

## Evaluation Tradecraft

*Evaluation Tradecraft* presents short articles about the art of evaluation in housing and urban research. Through this department of *Cityscape*, the Office of Policy Development and Research presents developments in the art of evaluation that might not be described in detail in published evaluations. Researchers often describe what they did and what their results were, but they might not give readers a step-by-step guide for implementing their methods. This department pulls back the curtain and shows readers exactly how program evaluation is done. If you have an idea for an article of about 3,000 words on a particular evaluation method or an interesting development in the art of evaluation, please send a one-paragraph abstract to [marina.l.myhre@hud.gov](mailto:marina.l.myhre@hud.gov).

---

# Fair Housing Testing: Selecting, Training, and Managing an Effective Tester Pool

Claudia L. Aranda  
Sarale H. Sewell  
Urban Institute

---

## Abstract

*The paired-testing methodology originated as a tool for fair housing enforcement and has been used in the multiple housing discrimination research studies funded by the U.S. Department of Housing and Urban Development since the late 1970s. In a paired test, testers who are comparably matched on personal, financial, and homeseeking characteristics—except for the characteristic being investigated, such as race or ethnicity—independently record information received by a housing provider. Each tester in the pair collects data that can detect and document the incidence and forms of discrimination at multiple points in the homeseeking process. Whether a fair housing testing study is designed for enforcement or research purposes, its successful implementation requires an effective tester pool. This article highlights important steps in tester selection, training, and management, all of which have been executed by the Urban Institute’s Field Operations Team since the spring of 2011 while supervising the completion of more than 13,000 paired tests across multiple housing discrimination studies regarding race and ethnicity, familial status, disability, sexual orientation, and gender identity.*

## Introduction

Since the late 1970s, the paired-testing methodology has been used in housing discrimination studies funded by the U.S. Department of Housing and Urban Development (HUD) to measure patterns of adverse treatment across the housing market. The methodology has been adapted for both research and enforcement purposes to investigate differential treatment on the basis of specific characteristics. For example, in a test designed to estimate the level of discrimination against families with children, two comparably qualified homeseekers—one with children and one without children— inquire about available housing. Each tester documents the information he or she obtains and the level of service provided—from contacting an agent; securing an appointment; meeting with an agent to view available units; and learning about move-in dates, monthly rent, security deposits, utilities, and any required fees. The results of the paired tests are then compared to determine whether and how the treatment experienced by testers with children differs systematically from that experienced by testers without children. Since forms of discrimination can be less blatant than they once were, housing testing studies can reveal important insights into marketwide behaviors and uncover systemic practices that would otherwise go undetected.

Since the spring of 2011, the Urban Institute’s Field Operations Team has supervised the completion of more than 13,000 paired tests across multiple housing discrimination studies (HDS) on race and ethnicity (HDS2012), familial status (HDS-Families), disability (HDS-Disabilities), and sexual orientation and gender identity (HDS-LGT). During the course of these studies, the Urban Institute contracted with testing organizations based in more than 40 cities across the country to coordinate tests. Although most of these groups have been fair housing organizations with active testing programs, some have had limited or no previous testing experience.<sup>1</sup> For all these HDS studies, the Field Operations Team was led by a director of field operations and regional coordinators based at the Urban Institute who were responsible for training local test coordinators, overseeing tester recruitment, training testers, supervising testing and test report preparation, reviewing test reports, maintaining daily contact with test coordinators at each site, and monitoring incoming data (submitted via an online data collection system). Careful oversight and regular communication enabled the Field Operations Team to anticipate operational challenges and correct problems as soon as they developed at any study site. The implementation lessons of the many HDS studies can help illuminate the “best practices” in building and sustaining a tester pool capable of completing the meticulous work that paired testing requires. The successful completion of any fair housing testing study requires (1) a careful tester-selection process, (2) a rigorous training program, and (3) and effective management, all of which are discussed in the forthcoming sections.

