The Impact of Implicit Stereotyping on Offender Profiling: Unexpected Results From an Observational Study of Shoplifting
Dean A. Dabney, Laura Dugan, Volkan Topalli and Richard C. Hollinger
*Criminal Justice and Behavior* 2006; 33; 646
DOI: 10.1177/0093854806288942

The online version of this article can be found at:
http://cjb.sagepub.com/cgi/content/abstract/33/5/646
THE IMPACT OF IMPLICIT STEREOTYPING ON OFFENDER PROFILING

Unexpected Results From an Observational Study of Shoplifting

DEAN A. DABNEY
Georgia State University

LAURA DUGAN
University of Maryland

VOLKAN TOPALLI
Georgia State University

RICHARD C. HOLLINGER
University of Florida

Much debate centers on the use of offender profiling as a technique to differentiate criminals from law-abiding citizens. Profiling advocates argue that it is appropriate to reference past experiences and information about known offenders to identify behavioral and demographic correlates that can then be applied to a given population of offenses or offenders. The viability of this argument rests on the assumption that past experiences and information about known offenders are free of bias. Data from an observational study of shoplifting are analyzed to assess this assumption systematically. Results indicate that trained observers, when allowed to deviate from a clearly specified random selection protocol, oversampled shoppers on the basis of race, gender, and perceived age, thus misrepresenting these factors as predictors of shoplifting behavior. Implications for the training of law enforcement and loss prevention officers are discussed.

Keywords: profiling; shoplifting; implicit stereotyping; offender profiling; stereotyping

Offenders make up a small fraction of the population but intermingle among large numbers of law-abiding citizens as they commit their chosen offenses. Formal social control agents, most notably police officers and private security personnel, patrol social space with the goal
of identifying and apprehending individuals before, during, or after the commission of their offenses. The practice of “offender profiling” describes the efforts of social control agents to narrow the field of potential suspects. Its goal is to use what we “know” about crime and criminals to more efficiently and accurately differentiate the needle (criminal) from the haystack (noncriminals). To accomplish this task, police and other security officials draw on existing apprehension reports to identify common factors found in the offender pool. These factors become part of a screening process that law enforcement agents employ when they are out on patrol. In addition, authorities look for behavioral or demographic markers shown or believed to be present in the population of known offenders. Although offender profiles include a host of markers, it is the viability of demographic factors, most notably race, sex, and age, that has drawn the most attention.

The legitimacy of this practice rests on two very tenuous assumptions: namely, that the existing apprehension profiles are representative of the larger offender population in question and that social control agents can employ such profiles without personal bias. If surveillance and/or apprehension bias is found to be present in the existing data on known “offenders,” then the legitimacy of the “offender” profile is significantly compromised. For example, it may be that a disproportionate representation of non-White individuals is reflected in apprehension statistics not because they offend more often than Whites but because they draw more attention from police and security personnel. In this case, racial stereotypes and prejudices, rather than accurate representations of offender conduct, serve as the basis for the profile. Importantly, stereotyping need not be explicit to affect the practice of profiling. In fact, unconscious or unmotivated stereotyping may have a larger impact on this practice than previously realized. The present study directly examines the issue of implicit stereotyping and

AUTHORS’ NOTE: Please address all correspondence to Dean A. Dabney at Georgia State University, Department of Criminal Justice, PO Box 4018, Atlanta, GA 30302-4018; e-mail: ddabney@gsu.edu. The authors wish to thank the research assistants who tirelessly helped with the data collection processes. We also wish to thank the drug store chain and its management that permitted us to conduct this groundbreaking research. We hope that this type of academician-practitioner relationship could set the groundwork for other research projects on important policy related topics.
offender profiling. The primary goal in collecting these data was to identify actual shoplifters and record their behavioral and demographic characteristics as distinct from those of law-abiding shoppers.

LITERATURE REVIEW

THE USE OF BEHAVIORAL AND DEMOGRAPHIC PROFILES IN LAW ENFORCEMENT

Law enforcement officers feel justified in relying on the outcomes of past enforcement experiences to shape the direction of their present and future patrol protocols. But scholars point out that this assumption is unsound if past enforcement practices are shown to be biased, noting that statistical probabilities based on the racially biased enforcement patterns of the past will simply replicate those biases in the future. Put another way, Barlow and Barlow (2002) remark:

When police officers use race as a factor in criminal profiling based on presumed statistical probabilities, they contribute to the very statistics upon which they rely (Harris, 1999; Hughes, 2000; Leadership Conference on Civil Rights, 2001). Therefore, police officers justify profiling, stopping, searching, and thus arresting African Americans disproportionately precisely because African Americans are profiled, stopped, searched, and arrested disproportionately. (p. 338)

The self-fulfilling prophecy that follows from this circular logic has led to an explosion of social commentary on the topic of police profiling practices focused on the most common form of police-citizen interaction: traffic stops. Members of the media and legal communities have been quick to weigh in. For example, a Lexis/Nexis query using the search terms racial profiling and driving while black uncovered 18,425 news stories and law review articles on the topic from 1994 to 2000 (Meehan & Ponder, 2002). It is not surprising that this commentary has spawned state-sponsored preventative action. As of March 2001, formal orders were in place in all but 10 states mandating that law enforcement officers systematically collect race or ethnicity data on some or all of the persons that they stop, search, or arrest (Bureau of Justice Statistics, 2001). Similar decrees have spread across federal, county, and local-level jurisdictions (Ramirez, McDevitt, & Farrell,
The goal of these efforts has been to determine if police officers do, in fact, engage in racial profiling and, if so, whether it is a statistically valid means of identifying criminals. In most jurisdictions, law enforcement personnel or legislators were responsible for mandating data collection protocols on traffic stops. To date, more than a dozen analyses have been conducted on the resulting datasets.

