The Impacts of Implicit Bias Awareness Training in the NYPD

July, 2020

Robert E. Worden
Sarah J. McLean
Robin S. Engel
Hannah Cochran
Nicholas Corsaro
Danielle Reynolds
Cynthia J. Najdowski
Gabrielle T. Isaza

Table of Contents

Chapter

Acknowledgements

We gratefully acknowledge the helpful suggestions of Lorie Fridell and Anna Laszlo, of Fair and Impartial Policing, LLC, and those of the NYPD Monitor, Peter Zimroth, and the Monitor's experts: Jennifer Eberhardt, John MacDonald, and Anthony Braga. Special thanks go to Tanya Meisenholder, Deputy Commissioner at the NYPD, for her help in making a rigorous evaluation design feasible, and to Lt. Tara Coffey, Lt. Daniel Weadock, and Sergeant Stephen Ward. Thanks also to Kenan M. Worden for assistance with the statistical analysis. Finally, we are grateful for the funding provided by Arnold Ventures (formerly the Laura and John Arnold Foundation) to conduct this research.

This research was commissioned by the International Association of Chiefs of Police (IACP) / University of Cincinnati (UC) Center for Police Research and Policy, and funded by Arnold Ventures (formerly the Laura and John Arnold Foundation). The findings, conclusions, and opinions in this report are those of the authors and do not necessarily represent the official position or policies of the IACP, Arnold Ventures, or the New York City Police Department. Please direct questions or comments regarding this report to Robert E. Worden, The John F. Finn Institute, 421 New Karner Road, Suite 12, Albany, NY 12205, 518-456-6323 ext. 308, rworden@finninstitute.org.

Executive Summary

In February of 2018, the New York City Police Department (NYPD) began in-service training on implicit bias for its 36,000 sworn personnel, using the Fair and Impartial Policing (FIP) curriculum. A team of researchers from the John Finn Institute for Public Safety and the IACP/UC Center for Police Research and Policy partnered with the NYPD to conduct evaluation research on the impacts of the training. The evaluation concentrated on the effects of the training among patrol officers assigned to commands in the Patrol Services Bureau, Transit Bureau, and Housing Bureau, whose training commenced in May, 2018 and concluded in April, 2019.

We assessed the immediate effects of the training on officers' beliefs and attitudes: their knowledge about the science of implicit bias and the potential implications for policing, and their attitudes about the salience of bias and discrimination as a social problem, and the importance of policing without prejudice. A survey was administered on the day of FIP training, either prior to or following the training on alternating days. We drew inferences about immediate training effects from the differences in pre- and post-training survey responses. The effect of the training on officers' knowledge about implicit bias was of moderate magnitude, though many officers' comprehension of the science of bias was limited. The effects of the training on officers' attitudes toward discrimination, and their motivation to act without prejudice, were fairly small, though prior to the training, most officers considered discrimination a social problem and felt individually motivated to act without bias. Officers regarded the training as beneficial: 70 percent reportedly gained a better understanding of implicit bias and more than two-thirds reportedly learned new strategies and skills that they expected to apply to their work. Nearly half rated the likelihood of using all five bias-management strategies as either a 6 or 7 on a 7-point scale anchored at 7 as 'very likely.'

We conducted a follow-up survey about officers' beliefs and attitudes and their actual utilization of FIP strategies, which was administered from June through August of 2019, ranging from 2 to 13 months following the training. Asked whether they attempted "to apply the FIP training in your duties over the last month," 42 percent said they had not, 31 percent said they attempted to use the bias-management strategies *sometimes*, and 27 percent said they attempted using them *frequently*. Comparing the follow-up survey responses to those on the days of training, we also detected some decay in the immediate effects of the training on officers' comprehension of the science of implicit bias.

The impact of police training is likely to be greater when it is supported by other organizational forces, of which immediate supervisors may be the most important. We surveyed sergeants post-training. We found that most sergeants view monitoring for bias as one of their responsibilities, and that they are willing to intervene as needed with

individual officers. One-quarter reported that they had intervened with an officer whose performance warranted intervention. Slightly more than half of the sergeants reportedly address issues of implicit bias during roll calls, thereby reinforcing the training.

Insofar as officers' unconscious biases may influence their enforcement decisions, and to the extent that officers apply their training in FIP strategies to manage their unconscious biases, we hypothesized that the training would lead to reductions in racial/ethnic disparities in enforcement actions, including stops, frisks, searches, arrests, summonses, and uses of force. We examined enforcement disparities at multiple levels of analysis – at the aggregate level of commands and the level of individual enforcement events. To isolate the effect of the training from other factors, the NYPD adhered to a protocol for a randomized controlled trial that provided for grouping commands into clusters scheduled for training by random assignment. This experimental control was supplemented by statistical controls in the analytical models. Overall, we found insufficient evidence to conclude that racial and ethnic disparities in police enforcement actions were reduced as a result of the training.