## Tester Selection

One of the first tasks any fair housing testing study must accomplish is the successful recruitment and selection of capable and committed testers. On each of the HDS studies, project staff have expended considerable effort recruiting testers who could be matched on age, gender, and other

---

<sup>1</sup> For the HDS-Disabilities study, more than one-third of the local testing organizations were disability advocacy groups and centers for independent living with no previous testing experience.

relevant characteristics to compose suitable tester pairs. Even organizations with robust testing programs have needed to recruit additional testers to complete the required number of HDS tests, particularly because some studies have required sites to conduct between 200 and 600 in-person paired tests. Selected testers ideally will have sufficient availability to complete multiple tests on a study.<sup>2</sup> On more recent HDS studies, the project team has established caps for the number of tests any single tester can conduct. A tester cap can help limit the extent to which the characteristics and behavior of any tester or tester pair can affect the study findings. Because of the amount of tester attrition that testing organizations experienced during HDS2012, groups participating in subsequent studies have been strongly encouraged to recruit 15 to 20 percent beyond the expected recruitment goals. Across all the recent HDS studies, the level of tester attrition has been attributed, in part, to the chronic underemployment of testers—when testers are offered permanent employment, they leave their short-term, part-time jobs as testers. Some level of tester attrition is also expected immediately after the tester training session (when testers learn how detailed the protocols are) and after testers conduct their first practice test (when some testers realize they are uncomfortable assuming a set of assigned but untrue characteristics). Because the recruitment and training of testers is laborious and costly, the assembly of a large tester pool early in the study can help forestall the need for subsequent recruitment drives and trainings, which can delay the completion of testing.

When conducting outreach for testers, HDS local testing organizations have been strongly encouraged to delve deeply within their existing networks of social service agencies, community groups, student associations, and nonprofit organizations to identify prospective tester candidates. To achieve the diverse tester pool that HDS required, testing organizations also expanded their typical recruitment efforts, forging new relationships with organizations and community leaders.<sup>3</sup> For example, each HDS study established specific targets for racial and ethnic representation based on metropolitan area census data. Achieving these goals, particularly the goal numbers for Hispanic and Asian-American testers, proved difficult for many sites. As a result, the project team based at the Urban Institute provided support to local testing organizations by connecting with Hispanic and Asian-American community groups and national umbrella organizations to try to increase the diversity of the tester pool. Given the level of confidentiality that fair housing testing requires, project staff exercised vigilance at every step during the recruitment process to ensure that partner organizations helping identify prospective testers understood the protocols to guard against disclosure. General advertising was strongly discouraged, and anytime organizations prepared e-mails or flyers for targeted outreach to specific groups or communities, materials were submitted for approval to the project management team. The comprehensive communication strategy avoided the use of such terms as “testing,” “fair housing research,” and other terms that might disclose the sensitive nature of the work.

---

<sup>2</sup> Depending on the number of tests to be conducted and the size of the study’s budget, it may not be worthwhile to expend the time and resources necessary to train a tester who may have availability to conduct only a few tests.

<sup>3</sup> Unlike some enforcement organizations, which only use testers who are obviously of their race or ethnicity, HDS testers have been diverse in skin color and accents. After the fieldwork for HDS2012 was completed, a team of coders assessed the racial and ethnic identifiability of each tester based on the tester’s name, recorded voice, and a standardized photograph. When overlaid with testing data, the identifiability analysis showed that minorities whose ethnicity is more readily identifiable experienced more discrimination than those who could be mistaken for White (Turner et al., 2013).

After outreach efforts identified a group of prospective testers, project staff conducted in-person interviews with candidates to determine which individuals were capable of fulfilling the responsibilities of the role. Because the testing process involves complex assignments and detailed protocols, testers must be selected carefully, according to their ability to perform the work. Project staff should consider the following criteria when assessing whether an applicant can be selected as a tester.