Each of these inquiries finds that some level of disparity exists between Whites and non-Whites in terms of the level of traffic stops, searches, and/or arrests. Several studies find evidence of age and/or gender-based bias. A number of these inquiries point to low success or hit rates for these stops (i.e., stop or search uncovering wrongdoing) as evidence that disparities reflect racial discrimination on the part of the police, which in turn produces inefficient and unfair results in terms of arrest outcomes.

Third-party scholars have been hesitant to draw definitive conclusions about the race-based enforcement practices of law enforcement agencies under study, noting that these efforts have yielded a great deal of descriptive data but have been unable to overcome several critical design flaws (e.g., Engel, Calnon, & Bernard, 2002; J. Taylor & Whitney, 1999; Withrow & Jackson, 2001; Zingraff et al., 2000). These include but are not limited to questionable data collection, poorly designed protocols, inconsistent compliance, and data entry errors. Furthermore, critics note that researchers have yet to identify an acceptable base rate that can capture a group’s relative risk of being stopped, searched, and/or arrested.

Most scholars agree that the methodological shortcomings of these traffic stop studies leave us unable to determine whether police officers engage in unfounded and systematic age, sex, and racial-based profiling practices while out on patrol. Moreover, the data do not allow us to discern the process or outcomes of the alleged officer bias. These inadequacies notwithstanding, there does exist voluminous anecdotal evidence and commentary that alleges the presence of officer bias (Belknap, 1996; Cole, 1999; Walker, Spohn, & DeLone, 1996).

PROFILING SHOPLIFTERS

Acts of shoplifting, like traffic violations, occur with great frequency in the United States. Diligent retailers have shown that they
are willing and able to apprehend large numbers of thieves (Hayes & Blackwood, 1999; Hollinger & Davis, 2002). Given that shoplifters commit their crimes under the cover of other shoppers, retail shelving, and a nondescript outward appearance, store personnel struggle to identify would-be or active thieves amid a sea of honest customers. These factors come together to make shoplifting an obvious outlet for proactive offender profiling practices.

Experts have long strived to gain a comprehensive understanding of who shoplifts and how they perpetrate their crimes. These efforts are partially represented in the wealth of research that reports on the demographic and behavioral characteristics of shoplifters. Most of these studies rely on a self-report methodology that taps the experiences of a given sample of youth (for a comprehensive review of these studies, see Farrington, 1999; Klemke, 1992; Krasnovsky & Lane, 1998). Such a design does not allow us to measure the issue of surveillance or apprehension bias. There also exist several studies that rely on apprehension data that are provided by police, courts, and store authorities (e.g., Farrington & Burrows, 1993; Adams & Cutshell, 1984; Hayes & Blackwood, 1999, respectively). In general, these studies suggest that shoplifters are disproportionately male, minority, and from lower class backgrounds. Several scholars, most notably Farrington and Burrows (1993) and Davis, Lundman, and Martinez (1991), warn that significant surveillance and apprehension bias permeates these official data sources. In recent years, the mainstream media have expanded on this nuance, alleging that retail corporations subject minority shoppers to heightened levels of in-store surveillance and arrest. Such a movement has spawned the term shopping while black (Williams, Henderson, & Harris, 2001, p. 22).

Shoplifting prevention, much like traffic patrol, has become the focal point of a heated social and legal debate, with offender profiling coming under attack. Interviews with store security personnel reveal that employees admit to using behavioral factors (type of clothing, neatness of appearance, nervous behavior, loitering in certain departments) along with demographic factors (race, gender, age, and class) to profile potential shoplifters (May, 1978; Murphy, 1986). The substance of these profiles fits generally the stereotypical images of criminals, targeting the young, minorities, and persons of lower class status.
Several staged shoplifting experiments were conducted in the 1970s and 1980s to determine how law-abiding customers react when they witness an act of shoplifting (e.g., Bickman, 1975; Bickman & Green, 1977; Felder & Bert, 1984; Klentz & Beaman, 1981; Steffensmeir & Steffensmeir, 1977). These bystander studies clearly demonstrate that offender stereotyping manifests itself in the behaviors of law-abiding customers as well. Customers were shown to exhibit different reporting practices of shoplifting based on the age, race, gender, and perceived social class of the researcher who committed staged acts of theft in their presence. It is interesting that the demographics of the law-abiding customers were also found to shape reporting practices.

Critics allege that shoplifter profiles are flawed and that they reinforce the stereotypes and prejudices that permeate our social world. The previously mentioned studies involving samples of store personnel and bystanders to shoplifting draw heavily on the labeling theory tradition in sociology to account for their findings. The scholars that conducted this research contend that the shoplifting profiles are based on prejudicial practices and/or statistics and thus serve to reify cultural stereotypes. Research on labeling and cultural stereotyping sheds light on the issue of stereotyping.