It is very difficult to isolate the effects of the training from other forces that produce disparate enforcement outcomes. Training impacts might be a signal that is easily lost in the noise of everyday police work. Estimating the effect of a single training curriculum on officers' decisions to invoke the law or otherwise exercise police authority may well be akin to finding the proverbial needle in a haystack. Furthermore, it has been presumed but not demonstrated that enforcement disparities stem, at least in part, from officers' implicit biases. Though research has shown that police officers, like the general public, hold unconscious biases, no scientific evidence directly links officers' implicit bias with enforcement disparities. To the contrary, the evidence – which is thin, to be sure – suggests that officers practice controlled responses even without implicit bias training. If disparities stem from forces other than implicit bias, then even a well-designed training that is flawlessly delivered cannot be expected to alter patterns of police enforcement behavior.

About the Authors

Robert E. Worden is the Director of the John F. Finn Institute for Public Safety, Inc., and Associate Professor of criminal justice at the University at Albany, State University of New York. He holds a Ph.D. in political science from the University of North Carolina at Chapel Hill. Professor Worden's scholarship has appeared in *Justice Quarterly*, *Criminology*, *Law & Society Review*, and other academic journals, and it has been funded by the National Institute of Justice, the Bureau of Justice Assistance, the Laura and John Arnold Foundation, the New York State Division of Criminal Justice Services, and other sponsors. Professor Worden serves on the New York State Law Enforcement Accreditation Council, to which he was first appointed in 2001. He previously served on the National Research Council's Committee to Review Research on Police Policies and Practices, whose report, *Fairness and Effectiveness in Policing: The Evidence*, was published by the National Academies Press in 2004. Professor Worden is the coauthor of *Mirage of Police Reform* (University of California Press, 2017) and *The Power to Arrest* (Springer, 2019). In 2018, he was recognized by the Police Section of the Academy of Criminal Justice Sciences with the O.W. Wilson Award for outstanding contributions to police education, research, and practice.

Sarah J. McLean is the Associate Director and the Director of Research and Technical Assistance at the John F. Finn Institute for Public Safety, Inc. She holds a Ph.D. in criminal justice from the University at Albany, with a specialization in policy and process. At the Institute, Dr. McLean designs and manages evaluative research on criminal justice strategies and interventions, such as gang violence interventions, juvenile prevention programs, and chronic offender initiatives. She is the coauthor of *Mirage of Police Reform* (University of California Press, 2017). Dr. McLean has formal training in qualitative techniques and extensive experience in engaging in qualitative research and analysis. She previously worked as a research associate at the Hindelang Criminal Justice Research Center of the University at Albany, and prior to that as a research associate at the National Center for Mental Health and Juvenile Justice (NCMHJJ).

Robin S. Engel, Ph.D., is Professor of Criminal Justice at the University of Cincinnati (UC) and Director of the *International Association of Chiefs of Police / UC Center for Police Research and Policy*. She recently served as UC's Vice President for Safety and Reform, where her administrative duties included oversight of the daily operations and reform efforts of the University of Cincinnati Police Division. Dr. Engel engages in police research and evaluation, with expertise in empirical assessments of police behavior, police-community relations, and crime reduction strategies. She promotes evidence-based practices in policing by establishing academic-practitioner partnerships. She has

served as Principal Investigator for over 80 research grants, and is ranked among top academics in the field of criminal justice/criminology based on publications in prestigious peer-reviewed journals. In 2017, Dr. Engel was awarded the *Distinguished Alumni Award* from the School of Criminal Justice at the University at Albany, and serves as a governor appointee for the *Ohio Collaborative Community-Police Advisory Board*. She also currently serves as the co-chair of IACP's *Research Advisory Committee*.

Hannah Cochran is a Research Analyst with the John F. Finn Institute. She holds a B.A. in both Political Science and Studio Art from Connecticut College, and an M.A. in Criminal Justice from the University at Albany, SUNY. She is the coauthor of *The Power to Arrest* (Springer, 2019). Her research interests include police culture, criminal investigation, and police-prosecutor relationships.

Nicholas Corsaro, Ph.D., is Associate Professor of Criminal Justice at School of Criminal Justice at the University of Cincinnati. He has served as a Director of Research in a number of different agencies and organizations. His evaluations of strategic policing have focused on open-air drug markets, gang violence, gun homicide, and police-staffing needs. He has also examined the role of police contracting on crime rates. His research appears in a large number of criminal justice, evaluation research, and public health peer-reviewed journals.

Danielle Reynolds is a Senior Research Analyst with the Finn Institute. She is also a Ph.D. candidate at the University of Albany. She holds an M.A. in Criminal Justice from John Jay College of Criminal Justice. Danielle's dissertation examines police supervision as a means to shape and guide officers' discretionary behavior. Her research interests include police supervision, police decision-making and behavior and police organization and subculture.