- **Affiliation with the housing industry.** Because of the sensitive nature of the work, applicants who wish to be testers should not work for or have immediate family who work for any segment of the housing industry, such as property management companies, insurance companies, appraisal companies, real estate firms, lending institutions, or other housing providers. When HDS testers notify project staff of such an affiliation, staff thank the applicant for his or her time and terminate the interview.
- **Confidentiality.** A successful testing program requires confidentiality—if housing providers were to learn that testing was under way, they could temporarily alter their practices or intensify efforts to identify potential testers. During the HDS interview process, project staff explain to potential testers that the information shared during the interview is to be kept strictly confidential, regardless of whether a candidate is ultimately selected to participate. After test coordinators determine the candidate meets the initial eligibility criteria, applicants are required to sign a confidentiality agreement.
- **Objectivity.** The importance of objectivity should be stressed during the initial stages of candidate interviews. Prospective testers must be able to conduct each test without making any assumptions about which housing providers are more likely to discriminate. Testers should be able to make fair and honest assessments of their experiences. Candidates who are unsure of their ability to remain objective throughout the testing process should not be considered.
- **Ability to be matched.** The selection of HDS applicants also hinges, to a large extent, on the ability to form tester pairs based on key characteristics, such as race and ethnicity, gender, and age. The personalities of testers also should be considered in the creation of pairs—individuals who are more outgoing should be matched with each other while more passive or reserved individuals should be established as matched pairs.
- **Ability to play a role.** Testers will be asked to assume certain personal and financial characteristics on tests that do not necessarily match their own. This set of characteristics includes an assigned household composition (marital/relationship status, number of children in the household, etc.), assigned employment (occupation, name and address of employer), and an assigned household income (the combined income of the tester and of any spouse/partner assigned that will be sharing the household). During the interview process, project staff will introduce the study and explain the role of the tester and expectations for the study. Testers are instructed to wear “clean and casual” attire and are expected to behave appropriately and credibly when playing the role of an interested homeseeker. If applicants are not willing or able to assume these characteristics on tests, then they should not be used for the study.

## Training

Regardless of the testing experience of organizations participating in HDS, they all have participated in a comprehensive training program with in-person training sessions and supplemental webinars, designed in consultation with each study's expert advisors and conducted by Urban Institute staff. The training program helps prepare local project managers and test coordinators to skillfully coordinate tests while providing guidance and support for testers, including those who may require specific accommodations and modifications. For example, during the HDS-Disabilities test coordinator training sessions, project staff outlined the approved modifications to the study protocols. Because note-taking was mandatory for all site visits, testers in wheelchairs with manual limitations were permitted to use other means to write down important information; some testers in wheelchairs used tablets or other electronic devices (which some found easier than writing by hand) or they were permitted to ask the housing provider to help them take notes. By explicitly defining how such modifications could be implemented, the study's field team helped ensure procedures were consistently used among the 30 sites coordinating in-person tests.

As part of the HDS tester training program, all testers participate in an in-person session (typically 5 to 6 hours long) that establishes protocols and guidelines, explains data collection forms, and reviews particularly challenging scenarios that may arise in the field.<sup>4</sup> The training delineates general codes of conduct and serves as a forum for any questions testers have before their first experience in the field. In addition, role-playing, watching short film clips, and taking pop quizzes increase tester engagement and reinforce key protocols. Testers also are strongly encouraged to read the entire tester manual, which includes a comprehensive index of everything covered during the training and sample test narratives that provide a chronological account of a tester's interaction with a housing provider. The HDS training program also requires that testers complete a practice test from start to finish. During practice tests, testers are briefed by test coordinators on their assignment, conduct a site visit, complete test forms, write a detailed narrative, and attend a debriefing session, exactly as they would during real tests. Testers also are trained on the use of the online data collection system. When training testers for a new study, test coordinators are encouraged to work closely with testers before and after their site visits to ensure that all protocols are followed. The practice test is an important opportunity for test coordinators to provide feedback to testers and to reinforce the high standards the work requires. Depending on the outcome of the initial practice test, some testers may be given a subsequent assignment before they can become an active member of the tester pool. By making this initial investment in time during the training phase of a study, testing coordinators can increase the quality of the tests to be completed.

---

<sup>4</sup> Because many HDS study sites have required more than one tester training session, Urban Institute staff have conducted subsequent sessions remotely via webinar. Throughout data collection on various studies, Urban Institute staff also conducted "refresher" sessions, highlighting key protocols via webinar or conference call.

## Management

Effective management is another key ingredient to maintaining a sufficient, capable tester pool. The HDS testing organizations that have consistently met the study goals on schedule and within the approved budget have excelled in working with and retaining their testers. The following tenets are important to a successful management strategy.