SOCIAL PSYCHOLOGY OF LABELING, STEREOTYPING, AND PROFILING

Drawing on the symbolic interactionist tradition, labeling theorists use concepts such as the “looking-glass self” (Cooley, 1902, p. 281), “dramatization of evil” (Tannenbaum, 1938, p. 21), “stigmatization” (Goffman, 1963, p. 1), “master status” (Becker, 1963, p. 33), “primary and secondary deviance” (Lemert, 1967, p. 40), and “categorical devaluation” (Schur, 1984, p. 24) to describe how stereotypes draw our attention to certain types of people and how the resulting experiences, ranging from a suspicious glance to a trip through the criminal justice system, shape one’s self-concept. Labeling theorists argue that offender stereotyping is commonplace and breeds profiling patterns that reinforce the divisions between law-abiding and law-violating persons.

In the social psychology literature on cultural stereotyping, scholars note that profiling is at its core an act of social inference, albeit a structured one, no different from the myriad of other social thought
processes that people employ every day. People engage in social thought processes to discern the intent, nature, and possible behaviors of others all the time. Profiling is a specific and directed form of social cognition aimed at determining who has or is likely to engage in crime. It is different from the other social thought processes that we engage in (such as forming an attitude about a presidential candidate or deciding whom to invite to dinner) because agents of formal social control direct it toward a specific purpose. In such a context, the power of social inference is being used as a crime fighting tool, and its impact on society is systematic but potentially flawed.

As a category of social cognition, profiling is a specialized, directed form of attribution formation. Attributions (Heider, 1958; Jones & Davis, 1965; Kelley, 1967) help us infer the intentions of others and how our behavior or threat of behavior have affected them (S. E. Taylor & Lobel, 1989; Thompson & Janigan, 1988). The process of social engagement and social assessment takes place constantly as people interact with others, and it relies heavily on the formation of attributions. When law enforcement officials and security personnel are charged with the task of differentiating offenders from nonoffenders in social spaces (such as department stores or other commercial establishments), they are engaging in a structured form of attribution formation: profiling.

To form an attribution, individuals must gather social information and cues through observation and interaction, select that information which is pertinent to the immediate situation, and finally integrate it into existing beliefs and schemas (Berkowitz, 1988). Once formed, attributions guide one’s subsequent behavior. The process is a valuable one that facilitates daily social functioning. However, it has drawbacks, the most notable of which is the susceptibility of attributions to existing belief systems such as stereotypes. Stereotypes are preconceived notions regarding the nature of individuals based on our attitudes toward particular groups (be it racial, gender, or otherwise) and our schemas about how those groups behave. However, social inferences based on stereotypes can often produce faulty judgments (attributions) that can then lead to erroneous decisions and inappropriate behaviors.

The tendency to rely on stereotypes to guide behavior and attributions has been found to be powerful and subtle, as we often fall
back on stereotyping processes without even knowing it. The best examples of this are studies in the social psychology literature regarding the nature of implicit stereotypes. Recent advances in the literature of social cognition have led to the distinction between implicit and explicit stereotypes in social judgment, particularly with respect to race and gender (Banaji & Hardin, 1996; Devine, 1989; Fiske, 1998; Wittenbrink, Judd, & Par, 1997). Explicit stereotypes originate through conscious, directed information processing, generally congruent with previously determined or established attitudes toward a group. Conversely, implicit stereotypes are a form of automatic information processing that operate unconsciously on social judgments, separate from the kind of conscious information processing that produces explicit stereotyping biases. Implicit stereotypes function almost reflexively, despite conscious efforts on the part of individuals to resist their use (Greenwald, 1990). Indeed, a number of studies have found individuals are often unaware that such processes are at work and do not realize they have influenced their social perceptions or judgments (Bargh, 1999; Fazio & Dunton, 1997; Fazio, Jackson, Dunton, & Williams, 1995; Von Hippel, Sekaquaptewa, & Vargas, 1997).

Because such processes differ from those evidenced in explicit stereotyping, direct measurement of implicit stereotypes (through the use of paper-and-pencil racial attitude questionnaires, for example) has been largely ineffective. As a result, social cognitivists who study implicit thought processes have turned to the use of indirect measures of stereotyping. Furthermore, few strategies have been shown to be effective in limiting implicit cognitive effects (Blair, Ma, & Lenton, 2001). Those that have require the systematic application of tactics designed to mitigate automatic processes (such as those associated with implicit stereotyping). These techniques require conscious effort and time-consuming training (Dasgupta & Greenwald, 2001; Mitchell, Nosek, & Banaji, in press; Nosek & Banaji, 2001).

Given the difficulty of mitigating implicit stereotyping processes in attribution formations, one would fully expect the process of profiling to be seriously flawed, even when agents of social control are specifically instructed to remove such bias from their profiling efforts. In fact, there is evidence that police officers are no better
than laypersons at discerning deceptive or potentially illegal behavior
(Akehurst, Koehnken, Bull, & Vrij, 1996), suggesting that implicit
stereotyping may be at work.

**METHOD**

The findings presented in this article are the unexpected product
of a larger research project to generate an unbiased demographic and
behavioral profile of active shoplifters. Unconvinced of the viability
of self-report or official report methodologies, we turned to observa-
tional techniques. The goal was to conduct unobtrusive observation
of a large sample of persons as they move about a select retail envi-
ronment as a means to better articulate the descriptive features of
shoplifting. Along the way, we were confronted with a completely
unexpected anomaly in the data that turned our attention to the issue
of surveillance and apprehension bias. It also reinforced the impor-
tant role that serendipity can play in the research process.