Cynthia J. Najdowski is an Associate Professor in the Department of Psychology at the University at Albany. Her research explores how social psychological phenomena shape criminal justice interactions in ways that produce miscarriages of justice for minorities, women, and children. Her previous work has been recognized with several national grants and awards and published in interdisciplinary journals such as *Law and Human Behavior* and *Psychology, Public Policy, and Law*. She recently published a volume entitled *Criminal Juries in the 21st Century* with Oxford University Press, which reviews contemporary challenges in the criminal jury system, and previously co-edited a book entitled *Children as Victims, Witnesses, and Offenders*.

Gabrielle T. Isaza is a Research Associate for the IACP/UC Center for Police Research and Policy and a doctoral candidate in the School of Criminal Justice at the University of

Cincinnati. Her areas of research include police effectiveness; survey design; and evaluations of police training programs, including de-escalation and implicit bias. More generally, Gabrielle is interested in fostering the practice of police-academic partnerships.

Chapter 1
Introduction

Racial and ethnic disparities in the outputs of policing and law enforcement are the rule rather than the exception – in vehicle and pedestrian stops, custodial arrests, citations or summonses, searches, and the use of force. The factors that give rise to these disparities are numerous, and their independent contributions to the disparities are difficult to estimate reliably. By many scholarly and popular accounts, however, unconscious or implicit bias is one factor that contributes to disparate treatment.[1] The President's Task Force on 21st Century Policing acknowledged the role that implicit bias may play in producing disparities, and it called for training to reduce the presumptive impact of implicit bias.[2]

The findings of social psychological research cast implicit bias as a likely influence on police behavior. The research has shown that in forming judgments and making decisions, people often rely on some common heuristics that simplify their cognitive tasks and enable them to reach conclusions quickly. Such heuristics are susceptible to predictable patterns of error. This research was popularized in Daniel Kahneman's 2011 best-selling book, *Thinking, Fast and Slow*, which distinguished "System 2" thinking that is slow, deliberative, and effortful, from the more common "System 1" thinking that is fast, intuitive, and automatic.[3] Some of the latter judgments are made outside of conscious awareness and rest on stereotypes – even stereotypes that the decision-maker would reject as inaccurate and not consider in the context of more deliberate decision-making. This common form of human judgment, along with widespread stereotypes, form the bases for unconscious or implicit bias.

The biases that can emerge are as numerous as the stereotypes and can affect judgments about many groups: women; people who are overweight; people with disabilities; people of particular religions; and of course, people of various races,

---

[1] See, e.g.: Jennifer L. Eberhardt, Phillip Atiba Goff, Valerie J. Purdie, and Paul G. Davies, "Seeing Black: Race, Crime, and Visual Processing," *Journal of Personality and Social Psychology* 87: 876-893; Lorie A. Fridell, "Racially Biased Policing: The Law Enforcement Response to the Implicit Black-Crime Association," in Michael Lynch, E. Britt Patterson, and Kristina K. Childs (eds.), *Racial Divide: Race, Ethnicity and Criminal Justice* (Monsey, NY: Criminal Justice Press, 2008), pp. 39-59; Jack Glaser, *Suspect Race: Causes and Consequences of Racial Profiling* (New York: Oxford University Press, 2015); Justin Nix, Bradley A. Campbell, Edward H. Byers, and Geoffrey P. Alpert, "A Bird's Eye View of Civilians Killed by Police in 2015: Further Evidence of Implicit Bias," *Criminology & Public Policy* 16: 309-340; cf. Lois James, Steven M. James, and Bryan J. Vila, "The Reverse Racism Effect: Are Cops More Hesitant to Shoot Black Than White Suspects?" *Criminology & Public Policy* 15 (2016): 457-479.

[2] The President's Task Force on 21st Century Policing, *Final Report of the President's Task Force on 21st Century Policing* (Washington: Office of Community Oriented Policing Services, 2015), p. 10. Recommendations for training are on pp. 11, 56, and 58.

[3] Daniel Kahneman, *Thinking, Fast and Slow* (New York: Farrar, Straus, and Giroux, 2011).

ethnicities, or nationalities. Judgments are particularly susceptible to unconscious biases when they must be made under time pressure and/or with information that is fragmentary and ambiguous. That social stereotypes can affect the judgments even of people who consciously reject them and who are avowedly unprejudiced make implicit bias especially pernicious. In policing, stereotypes that associate people of color, and especially Black people, with violence and other criminal behavior may lead law enforcement personnel to act unknowingly with bias.

As an antidote for this affliction, many police agencies across the U.S. have adopted implicit bias training, the prevalence of which was recently documented in the first known survey, conducted by CBS News.[4] Of the 109 participating departments, 105 (96%) reportedly provide implicit bias training for their officers; in 93 of those departments (89%), the training is mandatory. What is particularly noteworthy is the rapid spread of this training in recent years. Although 28 departments began teaching officers about implicit bias more than a decade ago, the majority implemented such training more recently, with 61 departments adopting such training within only the past five years. Neither the agencies themselves nor the law enforcement field know whether the training is efficacious, however; the full range of hypothesized effects of implicit bias training have not been systematically evaluated.