- **Communication.** Testing organizations that clearly communicate with testers about expectations, study goals and timelines, and payment and reimbursement processes—and any unexpected changes—have had the strongest relationships with their testers. The most successful test coordinators establish the expectation of regular communication during the completion of a test by texting, calling, or e-mailing testers to remind them of a scheduled site visit and requiring testers to contact them as soon as a test has ended. Such timely communication can make the difference between a failed test and one in which both testers successfully complete their site visits. In addition, before and at the conclusion of each test, staff conduct a one-on-one briefing and debriefing session, respectively. Testers are briefed in detail about their assignment to make sure they understand their profile and are reminded of the key protocols. After the tester completes the test, the test coordinator debriefs the tester about his or her experience, reviewing test forms, providing feedback, and answering any questions. If necessary, this session also will serve as an opportunity to review key test protocols if the tester made any errors; the complexity of the study protocols can affect the rate of error, which is usually highest on the initial tests that a tester completes. Even after a particular study has come to an end, organizations can notify testers of agency updates and upcoming testing opportunities, ascertaining any changes to testers' schedules and their ability to accept assignments.
- **Organization.** It is imperative that test coordinators remain highly organized, especially when their local testing organization has a large number of tests to complete. Test coordinators must stay informed about changes to testers' availability to ensure that tests will be completed according to schedule. In addition, test coordinators must perform a quality review of the test forms to address protocol errors with testers immediately after they occur, which will help minimize the number of repeated mistakes.
- **Efficient administrative procedures.** By establishing efficient and timely procedures for scheduling testers, processing invoices, and issuing payments, testing organizations help minimize tester frustration and increase the likelihood testers will accept assignments on an ongoing basis. HDS organizations that have experienced significant delays in issuing tester payments not coincidentally have also experienced a higher level of tester attrition. Groups that process payments according to an established payroll schedule help promote strong tester/test coordinator relations and maximize the use of available testers.
- **Setting expectations.** Throughout the recruitment and selection process of prospective testers, test coordinators can help set the tone for an entire study. For each of the HDS studies, testing organizations have been encouraged to set clear expectations at the outset, providing selected testers with all relevant information about the study, including training dates, compensation, and specific study requirements. By being explicit about the study's processes and rigorous

standards, test coordinators can identify testers who are most capable of completing test assignments and help reduce tester confusion and conflict. In addition, test coordinators must enforce any protocols or study requirements as necessary to ensure that tests are completed in a timely manner and meet the highest standards. Beyond setting expectations for testers' behavior, project staff also must exhibit the same high standards in their own conduct, adhering to study guidelines in test coordination and best practices in tester management.

## **Conclusion**

As the experience of the recent HDS studies demonstrate, fair housing testing studies, whether they are designed for enforcement or research purposes, require the assembly of a capable pool of available testers. Testing organizations must meet the challenge of identifying individuals with the capacity of maintaining confidentiality and objectivity while successfully adhering to study protocols. Selected testers must then complete a rigorous training program, which includes conducting at least one practice test. By making a significant investment of time during the training phase of a study, test coordinators will help testers complete higher quality tests with fewer errors, reducing the time needed for corrections and minimizing tester frustration. Sustaining a robust pool of testers also requires a comprehensive management strategy. Testing organizations must practice effective communication, implement efficient administrative procedures, and set high expectations for both project staff and testers alike. By maintaining a pool of credible, committed testers, testing organizations can successfully complete fair housing testing studies, helping identify emerging industry trends and uncovering discriminatory patterns and practices in both rental and sales markets.

## **Acknowledgments**

The authors thank the many individuals who have served as members of the HDS Field Operations team, including the extraordinary staff based at local organizations across the country. The perseverance of testers, including those who traveled significant distances (sometimes to other cities and states) to help conduct tests, provided an inspiring example for the entire project team based at the Urban Institute. The HDS studies could not have been successfully completed without their energy and dedication.

## **Authors**

Claudia L. Aranda is a senior research associate in the Metropolitan Housing and Communities Policy Center at the Urban Institute.

Sarale H. Sewell is a research associate in the Metropolitan Housing and Communities Policy Center at the Urban Institute.

## References

Turner, Margery Austin, Rob Santos, Diane K. Levy, Doug Wissoker, Claudia L. Aranda, and Rob Pitingolo. 2013. *Housing Discrimination Against Racial and Ethnic Minorities 2012*. Washington, DC: U.S. Department of Housing and Urban Development.