**SETTING**

The study was conducted in a retail pharmacy/drug store located
in Atlanta. The setting was chosen for two reasons. First, these retail
outlets cater to all customers, irrespective of class, race, age, and
socioeconomic status; they carry a wide range of products that span
a broad price range, are open long hours, and are now found virtu-
ally everywhere. Second, the typical shopping trip to a drug store is
relatively short and focused, making it amenable to our observa-
tional platform.

The Atlanta metropolitan area was chosen because (a) the selected
retail drug chain had 151 store locations within or around the city lim-
its and (b) the city of Atlanta possesses the type of diverse population
to permit the assessment of the research goals stated herein. Corporate
marketing profile reports and 1990 Census data were accessed to iden-
tify a store location that serviced a demographically diverse population
of shoppers. The final setting was a single store positioned on the
social crossroads between several ethnic and racially diverse neigh-
borhoods. It was on a major commercial thoroughfare serviced by a
bus route and was open 24 hr a day, 7 days a week, allowing variation of the observational periods. In addition to several subdivisions consisting of single-family dwellings, there were a university, government housing project, numerous apartment complexes, and several assisted living communities located within a 3-mile radius of the store.

MATERIALS/EQUIPMENT

Data collection was accomplished using a sophisticated array of 16 closed-circuit, high-resolution color closed circuit television (CCTV) cameras, permitting unfettered observation of persons as they moved about in the 7,500 square feet of retail space. At the heart of the system were four small, state-of-the-art CCTV cameras that were concealed behind small smoke tinted domes and mounted in the store ceiling. The authors, working in consultation with corporate loss prevention staff, carefully selected the location of every camera to minimize the number of “dead spots” in the store where shoppers could not be closely observed. The objective was to create an observation field in which the entire shopping trip could be observed from start to finish without any interruption in the vision field.

The camera system allowed all shopper actions to be viewed in real time and videotaped for later analysis. No audio recording was done. The CCTV cameras were operated and monitored by trained research assistants from the control room located in a second floor office in the front corner of the store (see Figure 1, in which triangles represent fixed cameras and circles represent rotating cameras). Ten observers were employed during the course of the 12-month data collection effort. Each observer was subject to a strict training regimen. This involved familiarization with the data collection instrument, training and retraining in the use of the camera system and sampling procedures, and a supervised trial observation that lasted 4 to 8 hours.

Customer data were recorded using a standardized but evolving data template. The instrument allowed for the systematic recording of demographic, behavioral, and situational data for each observation. Among other things, this template allowed for the date, time, number of shoppers, and other descriptives of each shopping trip to be recorded. The ethnicity, gender, approximate age, and a rough estimate of social class (based largely on dress) were also documented. Finally,
a section of the instrument allowed for a broad range of information to be recorded in those instances where the shopper was observed committing a theft (e.g., specific theft techniques and merchandise descriptions). The videotapes and observational data were subjected...
to a follow-up interrater reliability exercise to determine the consistencies of the behavioral and demographic estimations.

SAMPLING

Given the number of shoppers that enter a busy suburban drug store, it was necessary to use a structured sampling protocol. The initial sampling protocol required that every third shopper who entered the store be followed by CCTV during his or her entire shopping trip, however long it should last. It was deemed important that no effort be made to purposively select any particular age, race, gender, or social class of shoppers. This protocol was followed for the first 6 months of data collection. However, during the first few months, it became apparent that this purely random assignment criterion was costly and time prohibitive.

A revised protocol was adopted, under which observers were instructed to follow every third shopper who was dressed in such a manner that they were afforded a reasonable opportunity to actually conceal merchandise. This meant that shoppers wearing tight clothing, such as exercise clothing, with no pockets, jackets, bags, and the like were not subject to surveillance. However, if the selected third person had no visible signs of potential concealment through baggage or baggy clothes but within 15 seconds of entering the store was seen exhibiting behavioral cues well documented (Baumer & Rosenbaum, 1984; Hayes, 1993; Helena, 1995) to be associated with shoplifting, then the observers were instructed to include the shopper in the sample. Thus, the observers were given one additional degree of selection discretion. This was a strictly defined point of discretion, as the observer had to identify and document one of three behavioral cues that served as the basis for this determination. The shopper had to show visible signs that he or she (a) was aware of or looking for shoplifting countermeasures (e.g., camera, store employees, electronic article surveillance [EAS] alarm system), (b) was engaging in repeated head and neck movements that indicated “scanning” or “scoping” the store for onlookers (see Hayes, 1993, p. 54), or (c) was nervously playing with product packaging (e.g., reducing its size or shape, removing security tags). Each observer was subject to an additional 2 to 4 hours of training and follow-up spot checks to
assure that this new protocol was clearly understood before it was implemented. This revised sampling protocol was then used for the next 6 months of the project.

The idea for the current article came about after we began to notice a significant, but unexpected, difference in the racial composition of the sample generated via the random versus the nonrandom selection criteria. Specifically, greater proportions of African Americans and Hispanics were included in the sample after the new discretionary assignment protocol was implemented. Subsequently, this article describes efforts to determine whether stereotypical tendencies are so imbedded in U.S. culture that even researchers who were trained to be completely objective are still vulnerable to them.4 In effect, the individuals we hired to act as observers became the de facto participants in this research. This process allowed us to examine profiling based on three categories: race, sex, and age stereotypes.