In 2018, the New York City Police Department (NYPD) joined the scores of other police departments in delivering in-service implicit bias training to its officers.[5] We have analyzed the impacts of the NYPD's implicit bias training on patrol officers' beliefs, attitudes, and enforcement behavior. This report details our methods and findings.

Implicit Bias in Policing

Plausible as it is, we should note that the role of implicit bias in generating disparities in enforcement is largely a matter of informed speculation rather than demonstrable scientific fact.[6] No systematic empirical research has established a direct connection between implicit bias, on one hand, and enforcement behavior that disfavors Blacks or other racial/ethnic groups, on the other hand. What might initially appear to

[4] CBS News, *We asked 155 police departments about their racial bias training. Here's what they told us*, August 7, 2019. Retrieved from https://www.cbsnews.com/news/racial-bias-training-de-escalation-training-policing-in-america/

[5] The NYPD also incorporated implicit bias training into its curriculum for recruits.

[6] In her ruling in the Floyd case, Judge Scheindlin speculated that "It would not be surprising if many police officers share the latent biases that pervade our society. If so, such biases could provide a further source of unreliability in officers' rapid, intuitive impressions of whether an individual's movements are furtive and indicate criminality. Unconscious bias could help explain the otherwise puzzling fact that NYPD officers check 'Furtive Movements' in 48% of the stops of Blacks and 45% of the stops of Hispanics, but only 40% of the stops of Whites." *Floyd et al. v. City of New York, 08-CV-1034*, pp. 44-45.

be a straightforward, intuitively compelling hypothesis is complicated by a number of factors. If disparities in enforcement do not stem at least in part from implicit bias, then training in implicit bias will not serve to attenuate those disparities, even if the training is well-designed, properly delivered, faithfully received, and conscientiously applied. We first consider the challenges in measuring implicit bias, and then review the research from which inferences about the effects of implicit bias in policing have been drawn.

Measuring Implicit Bias

Prejudice and discriminatory beliefs have always posed challenges in measurement, and as social acceptance for discrimination has decreased over time, research subjects have become increasingly reluctant to disclose their prejudices to researchers. Unconscious bias poses an even greater challenge to empirical research, since subjects cannot disclose cognitive processes of which they are unaware. Consequently, social psychologists have devised indirect measurement schemes to calibrate subjects' unconscious biases.

One method that is commonly used is known as sequential priming:

> Priming involves presenting some stimulus with the aim of activating a particular idea, category, or feeling and then measuring the effects of the prime on performance in some other task. Because the human mind is organized as networks of associations, activating any one idea has the effect of spontaneously drawing to mind associated thoughts, memories, and feelings. Priming can be used as a means of mapping the networks of associations for an individual because the same primes tend to activate different associative links for different people.[7]

One version of a sequential priming procedure is a "shooter task."[8] This exercise involves simulated situations in which subjects are prompted to assess the threat posed in a visual stimulus – that is, the image of a person holding either a firearm or an innocuous object – and decide to "shoot" (by pressing a key on a keyboard) or not shoot "as quickly as possible." The race of the person in the image is systematically manipulated to allow researchers to determine whether the accuracy of the subjects' judgments to shoot or not, and the speed with which they render a judgment, is affected by the race of the target person. Unconscious, automatic associations are thereby revealed. Typically administered to samples of college students, such research

[7] C. Daryl Cameron, Jazmin L. Brown-Iannuzzi, and B. Keith Payne, "Sequential Priming Measures of Implicit Social Cognition: A Meta-Analysis of Associations With Behavior and Explicit Attitudes," *Personality and Social Psychology Review* 16 (2012): 330-350.

[8] Joshua Correll, Bernadette Park, Charles Judd, and Bernd Wittenbrink, "The Police Officer's Dilemma: Using Ethnicity to Disambiguate Potentially Threatening Individuals," *Journal of Personality and Psychology, 83* (2002): 1314-1329.

finds that research subjects are quicker to shoot the Black "suspect," and they are more likely to erroneously shoot the unarmed Black suspect.

