immigrants. Starting in New York City with a core membership of restaurant workers who lost their jobs in the 2001 terrorist attack on the World Trade Center, the organization has developed into an influential presence throughout New York City's restaurant industry and then a multi-city group of similar organizations in a dozen states (Jayaraman 2013).

In achieving that growth, ROC-U has relied in part on a sequence of research activities first developed in New York and subsequently repeated in other cities. In New York, the sequence began with a research study in 2004 based on a structured survey of 530 workers, 45 semi-structured interviews with workers, and 35 semi-structured interviews with restaurant operators (ROC-NY 2005). Along with widespread issues of low wages, wage theft, and on-the-job harassment, this study highlighted occupational segregation in which immigrants and people of color were employed almost exclusively in low-paid positions in restaurant kitchens while white workers with similar skills, qualifications, and experience were overrepresented in server and manager jobs in the same restaurants' dining rooms. This pattern was especially stark in the city's upscale, "fine dining" establishments, where the earning opportunities were greatest.

To pursue that finding, ROC-NY joined with economist Marc Bendick, Jr. to conduct paired live-tester audits on the hiring of servers by fine dining restaurants in Manhattan (ROC-NY 2009; Bendick et al. 2010). During 2006 and 2007, the study recruited 37 volunteer testers primarily from among ROC members, partially

with the goal of actively engaging these individuals in the organization to strengthen their sense of affiliation. Testers were paired into teams of one white person and one person of color of the same gender, trained in effective interviewing techniques, provided with resumes showing equivalent qualifications, and assigned to apply for server positions at Manhattan's top restaurants, either responding to server openings advertised on-line or making "cold calls" at restaurants randomly selected from published lists of the city's most celebrated dining places. The study documented substantial differences in treatment adverse to persons of color at 31% of the restaurants audited and calculated that testers of color were only 55% as likely as their testing partners to receive job offers.

The study also illustrated employers' discriminatory behavior through narratives such as the following (Bendick et al. 2010, pp. 810–811):

> Answering a Craigslist advertisement, a white woman with no accent applied at an upscale Italian restaurant. She was promptly sent to an assistant manager, who, during an 18 minute interview, called her resume impressive, said that she presented herself well and that she'd "fit right in," and offered her specific work shifts. He emphasized that she would have opportunities to advance into management and that the restaurant would pay part of her health insurance. Meanwhile, a Chinese American woman with no accent, who had arrived half an hour before the white woman, was sent away with an interview appointment for the following day. During that interview, which lasted nine minutes, the same manager who had interviewed the white woman denied ever hearing of the restaurants on her resume and questioned whether she had worked in elegant establishments. He concluded that he would call her after consulting with other managers, but he never did.

The findings of this study were released at a well-attended "industry summit" hosted at his own restaurant by a celebrity chef who was a long-time ROC supporter. That release received considerable news media coverage, especially in the restaurant industry trade press, usually featuring vivid anecdotes that had been presented at the summit by testers themselves. The findings have been frequently cited by ROC throughout its organizing and lobbying activities in New York and elsewhere, whether appealing to potential worker members, restaurant operators, the news media, or public officials (ROC-U 2014). They are also reflected in ROC's list of "high road" restaurants they recommend to issue-conscious restaurant consumers (ROC-U 2016).

2.6 Barriers to Be Overcome

As multiple chapters in this book illustrate, the growing conceptual and methodological sophistication of much contemporary auditing research is impressive. However, the history reviewed in this chapter suggests that sophistication, experimental control, and academic credibility are often enhanced in ways that sacrifice other aspects of auditing that also contribute to the method's unique power. The history of scholar activism reviewed in this chapter suggests that such sacrifice is costly to both society and science. It is also often not necessary.

Of course, this history also documents that discrimination auditing in the activist scholar tradition is not without its challenges. Researchers often need to learn patience, a more user-friendly way of communicating, and sometimes humility to work effectively with non-academic partners whose perspectives and priorities often differ from their own. Incentives in the academic world tend to militate against overtly "applied" scholarship; involvement with community groups and in policy issues is sometimes viewed by academic colleagues, such as tenure committees, as a distraction at best and as indicating academic unworthiness at worst. Where activist audit studies achieve scholarly publication, that often occurs in interdisciplinary or second-tier journals that carry less academic prestige.

Furthermore, although some university Institutional Review Boards (IRBs) have enthusiastically endorsed audit studies, others have blocked them for political or other non-scientific reasons. One example is provided by Psychology Professor Jane Connor at the State University of New York at Binghamton (Connor 2000). In 1998, Connor taught a course on the Psychology of Racism in which her students watched a *Prime Time Live* television segment which followed two actors – John (White) and Glen (Black) – through a day of settling into a new city. Dressed similarly and coached to speak and behave in similar ways, the actors went apartment hunting, job seeking, and shopping while the video documented their strongly contrasting treatment.