Once we suspected that observers with discretion might be more likely to follow shoppers who matched their preconceived idea of probable shoplifters, the decision was made to extend the data collection effort another month to detect evidence of profiling. Non-White observers were added to assure that the final data set would include a sufficient number of cases who were followed by both White and non-White observers and who followed subjects both randomly and nonrandomly. The final data set consists of observational data on 1,555 shoppers gathered by 10 different trained observers. Table 1 shows the distribution of shoppers’ and observers’ race selection protocol across shoppers. Most apparent is that the discrete group did indeed select more non-White shoppers than the random group. Also, although far fewer of the shoppers were followed at random by non-White observers, the distribution yields enough statistical power to warrant further analysis.

**DATA ANALYSIS**

An interrater reliability exercise was conducted to assess the validity of the data coding processes. Roughly 10% of the videotaped observations documented by each of the study’s original trained observers were randomly selected. A diverse group of five
TABLE 1: Distribution of the Observers’ and Shoppers’ Race Conditioned on Sampling Protocol

<table>
<thead>
<tr>
<th>Condition</th>
<th>Random (n = 509)</th>
<th>Discretionary (n = 1,046)</th>
<th>Total (n = 1,555)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>%</td>
<td>M</td>
</tr>
<tr>
<td>Non-White shopper</td>
<td>215</td>
<td>42.2</td>
<td>634</td>
</tr>
<tr>
<td>White shopper</td>
<td>294</td>
<td>57.8</td>
<td>412</td>
</tr>
<tr>
<td>Non-White observer</td>
<td>65</td>
<td>12.8</td>
<td>147</td>
</tr>
<tr>
<td>White observer</td>
<td>444</td>
<td>87.2</td>
<td>899</td>
</tr>
</tbody>
</table>

individuals then viewed these materials and were asked to estimate the sex, race/ethnicity, social class, and estimated age of each person appearing on the tape using the original data template. They were also asked to document whether the shoppers wore clothing or carried items conducive to shoplifting and/or exhibited evidence of the behaviors that the protocol deemed suspicious. The level of agreement between the original and follow-up data templates was then calculated by hand. Although the agreement for the raw social class and age variables dipped below the 70% level, all of the recodes used in the analysis plan (i.e., juvenile, adult, elderly, or lower/working vs. middle/upper classes) produced agreement levels in excess of 85%. Near perfect agreement was observed for the gender, baggy clothes and baggage, suspicious cues, and race/ethnicity variables.

Next, differences in the means of the shoppers’ characteristics, conditional on the independent variables, were tested. This analysis allowed us to determine whether observers operating under the discretionary selection protocol were more likely to follow non-White, male, or adolescent shoppers than were observers following the strict random protocol. It also allowed us to detect a selection pattern by the race and sex of the observers. Finally, a bivariate analysis allowed us to determine whether those shoppers who act or dress suspiciously are disproportionately non-White, male, or adolescent.

In the next phase of the analysis, we estimated the probability that our observers would follow a subject who is non-White, male, or adolescent, conditional on the observer’s race, sex, and whether they
followed the discretionary protocol. Because the dependent variables are dichotomous, we estimated the coefficients using a logit generalized linear model. If the coefficient for discretion was significant and positive in any of the models, then the results would suggest that non-Whites, males, and/or adolescents were more likely than others to be chosen by the observers given discretion. We then added the selection criteria, suspicious behavior and dress variables, to the model. If the coefficient estimate for discretion remained positive and significant, we could conclude that our observers systematically chose to follow non-Whites, males, or adolescents regardless of their dress or behavior. We concluded the investigation by introducing an interaction term into the model to determine whether observers were any more likely to exhibit biases in their decision criteria according to their race.

RESULTS

BIVARIATE FINDINGS

Results for the bivariate analysis are found in Table 2. Each row displays the distribution of shopper characteristics conditional on the criteria at the top of the column. Because each characteristic is dichotomous, its mean represents the proportion of those who meet the selection condition at the top of the column that are described by that variable. For example, the .55 in the first column, first row means that White observers selected 55% of the non-White shoppers included in the sample. The corresponding .50 found in the “No” column for White observers means that non-White observers selected 50% of the non-White shoppers who were in the sample. Note that the difference between these two statistics is not significant.

The unconditional distribution of the shopper’s characteristics indicates that just more than half of them were non-White or male (.55 and .53, respectively). Only 4% of the shoppers appeared to be adolescent. This distribution changes once we condition our other independent variables, beginning with the observer’s race or sex. Statistically, Whites were just as likely to select non-Whites as were non-White observers (.55 vs. .50, p = .19). However, female observers appear to have been
| Shopper Characteristics | White Observers | | Female Observers | | Discretion | | Suspicious Cues | | Suspicious Dress |
|---|---|---|---|---|---|---|---|---|
| | Yes | No | Yes | No | Yes | No | Yes | No |
| Non-White (.55) | .55 | .50 | .57*** | .48 | .61*** | .42 | .59** | .52 |
| Male (.53) | .52*** | .63 | .54 | .51 | .57*** | .46 | .51 | .54 |
| Adolescent (.04) | .04* | .07 | .05 | .04 | .05* | .03 | .06** | .03 |

*p ≤ .10. **p ≤ .05. ***p ≤ .01, two-tailed difference.
more likely than males to select non-White shoppers (.57 vs. .48, \( p < .01 \)). This appears to be the only statistical distinction between female and male observers. Non-White observers were significantly more likely than Whites to choose males (.63 vs. .52, \( p < .01 \)) and marginally significantly more likely to choose adolescents (.07 vs. .04, \( p = .09 \)).