Other methods are also designed to uncover the associative structures that may unconsciously influence judgment. The widely used Implicit Association Test (IAT) assesses the speed with which research subjects complete a series of tasks, the ease (measured by speed) of which turns on unconscious associations, such as those that connect male and female names with male and female faces, respectively, or those that connect Black and White faces with pleasant and unpleasant words.[9] The IAT has been extensively used to measure implicit racial bias, and it tends to show that most people have a bias of at least moderate proportions against Blacks. The IAT is not without its critics, however.[10] Other measures include the more time-economical Brief Implicit Association Test (BIAT), the Go-No-Go Association Test, the sorting paired features procedure, and the Affect Misattribution Procedure (AMP).[11]

Research indicates that, like other members of the general public, police officers hold implicit biases. Joshua Correll and his colleagues analyzed the speed and accuracy with which samples of police officers and a sample of (civilian) community members performed the shooter task. Correll et al. found that officers and community members alike "exhibited robust racial bias in response speed," suggesting unconscious bias.[12] Lois James and her colleagues administered the race-weapons IAT to Spokane (WA) patrol officers, finding that all but 4 percent exhibited an anti-Black unconscious bias; the association was moderate in magnitude for 40 percent and strong for 38 percent.[13]

---

[9] Anthony G. Greenwald, Debbie E. McGhee, and Jordan L.K. Schwartz, "Measuring Individual Differences in Implicit Cognition: The Implicit Association Test," *Journal of Personality and Social Psychology* 74 (1998): 1464-1480.

[10] See https://implicit.harvard.edu/implicit/faqs.html. For criticisms of the IAT, see, e.g., Frederick L. Oswald, Gregory Mitchell, Hart Blanton, James Jaccard, and Philip E. Tetlock, "Using the IAT to Predict Ethnic and Racial Discrimination: Small Effect Sizes of Unknown Societal Significance," *Journal of Personality and Social Psychology* 108 (2015): 562-571; and Hart Blanton, James Jaccard, Erin Strauts, Gregory Mitchell, and Philip E. Tetlock, "Toward a Meaningful Metric of Implicit Prejudice," *Journal of Applied Psychology* 100 (2015): 1468-1481.

[11] For comparisons of these procedures to one another, see Yoav Bar-Anan and Brian A. Nosek, "A Comparative Investigation of Seven Indirect Attitude Measures," *Behavioral Research* 46 (2014): 668-688.

[12] Joshua Correll, Bernadette Park, Charles M. Judd, Bernd Wittenbrink, Melody S. Sadler, and Tracie Keesee, "Across the Thin Blue Line: Police Officers and Racial Bias in the Decision to Shoot," *Journal of Personality and Social Psychology* 92 (2007): 1006-1023. Also see Joshua Correll, Sean M. Hudson, Steffanie Guillermo, and Debbie S. Ma, "The Police Officer's Dilemma: A Decade of Research on Racial Bias in the Decision to Shoot," *Social and Personality Compass* 8/5 (2014): 201-213; and Eberhardt et al., "Seeing Black: Race, Crime, and Visual Processing," *op cit*.

[13] Lois James, Stephen M. James, and Bryan J, Vila, "The Reverse Racism Effect: Are Cops More Hesitant to Shoot Black Than White Suspects?" *Criminology & Public Policy* 15 (2016): 457-479. Also see Lois James, Bryan J. Vila, and Kenn Daratha, "Results from Experimental Trials Testing Participant Responses to White, Hispanic and Black Suspects in High-Fidelity Deadly Force Judgment and Decision-Making Simulations,"

Implicit Bias and Police Behavior

Biased behavior does not always follow from implicit biases, however. In the shooter task administered by Correll et al., only the community members were more likely to shoot the unarmed Black person than the unarmed White person. The authors attributed the officers' infrequent errors to their professional training. Furthermore, when comparing officers' performance to community members' in the shoot/don't shoot simulation, Correll and his colleagues found that officers "were faster to make correct responses; they were better able to detect the presence of a weapon; and they set significantly higher criterion for the decision to shoot, indicating a less 'trigger-happy' orientation."[14] This difference may be accounted for by officers' enhanced *controlled processing*, associated with increased activity in the areas of the brain tasked with parsing paradoxical information, as a result of their expertise and training. The authors note that "in some cases … training leads participants to work harder, in cognitive terms, as they learn to marshal the attention and control necessary for optimal performance."[15] Although police officers are as susceptible as anyone to holding implicit biases, their biases do not inevitably lead to biased actions.

James et al., improved the verisimilitude of the shooter task by placing research subjects – including police officers – in use-of-force simulators like those that police agencies use for training. The subjects were equipped with modified Glock handguns and instructed to react to "dynamic, interactive, life-size video scenarios that we designed and had filmed and acted professionally to capture the complexity and emotional content of deadly encounters while maximizing experimental control."[16] Officers were slower to shoot Black suspects than White suspects, and less likely to mistakenly shoot unarmed subjects who were Black. Their shoot/don't shoot decisions were unrelated to their IAT scores.

Outside of the controlled conditions of the laboratory, researchers have attributed disparities in enforcement actions to implicit bias in the absence of any measures of implicit bias. For example, Nix and his colleagues analyzed the data on fatal officer-involved shootings compiled by *The Washington Post*, claiming that,

> Our analysis provides insight as to whether the implicit bias effect manifests itself in the real world where officer safety is an immediate concern. Although we could not determine whether officers were quicker or more likely to fire their weapon at minority suspects, we argue that if minorities were more likely to have *not* been

Journal of Experimental Criminology 9 (2013): 189-212; and Lois James, Bryan J. Vila, and David Klinger, "Racial and Ethnic Bias in Decisions to Shoot Seen through a Stronger Lens: Experimental Results from High-Fidelity Laboratory Simulations," *Journal of Experimental Criminology* 10 (2014): 323-340.