Visible-minority students in Connor's class generally described the actors' experiences as similar to their own. In contrast, many non-minority students questioned whether, because the film has been made a decade earlier and in a different region of the United State and was made for television rather than as a scientific study, similar results would be found in their city. Accordingly, Connor organized a follow-up independent study course in which her students designed an in-person audit study of retailers in the Binghamton area. However, over the course of 2 years and multiple revisions, the proposal was never approved by the university's IRB. Connors' experiences offer a cautionary tale on the ethical and political controversies that auditing can trigger.

The twin goals of social progress and the advancement of human knowledge will be best served if the entire social science research community – funders, institutions, communities, and researchers themselves – invest creativity and effort in overcoming such obstacles and restoring the diversity of auditing activities more typical of auditing's earlier years. Without that breadth, auditing researchers may earn the title of scholars but not activist scholars, and all parties – activists, scholars, and society at large – will be the poorer for it.

References

Alexander, M. (2012). *The new Jim Crow, Mass incarceration in the age of colorblindness*. New York: New Press.

Allport, G. (1954). *The nature of prejudice*. Reading: Addison Wesley.

Aronson, E., & Carlsmith, J. M. (1968). Experimentation in social psychology. In G. Lidzey & E. Aronson (Eds.), *The handbook of social psychology* (Vol. 2, 2nd ed., pp. 1–79). Reading: Addison-Wesley.

Ayres, I. (1991). Fair driving: Gender and race discrimination in retail car negotiations. *Harvard Law Review, 104*(4), 817–872.

Bendick, M., Jr. (1989). *Testing race discrimination in hiring: A research design*. Washington, DC: Bendick and Egan Economic Consultants, Inc.

Bendick, M., Jr. (1995, May 4). *Research evidence on discrimination and affirmative action in employment*. Testimony, Committee on the Judiciary, California State Assembly.

Bendick, M., Jr. (1999). Adding testing to the nation's portfolio of information on employment discrimination. In M. Fix & M. Turner (Eds.), *A national report card on discrimination: The role of testing* (pp. 47–68). Washington, DC: The Urban Institute.

Bendick, M., Jr., & Madar, C. (2009). *Transgender need not apply: Gender identity job discrimination in New York City's retail sector*. New York: Make the Road New York.

Bendick, M., Jr., & Nunes, A. (2012). Developing the research basis for controlling bias in hiring. *Journal of Social Issues, 68*, 238–263.

Bendick, M., Jr., Jackson, C., Reinoso, V., & Hodges, L. (1991). Discrimination against Latino job applicants: A controlled experiment. *Human Resource Management, 30*, 469–484.

Bendick, M., Jr., Jackson, C., & Reinoso, V. (1994). Measuring employment discrimination through controlled experiments. *Review of Black Political Economy, 23*, 25–48.

Bendick, M., Jr., Brown, L., & Wall, K. (1999). No foot in the door: An experimental study of employment discrimination against older workers. *Journal of Aging & Social Policy, 10*(4), 5–23.

Bendick, M., Jr., Rodriguez, R., & Jayaraman, S. (2010). Employment discrimination in upscale restaurants: Evidence from paired comparison testing. *The Social Science Journal, 47*, 802–818.

Biondi, M. (2003). *To stand and fight: The struggle for civil rights in postwar New York City*. Cambridge: Harvard University Press.

Boggs, R., Sellers, J., Bendick, M., & Jr, M. (1993). Use of testing in civil rights enforcement. In M. Fix & R. Struyk (Eds.), *Clear and convincing evidence: Measurement of discrimination in America* (pp. 345–376). Washington, DC: Urban Institute Press.

Branch, T. (1988). *Parting the water, America in the King years 1954–1963*. New York: Simon & Schuster.

Cherry, F. (1995). *The "stubborn particulars" of social psychology: Essays on the research process*. London: Routledge.

Cherry, F. (2004). Kenneth B. Clark and social psychology's other history. In G. Philogene (Ed.), *Racial identity in context: The legacy of Kenneth B. Clark* (pp. 13–33). Washington, DC: American Psychological Association.

Cherry, F. (2008). Thomas F. Pettigrew: Building on the scholar/activist tradition in social psychology. In U. Wagner, L. T. Tropp, G. Finchilescu, & C. G. Tredoux (Eds.), *Improving intergroup relations: Building on the legacy of Thomas F. Pettigrew* (pp. 11–23). Oxford: Blackwell.