The next three comparisons address a concern that those trained not to profile will disproportionately select non-Whites, males, and adolescents when given discretion. Indeed, those with discretion did have a higher probability of choosing non-Whites compared to the observers operating under the random selection protocol (.61 vs. .42, \( p < .01 \)). They were also significantly more likely to choose males (.57 vs. .46, \( p < .01 \)) and marginally significantly more likely to choose adolescents (.05 vs. .03, \( p = .06 \)). This, however, is not enough evidence to prove discrimination. If shoppers in these groups did, indeed, match the selection criteria of behaving or dressing suspiciously, then we would expect that the observers with discretion to more commonly choose them. The findings show that shoppers who behaved suspiciously were more likely to be non-White and adolescent. However, they were not disproportionately male. In fact, the results show that those shoppers who dressed suspiciously were more likely to be female.

In sum, the significance of the discretion variable in the above results suggest that the observers may have been exhibiting stereotypical biases when selecting more non-White, male, or adolescent shoppers. However, in defense of the observers, when we examined the videotapes of those shoppers who behaved or dressed suspiciously, non-Whites and adolescents were disproportionately represented. Thus, had the observers been following the appropriate selection criteria, we would logically expect them to oversample non-Whites and adolescent shoppers. However, we would also expect them to oversample females. Instead, they disproportionately selected only male shoppers, suggesting some gender profiling.

**MULTIVARIATE LOGISTIC ANALYSIS**

To better estimate the impact of discretion on selection, we controlled for the selection criteria. If no discrimination bias was present, then the selection criteria of suspicious cues and suspicious clothing should account for all of the difference that exists between the sampling dis-
tributions generated by observers following the discretionary selection protocol and those following the random protocol. However, if the discretion variable remains significant, then the findings will strongly suggest that the perceptual biases of the observers influenced their decision to oversample non-Whites, males, and adolescents.

Table 3 displays the odds ratios of the logistic regressions for all three outcomes with and without the selection criteria. All odds ratios greater than 1 show a positive association and those less than 1 show a negative association. Note that the likelihood ratio chi-square statistics for all models are significantly different from the intercept-only model.

The first column under each dependent variable shows the results of the model without controlling for the selection criteria. Consistent with the bivariate analysis, the discretion variable is positive and strongly significant in the first two models and marginally significant in the third, suggesting that those with discretion are more likely to choose non-Whites, males, and possibly adolescents. When we control for the selection, discretion remains significant in all three models.

Thus, above and beyond suspicious clothing and behavior, when the observers were granted discretion, they still tended to choose non-White shoppers (i.e., racial profile). The odds that such an observer would choose to follow a non-White shopper (regardless of dress or behavior) is more than twice that of the observers operating under a random selection protocol (odds ratio = 2.10). We also find evidence of gender-based profiling. Observers following the discretionary selection protocol have an odds of following males that is 1.55 times higher than that of the observers operating under the random selection protocol (regardless of dress or behavior). Finally, there is evidence that the observers were profiling shoplifters based on the perceived age of the person entering the store. After controlling for the selection criteria of suspicious cues, the odds that observers who were granted discretion followed adolescent shoppers was 1.75 times that of the observers operating under the random selection protocol.

Having discovered that observers with discretion are more likely to select non-White, male, and adolescent shoppers for observation, we next tested to see if this tendency varies by race of the observer. Although 10 observers are not necessarily representative of their entire race, one can argue that because they evaluated more than 1,500 shoppers, sufficient variation exists to make this exercise worth explora-
<table>
<thead>
<tr>
<th></th>
<th>Non-White</th>
<th>Male</th>
<th>Adolescent</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OR</td>
<td>SE</td>
<td>OR</td>
</tr>
<tr>
<td>White observer</td>
<td>1.120</td>
<td>.177</td>
<td>1.010</td>
</tr>
<tr>
<td>Female observer</td>
<td>1.308**</td>
<td>.156</td>
<td>1.272</td>
</tr>
<tr>
<td>Discretion</td>
<td>2.050***</td>
<td>.227</td>
<td>2.098**</td>
</tr>
<tr>
<td>Suspicious cues</td>
<td>—</td>
<td>—</td>
<td>1.325**</td>
</tr>
<tr>
<td>Suspicious dress</td>
<td>—</td>
<td>—</td>
<td>0.679</td>
</tr>
<tr>
<td>Likelihood ratio</td>
<td>53.72***</td>
<td>61.02***</td>
<td>28.57***</td>
</tr>
</tbody>
</table>

*p ≤ .10. **p ≤ .05. ***p ≤ .01, two tailed difference.
### Table 4: Odds Ratios for Logistic Regressions Predicting the Selection of Non-Whites, Males, and Adolescents With Race/Discretion Interaction Term

<table>
<thead>
<tr>
<th></th>
<th>Non-White</th>
<th>Male</th>
<th>Adolescent</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OR</td>
<td>SE</td>
<td>OR</td>
</tr>
<tr>
<td>White observer</td>
<td>1.876**</td>
<td>0.601</td>
<td>0.421**</td>
</tr>
<tr>
<td>Female observer</td>
<td>1.210</td>
<td>0.148</td>
<td>1.188</td>
</tr>
<tr>
<td>Suspicious cues</td>
<td>1.308*</td>
<td>0.148</td>
<td>1.045</td>
</tr>
<tr>
<td>Suspicious dress</td>
<td>0.685</td>
<td>0.193</td>
<td>0.330**</td>
</tr>
<tr>
<td>Discretion</td>
<td>4.383**</td>
<td>1.494</td>
<td>1.106</td>
</tr>
<tr>
<td>White Observer ×</td>
<td>0.435*</td>
<td>0.157</td>
<td>1.473</td>
</tr>
<tr>
<td>Discretion</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Likelihood ratio</td>
<td>66.66***</td>
<td>44.36***</td>
<td></td>
</tr>
</tbody>
</table>

*p ≤ .10, **p ≤ .05, ***p ≤ .01, two-tailed difference.