[14] Correll, Park, Judd, and Wittenbrink, "Across the Thin Blue Line," 1020.

[15] Ibid, p. 1008.

[16] James et al., "The Reverse Racism Effect," p. 464.

> attacking the police/other civilians, or more likely to have been unarmed, this would indicate the police exhibit implicit bias by falsely perceiving minorities to be a greater threat to their safety (i.e., threat perception failures; see Fachner and Carter, 2015).[17]

Threat Perception Failure (TPF), according to Fachner and Carter, is a "mistake of fact" that occurs "when an officer perceives that a suspect is armed due to the misidentification of a nonthreatening object (e.g., a cell phone) or movement (e.g., tugging at the waistband)," and constituted nearly half of the shootings involving unarmed subjects.[18] Among officer-involved shootings in Philadelphia, they found that Blacks were more likely than others to be involved in TPF shootings. An additional 35 percent of officer-involved shootings of unarmed subjects were instances of "physical altercations," in which "the suspect reached for the officer's firearm or overwhelmed the officer with physical force," which they did not consider as threat perception failures.[19]

Nix et al. contend that implicit bias is at the root of their findings, which

> … showed that citizens in the *other* racial/ethnic group were significantly more likely than whites to have *not* been attacking the officer(s) or other civilians, and that Blacks were more than twice as likely as Whites to have been unarmed when they were shot and killed by police. These findings suggest evidence of implicit bias in real-world scenarios. In line with previous police shooting simulation studies (see Correll et al., 2002; Cox et al., 2014; Payne, 2001), it seems that officers may have been more likely to experience threat perception failures in fatal shootings that involved minority civilians.[20]

These findings are not in line with previous shooting simulation studies of police subjects, however. The inference about the contribution of implicit bias to the observed disparities neglects the distinction between stereotype activation and application. A cultural stereotype may be activated by a stimulus but not applied behaviorally if the actor exerts control.[21]

---

[17] Nix et al., "A Bird's Eye View of Civilians Killed by Police in 2015," p. 317.

[18] George Fachner and Steven Carter, *An Assessment of Deadly Force in the Philadelphia Police Department*, Collaborative Reform Initiative (Washington, DC: Office of Community Oriented Policing Services, 2015), p. 30. They did not attribute this pattern to unconscious bias, though they recommended that the department provide training in implicit bias.

[19] Ibid. This point is made by David A. Klinger and Lee Ann Slocum, "Critical Assessment of an Analysis of a Journalistic Compendium of Citizens Killed by Police Gunfire," *Criminology & Public Policy* 16 (2017): 349-362.

[20] Nix et al., "A Bird's Eye View of Civilians Killed by Police in 2015," pp. 328-329.

[21] See, e.g., Patricia G. Devine, "Stereotypes and Prejudice: Their Automatic and Controlled Components," *Journal of Personality and Social Psychology* 56 (1989): 5–18; Regina Krieglmeyer and Jeffrey W. Sherman, "Disentangling Stereotype Activation and Stereotype Application in the Stereotype Misperception Task," *Journal of Personality and Social Psychology* 103 (2012): 205–224.

Similarly, a task force examination of 26 fatal mistaken-identity, police-on-police shootings, and of anecdotal accounts of other police-on-police confrontations, concluded that.

> … mistaken-identity, police-on-police shootings have tragically killed undercover and plainclothes officers—white, black, and Hispanic—without any obvious racial or ethnic pattern; but the reality is strikingly different for off-duty officers. As far as we can determine, 1982 was the last year in which an off-duty, white police officer was killed in a mistaken-identity, police-on-police shooting anywhere in the United States. Since then, nine off-duty officers of color have been killed in such shootings, including both Officer Ridley and Officer Edwards in New York State.[22]

From this pattern and testimony of law enforcement officers, the task force drew the inference that implicit bias underlay this pattern:

> We find the scientific evidence persuasive that police officers share the same unconscious racial biases found among the general public in the United States. Specifically, we are persuaded by evidence that both police officers and members of the general public display unconscious biases that lead them to be quicker to "shoot" images of armed black people than of armed white people in computer-based simulations testing shoot/don't-shoot decision-making.[23]

One factor that may confound the inferences in these studies is neighborhood context. Research that takes account of the level of violence in the area in which police confront suspects is limited, to be sure. However, when laboratory experiments account for the dangerousness of the context, the bias against Black subjects disappears.[24] Rather, researchers have found that in dangerous environments, participants' tendency to shoot White subjects rises to the level of their tendency to shoot Black subjects. Research considering use of force incidents more generally has found that accounting for high crime areas reduced the formerly significant relationship between suspect race and levels of police force, even when controlling for citizen resistance.[25] Additional research on the neighborhood context of police use of force found that the effect of suspect race was confounded by neighborhood context, also controlling for situational factors. The authors concluded that "racial cues can and do signal threat …. But racial threat perception may be one manifestation of a more comprehensive threat-detection

---

[22] New York State Task Force on Police-on-Police Shootings, *Reducing Inherent Danger: Report of the Task Force on Police-on-Police Shootings* (Author, 2010), p. iii. Also see Al Baker, "Bias Seen in 'Police-on-Police' Shootings," *New York Times*, May 27, 2010.