Cherry, F. (2009). Social psychology and social change. In D. Fox, I. Prilleltensky, & S. Austin (Eds.), *Critical psychology: An introduction* (2nd ed., pp. 93–109). Thousand Oaks: Sage.

Cherry, F., & Borshuk, C. (1998). Social action research and the Commission on Community Interrelations. *Journal of Social Issues, 54*, 119–142.

Clark, L. (1989). Ensuring equal employment opportunity through law. In D. L. Bawden (Ed.), *Rethinking employment policy* (pp. 155–167). Washington, DC: Urban Institute Press.

Collier, G., Minton, H. L., & Reynolds, G. (1991). *Currents of thought in American social psychology*. New York: Oxford University Press.

Connor, J. M. (2000). *Studying racial bias: Too hot to handle?* National Center for Case Study Teaching in Science. Case Collection http://sciencecases.lib.buffalo.edu/cs/collection/detail.asp?case_id=458&id=458

Cosby, F., Bromley, S., & Saxe, L. (1980). Recent unobtrusive studies of black and white discrimination and prejudice. *A literature Review, Psych Bulletin, 87*, 546–563.

Crabtree, C. (2018). An introduction to conducting email audit studies. In S. M. Gaddis (Ed.), *Audit studies: Behind the scenes with theory, method, and nuance*. Cham: Springer International Publishing.

Cross, H. (1990). *Employer hiring practices: Disparate treatment of Hispanic and Anglo job seekers*. Washington, DC: The Urban Institute.

Edelman, B., Luca, M., & Svirsky, D. (2017). Racial discrimination in the sharing economy: Evidence from a field experiment. *American Economic Journal: Applied Economics, 9*, 1–22.

Elms, A. C. (1975). The crisis in confidence in social psychology. *The American Psychologist, 30*, 967–976.

FEC. (1993). *Employment testing conference*. Washington, DC: Fair Employment Council of Greater Washington, Inc..

Fine, G. A. (Ed.). (1995). *A second Chicago School? The development of postwar American sociology*. Chicago: University of Chicago Press.

Fix, M., & Struyk, R. (Eds.). (1993). *Clear and convincing evidence, measurement of discrimination in America*. Washington, DC: The Urban Institute.

Fix, M., & Turner, M. A. (Eds.). (1999). *A national report card on discrimination in America: The role of testing*. Washington, DC: The Urban Institute.

Gaddis, S. M. (2018). An introduction to audit studies in the social sciences. In S. M. Gaddis (Ed.), *Audit studies: behind the scenes with theory, method, and nuance*. Cham: Springer International Publishing.

Galster, G. (1993). Use of testing in investigating discrimination in mortgage lending and insurance. In M. Fix & R. Struyk (Eds.), *Clear and convincing evidence: Measurement of discrimination in America* (pp. 287–334). Washington, DC: Urban Institute Press.

Giles, H. H., & Van Til, W. (1946). School and community projects. *The Annals of the American Academy of Political and Social Science, 244*, 34–41.

Gilpin, P. J., & Gasman, M. (2003). *Charles S. Johnson: Leadership beyond the veil in the age of Jim Crow*. Albany: State University of New York Press.

Gordon, L. N. (2015). The individual and "the general situation," The tension barometer and the race problem at the University of Chicago, 1947–1954. *Journal of the History of the Behavioral Sciences, 46*, 27–51.

Greenberg, C. (1997). Negotiating coalition: Black and Jewish civil rights agencies in the twentieth century. In M. Adams & J. Bracey (Eds.), *Strangers and neighbors: Relations between Blacks and Jews in the United States* (pp. 476–494). Amherst: University of Massachusetts Press.

Heckman, J. J. (1998). Detecting discrimination. *The Journal of Economic Perspectives, 12*, 101–116.

Hodgson, G. M. (2003). John R. Commons and the Foundations of Institutional Economics. *Journal of Economic Issues, 37*, 547–576.

Jackson, W. A. (1990). *Gunnar Myrdal and America's conscience: Social engineering and racial liberalism, 1938–1987*. Chapel Hill: University of North Carolina.

Jackson, J. P., Jr. (1998). Creating a consensus: Psychologist, the Supreme Court, and school desegregation, 1952–1955. *Journal of Social Issues, 54*, 143–177.

Jackson, J. P., Jr. (2001). *Social scientists for social justice: Making the case against segregation*. New York: New York University Press.

Jahoda, M., Deutsch, M., & Cook, S. (1951). *Research methods in social relations with especial reference to prejudice. Part Two: Selected techniques*. New York: The Dryden Press.

Jayaraman, S. (2013). *Behind the kitchen door*. Ithaca: Cornell University Press.

Jones, J. M., Dovidio, J. F., & Vietze, D. L. (2014). *The psychology of diversity*. Malden: Wiley Blackwell.