The results support the conclusion that despite intensive training and specific instructions to ignore shopper demographic characteristics in...
selecting potential or probable offenders, observers were unable to resist the power of implicit cultural stereotypes in shaping their selection of individuals. Specifically, when working within the limits of a discretionary sampling protocol built around exclusively behavioral characteristics, observers included disproportionate numbers of shoppers with demographic characteristics stereotypically attributed to shoplifters (non-White adolescent males). This suggests that offender stereotypes do translate into profiling behaviors. Though this determination cannot be made definitively on the basis of a single, post hoc analysis, the bad news is that if some offender stereotypes do in fact make their way into behavioral practices of trained research assistants, it will be difficult for police and security trainers to determine where to direct their cultural sensitivity and antiprofiling efforts. In addition, there is a very real possibility that years of racial profiling have produced a change in the behavior patterns of those most commonly targeted by such efforts. Young Black males, for example, may constantly look over their shoulders, a behavior that ironically even further attracts the attention of those focusing on behavior-based profiles.6

There currently exists disagreement as to how (or even whether) social control agents should engage in profiling practices, particularly in relation to police patrol and traffic stops. Authorities tend to endorse
directed patrol strategies that target known areas (i.e., hot spots) or individual characteristics that are shown to be associated with increased criminal propensity. This approach, it is argued, increases patrol efficiency and best protects the public from being victimized (Barlow & Barlow, 2002; Department of California Highway Patrol, 1998; Webb, 1999). On the other hand, civil libertarians and critical criminologists (e.g., American Civil Liberties Union of Northern California, 1999; Chambliss, 1994; Cole, 1999; Harris, 1999) observe that the use of directed patrol strategies and/or specialized crime control units can open the door for subjective bias on the part of social control agents. The controversy in this regard is “racial profiling” whereby a citizen’s race becomes the principal factor on which stops are initiated, regardless of whether they are warranted. With regard to shoplifting, previous research (see Buckle & Farrington, 1984; Dabney, Hollinger, & Dugan, 2004; see Klemke, 1992, for a review) indicates that demographic factors serve as poor predictors of which strata of the shopper population are most likely to engage in acts of theft. More problematic is the fact that analysis of store apprehension files (Davis et al., 1991) and law enforcement data (Farrington & Burrows, 1993; Farrington & Langan, 1992) raise the possibility that race and class bias may also play a role in shoplifting apprehension practices. From this perspective, nonobjective profiling practices (such as directed patrol strategies) seem ripe for abuse and injustice.

Our multivariate results support the above notion. However, there exists a more fundamental question at the heart of the profiling debate: Do the biased observational techniques produce results? Put another way, will “hit rates” (probability of observing a theft) increase if implicit stereotypes flourish? The data in Table 5 speak to this issue by presenting the probability of observing a theft, conditioned on the shoppers’ race, age, and sex, in the random and discrete subsamples. Referring to the “hit rate” data on race, note that the observers’ discrete strategy was successful, as more of the non-White shoppers were observed shoplifting in that group than in the random group (.11 versus .06, p < .05). However, one is reminded that the multivariate analysis, that adjusted for the sampling protocol and controlled for other relevant factors, showed no overarching race effect. The findings from this multivariate analysis are the more telling and “true” representation of what is transpiring on the shop floor.
TABLE 5: The Proportion Observed Shoplifting by Selection Criteria and Shopper Characteristics

<table>
<thead>
<tr>
<th></th>
<th>Random</th>
<th>With Discretion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Race</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-White</td>
<td>.06</td>
<td>.11*</td>
</tr>
<tr>
<td>White</td>
<td>.05</td>
<td>.03</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>.08</td>
<td>.10</td>
</tr>
<tr>
<td>Female</td>
<td>.03</td>
<td>.06</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adolescent</td>
<td>.13</td>
<td>.09</td>
</tr>
<tr>
<td>Nonadolescent</td>
<td>.05</td>
<td>.08*</td>
</tr>
</tbody>
</table>

*p < .05.

short, our more systematic analysis reveals that behavior—not race—predicts shoplifting. Elsewhere (Dabney et al., 2004), we delve into this issue in greater depth. One must be cautious when interpreting the high hit rate as evidence supporting racial profiling. Granted, higher numbers of non-Whites will be caught, but one must realize that this strategy also results in higher numbers of innocent shoppers being observed (and stopped if applied to apprehension protocols) and higher numbers of White shoplifters going undetected. One should not allow the ends to justify the means when the means knowingly produce inequality.7

When our trained observers were allowed to employ nonobjective processes to tag potential shoplifters, the resulting observation sets were heavily biased toward minorities, males, and youth, a finding predicted by the social cognition literature on automatic information processes and stereotyping. Furthermore, the extant literature on implicit stereotyping suggests that attempting to “train” such implicit leanings out of police officers and other agents of formal social control will most likely be futile. Even if such training were successful in the short run, continued exposure of these formal social control agents to offender environments thereafter would most likely lead to a re-establishment of implicit stereotyping processes. This, in turn, would yield a return to unfair and inefficient profiling practices (for systematic techniques for the mitigation of implicit stereotypes, see Dasgupta & Greenwald, 2001; Mitchell et al., in press; Nosek & Banaji, 2001). In short, one might successfully apprehend more African
Americans, but this strategy would simultaneously result in allowing larger numbers of White shoplifters to unfairly escape undetected.