[23] Ibid.

[24] Joshua Correll, Bernd Wittenbrink, Bernadette Park, Charles M. Judd, and Arina Goyle, "Dangerous Enough: Moderating Racial Bias with Contextual Threat Cues," *Journal of Experimental Social Psychology*: 47 (2011): 184–189.

[25] Lorie A. Fridell and Hyeyoung Lim, "Assessing the Racial Aspects of Police Force Using the Implicit- and Counter-Bias Perspective," *Journal of Criminal Justice* 44 (1998): 36–48.

process – a process that monitors the environment for a variety of threats."[26] These studies lend credence to the possibility that neighborhood context may account for the race-related differences in threat perception failures otherwise attributed to the Black-crime implicit bias association.

In 2017, a report on proactive policing was issued by a committee of experts appointed by the National Academies of Sciences, Engineering, and Medicine.[27] The committee reported that "There is, to our knowledge, no peer-reviewed work in psychology examining how any motivating factors, implicit or explicit, held by police influence their behavior towards subjects in the real world." The base of empirical evidence linking implicit bias to enforcement disparities is thin indeed. If implicit bias training fails to affect officers' enforcement choices, the reason might be that implicit bias does not influence officers' judgments.

Implicit Bias Training

Task forces and scholars have recommended implicit bias training for police with a view toward reducing the racial and ethnic disparities in enforcement, and it appears that police departments have heeded their calls. Agencies may also have responded to the demands of local communities and their leaders, though we are aware of no research on public opinion concerning implicit bias training.

In contrast, the popular press has demonstrated greater skepticism about the effectiveness of implicit bias training, cognizant of the limitations of extant knowledge. Writing in *New York Magazine*, Jesse Singal opined that "Perhaps no new concept from the world of academic psychology has taken hold of the public imagination more quickly and profoundly in the 21st century than implicit bias," and he cautioned that the most popular measure of implicit bias – the Implicit Association Test – may not measure implicit bias well or at all.[28] *The Atlantic* has played a particularly prominent role in reporting developments in and application of implicit bias training over the years. In 2017, *The Atlantic* examined the Salt Lake City Police Department's efforts to address the public's concerns with officers' implicit bias in the wake of the fatal police shooting of Patrick Harmon.[29] A key point laid out is that, by definition, implicit bias training for police "raises questions about the fundamental equation of policing, which is built on

---

[26] William Terrill and Michael Reisig, "Neighborhood Context and Police Use of Force," *Journal of Research in Crime and Delinquency* 40 (2003), p. 9.

[27] National Academies of Sciences, Engineering, and Medicine, *Proactive Policing: Effects on Crime and Communities* (Washington, DC: The National Academies Press, 2017), chap. 7, p. 21

[28] Jesse Singal, "Psychology's Favorite Tool for Measuring Racism Isn't Up to the Job," *New York Magazine*, January 11, 2017, https://www.thecut.com/2017/01/psychologys-racism-measuring-tool-isnt-up-to-the-job.html#_ga=2.178677415.2088139459.1574447252-658605711.1574447252

[29] Tom James, "Can Cops Unlearn their Unconscious Biases?" *The Atlantic*, December 23, 2017.

the notion of the fair cop." Stated differently: though science tells us that everyone is vulnerable to having implicit biases, the presumed impartiality of police as legal actors is inherently impugned when departments are unilaterally required to participate in implicit bias training. This may have larger implications for police in their daily duties with respect to their relationship with the public and capacity to enforce the law. The author notes the lack of evidence regarding the effectiveness, or even effect, of implicit bias training programs, despite their rapid spread across the United States.