Lahey, J., & Beasley, R. (2018). Technical aspects of correspondence studies. In S. M. Gaddis (Ed.), *Audit studies: Behind the scenes with theory, method, and nuance*. Cham: Springer International Publishing.

Lambert, B. W., & Cohen, N. E. (1949). A comparison of different types of self-surveys. *Journal of Social Issues, 5*, 46–55.

Landa, P., & Littman, G. (1950). *A pilot study to test discriminatory practices against ethnic minority groups in public eating accommodations, An audit to determine the degree of discrimination practiced against Negroes in luncheonettes*. Unpublished thesis, New York School of Social Work, Columbia University.

Pager, D. (2007). *Marked: Race, crime, and finding work in an era of mass incarceration*. Chicago: University of Chicago Press.

Pager, D. (2016). *Curriculum vita*. Downloaded November 23, 2016 from www.scholar.harvard.edu

President's Committee on Civil Rights. (1947). *To secure these rights*. Washington, DC: Government Printing Office.

Richards, G. (1997). *"Race," racism, and psychology: Towards a reflexive history*. London: Routledge.

ROC-NY. (2005). *Behind kitchen doors: Pervasive inequality in New York City's thriving restaurant industry*. New York: Restaurant Opportunity Center of New York and the New York City Restaurant Industry Coalition.

ROC-NY. (2009). *The great service divide: Occupational segregation and inequality in the New York City restaurant industry*. New York: Restaurant Opportunity Center of New York and the New York City Restaurant Industry Coalition.

ROC-U. (2014). *Occupational segregation & inequality in the US restaurant industry*. New York: Restaurant Opportunity Center–United.

ROC-U. (2016). *Restaurant opportunities center united diners guide to ethical eating*. Downloaded October 17, 2016 from www.rocunited.org/dinersguide

Rodriguez, M. N., & Avery, B. (2016). *Ban the box, U.S. cities, counties, and states adopt fair hiring policies*. Downloaded November 23, 2016 from www.nelp.org

Sanders, K. M. (2001). The Burlington self-survey in human relations: Interracial efforts for constructive community change, 1949–1951. *The Annals of Iowa, 60*, 244–269.

Schrecker, E. (1986). *No ivory tower, McCarthyism and the universities*. New York: Oxford University Press.

Schuman, H., Singer, E., Donovan, R., & Selltiz, C. (1983). Discriminatory behavior in New York restaurants: 1950 and 1981. *Social Indicators Research, 13*, 69–83.

Selltiz, C. (1955). The use of survey methods in a citizens' campaign against discrimination. *Human Organization, 14*, 19–25.

Selltiz, C., Jahoda, M., Deutsch, M., & Cook, S. W. (1959). *Research methods in social relations [Rev]. Holt*. New York: Rinehart & Winston.

Steeler, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *The American Psychologist, 52*, 613–629.

Swim, J. K., & Stangor, C. (1998). *Prejudice: The target's perspective*. San Diego: Academic.

Taylor, S. (2010). *"Transgendering" personal appearance policies*. Downloaded November 22, 2016 from www.maketheroad.org

Torre, M. E., & Fine, M. (2011). A wrinkle in time: Tracing a legacy of public science through community self-surveys and participatory action research. *Journal of Social Issues, 67*, 106–121.

Torre, M. E., Fine, M., Stoudt, B. G., & Fox, M. (2012). Critical participatory action research as public science. In H. Cooper (Ed.), *APA handbook of research methods in psychology: Vol 2. Research designs* (pp. 171–184). Washington, DC: American Psychological Association.

Turner, M., Fix, M., & Struyk, R. (1991). *Opportunities denied, opportunities diminished: Racial discrimination in hiring*. Washington, DC: The Urban Institute.

Watson, G. (1947). *Action for unity*. New York: Harper & Bros.

Wienk, R. E., Reid, C. E., Simonson, J. C., & Eggers, F. J. (1979). *Measuring racial discrimination in American housing markets: The housing market practices survey*. Washington, DC: Department of Housing and Urban Development, Office of Policy Development and Research.

Williams, R. M., Jr. (1947). *The reduction of intergroup tensions*. New York: Social Science Research Council.

Wormser, M. H., & Selltiz, C. (1951). *How to conduct a community self-survey of civil rights*. New York: Association Press.

Yinger, J. (1986). Measuring discrimination with fair housing audits: Caught in the act. *The American Economic Review, 76*, 881–893.

Yinger, J. (1995). *Closed doors, opportunities lost: The continuing costs of housing discrimination*. New York: Russell Sage Foundation.

Yinger, J. (1998). Evidence of discrimination in consumer markets. *The Journal of Economic Perspectives, 12*, 23–40.