Admittedly, the costs in money and time required to expose all current security agents to the type of systematic implicit stereotype-reducing training required would be prohibitive. Although hardly generalizable to all retail establishments or police patrol practices, our findings provide much food for thought to those who support the structured and limited use of directed patrol practices.

We recognize that several limitations hinder the overall contribution of our work. First, this article resulted not from an a priori research plan but rather as a response to an anomaly in the data that was discovered while pursuing a different research question. In particular, while using the observational data to predict the behavioral and demographic profile of shoplifters, we noticed that a race-based artifact in the data that were gathered only under the discretionary sampling protocol. This situation limited the degree to which we could explore the nature and dynamic of observers’ stereotyping practices. This notwithstanding, we were able to isolate and articulate race and gender biases that were introduced into the sampling process when observers were granted discretion in selecting shoppers entering the store. Additional research that relies on larger samples of observers and a more tightly controlled experimental design is needed to more completely explore the validity of our findings. Such projects should not only include larger samples of observers and shoppers but also allow for robust representation in the age, gender, and race of these individuals.

Our findings are also limited by the fact that we rely solely on observational techniques to estimate the age, race, social class, and other discreet characteristics of the persons being surveilled. Some of these characteristics are not easily approximated with perfect accuracy via visual observation (particularly when using such subjective measures as suspicious dress or behavior). Outward appearance is a less than ideal indicator of one’s social class or age, for example. There also exists the possibility that the nuances that shaped observers’ perceptions of what constituted suspicious behavior, baggy clothing, or other key indicators used here fall below or outside of the limited categories included on our data template, especially when applied to the race of the shopper (other variables such as age and
gender are less problematic). For that matter, different observers could have different thresholds on a given variable or the categories contained therein. To some extent, these measurement error issues are beyond the scope of our data. It is for this reason that we conducted the extensive interrater reliability analyses described above.

More important, this article speaks to the perceived race, class, age, gender, and so forth of the shoppers, not their actual demographic characteristics. Specifically, we have shown that perceived demographics do not always match actual demographics. As such, although it is comforting to gain affirmation of the measures used in this study, such affirmation is not critical to the research questions explored herein. More work is needed to assure that perceived demographic factors be reliably measured in subsequent studies of this kind. Future inquiries should consider the possibility of using a brief pencil-paper survey or other means to more accurately document the demographics of observed shoppers as they exit the store. Because perception of individuals is the crucial factor in these kinds of studies, such techniques should incorporate multiple observers and strive to achieve high interrater reliability levels among them. It would also be useful to conduct follow-up interviews with the observers to inquire into the criteria that shaped their decision to include or exclude shoppers from surveillance under a discretionary protocol. Such a strategy would also allow the researcher to engage head on the sources and dynamics of attribution formations and their associated stereotypes. This type of information would go a long way to helping scholars and practitioners alike better understand and implement effective social control protocols.

NOTES

1. A list of sources is available from the contact author.
2. These investigations often employ the use of visual and word association tasks, such as matching photographs of racial exemplars with negative and positive terms, to determine the degree to which unconscious, automatic stereotyping processes influence social judgments. These studies reveal the presence of unconscious automatic interference on deliberative judgments (Bargh, 1999; Devine, 1989; Greenwald & Banaji, 1995), similar to those required when one is attempting to profile potential offenders.
3. A significant number of persons who entered the store did not engage in traditional shopping activities but instead proceeded to the public restrooms or waited for their prescriptions to be filled. Although these individuals generally spent long periods of time on property,
they initiated little or no contact with store merchandise and thus posed little threat to shoplift. The observers were told to cease observing these individuals.

4. In the original analysis, had we not adjusted for the oversampling of Black and Hispanic shoppers by the observers, we might have erroneously concluded that non-White shoppers are more prone to retail theft (Dabney et al., 2004). Instead, in the refined analysis, we found that White females were less likely to steal than others, and Hispanic females were more apt to steal.

5. The basic model assumptions were met.

6. There also exists a slight possibility that observers who operated under the random selection protocol came to “learn” who shoplifts and then went about imposing those lessons onto their discretionary selection behaviors. This possibility is reduced by the fact that only three observers worked under both protocols, and these individuals witnessed a small number of shoplifting cases under random assignment.

7. As is the case with race, profiling on the basis of sex also produces enhanced hit rates, although not significant, as observed by a jump in predicted probabilities of .08 to .10 for men under the random versus nonrandom selection protocols. Conversely, our data for the age-graded variable show higher hit rates for adolescents under the random protocol (.13) than under the discretionary protocol (.09). Unexpectedly, significantly higher hit rates were observed for nonadolescents in the discrete sample.

REFERENCES