Citing the unevaluated outcomes of these programs, *The Atlantic* quotes several leaders in the field of bias research as calling for further research. Patricia Devine, a psychology professor and Director of the Prejudice and Intergroup Relations Lab at the University of Wisconsin, is one of those leaders. "If agencies skip key steps, Devine said, like arming participants with concrete strategies for monitoring their own biases, they won't work." Phil Goff, of John Jay College, maintains that training "should include rigorous testing after classes finish to see if officers' reactions, behavior, or perceptions were actually changed by the material." Devine told the *New York Times* that,

> … she was troubled by the spread of such training in the absence of probing, objective research. She said more study of officers' unintentional biases is necessary to evaluate how training can impact their behaviors. Additional data is needed, she said, to determine if officers retain what they are taught and if civilians are benefiting from fairer policing. 'You could have the best of intentions and you could do something that you think intuitively makes sense, but it can and often does backfire; it makes things worse ….'"[30]

Relatedly, in 2008, *The Washington Post* reported that the majority of diversity training programs implemented in mid-size to large U.S. companies were ineffective, and were sometimes counterproductive to increasing diversity in workplaces, based on the research of Alexandra Kalev at the University of Arizona and her colleagues.[31] In 2016, *The Washington Post* revisited Kalev et al.'s continued research on diversity training, which supported their original findings and underscored the importance of ensuring that training is optional, not mandatory.[32] Also reported was Kalev and colleagues' findings regarding organizational implementation of diversity-focused programs, noting that companies that adopt "diversity task forces" demonstrate a stronger commitment to increasing company diversity, and effectively install a more

[30] Al Baker, "Confronting Implicit Bias in the New York Police Department," *New York Times*, July 15, 2018.
[31] Shankar Vedantam, "Most Diversity Training Ineffective, Study Finds," *The Washington Post*, January 20, 2008. See Alexandra Kalev, Frank Dobbin, and Erin Kelly, "Best Practices or Best Guesses? Assessing the Efficacy of Corporate Affirmative Action and Diversity Policies," *American Sociological Review* 71(2006): 589-617.
[32] Jena McGregor, "To Improve Diversity, Don't Make People Go to Diversity Training. Really," *The Washington Post*, July 1, 2016. Also see Frank Dobbin and Alexandra Kalev, "Why Diversity Programs Fail," *Harvard Business Review* 94 (2016).

durable organizational reform. Task forces, embedded within the organizational framework as an evolving and persisting commitment, may be preferable to diversity trainings because they provide for a long-standing effort to educate and diversify staff and management, rather than a single 8-hour training.

Doubts regarding the efficacy of implicit bias trainings have engendered some proposals for alternative solutions. In 2019, *The Atlantic* quoted Stanford psychology professor and current member of the NYPD Monitor's team Jennifer Eberhardt: "One of the ways to correct these biases is not simply though bias training …. It's trying to understand what practices or what policies might be driving the disparity."[33]

Despite the lack of empirical evidence regarding the effects of implicit bias training, it continues to be the de facto prescription to treat society's disparities, including bias in college admissions,[34] maternal health care,[35] and the daily practices of restaurant industry workers.[36] *The Atlantic* reported on Devine's Prejudice Lab and the unconscious-bias trainings developed and administered therein.[37] While most programs of this nature have little empirical evidence to support their utility, the author notes that there is evidence to suggest that Devine's training produces observable, expected outcomes. Devine's implicit bias curriculum is targeted towards a broader audience, and aims to "break the habit" of prejudice by making trainees aware of their biases, providing them with motivation to change, and equipping them with strategies to implement that change. The effect of these trainings is reported in *The Atlantic* as altering hiring patterns at the University, with a substantial increase in the proportion of female staff in those departments that underwent training, and stagnant numbers in the untrained departments. The effects appeared to be lasting ones, as two years after the training, "students who took part in a public forum on race were more likely to speak out against bias if they had participated in the training."[38]

---

[33] Annika Neklason, "A Practical Approach to Police Bias," *The Atlantic*, June 29, 2019.

[34] Megan Zahneis, "A Judge Advised Harvard to Give Its Admissions Officers Training to Stop Bias. Will That Help?" *The Chronicle of Higher Education*, October 4, 2019.

[35] Sarah Ruiz-Grossman, "California Takes New Steps To Stop Black Women from Dying in Childbirth," *The Huffington Post,* October 8, 2019; Annie Waldman, "New York City Launches Initiative to Eliminate Racial Disparities in Maternal Death," *ProPublica*, July 30, 2018. Retrieved from: https://www.propublica.org/article/new-york-city-launches-initiative-to-eliminate-racial-disparities-in-maternal-death

[36] Yuki Naguchi, "Starbucks Training Focuses on the Evolving Study of Unconscious Bias," *National Public Radio*, May 17, 2018. Retrieved from: https://www.npr.org/2018/05/17/611909506/starbucks-training-focuses-on-the-evolving-study-of-unconscious-bias.

[37] Jessica Nordell, "Is this how discrimination ends?" *The Atlantic,* May 7, 2017.

[38] The internal validity of this finding is weak, however. Furthermore, hiring decisions are typically more deliberative in nature than those of street-level policing.

Our Study of the NYPD

This evaluation of implicit bias training augments the empirical evidence on its impacts. In February of 2018, the New York City Police Department (NYPD) began in-service training on implicit bias for its 36,000 sworn personnel, and the department agreed to cooperate with us so that we could conduct research on the impacts of the training on patrol officers' beliefs, attitudes, and enforcement behavior.[39] The NYPD had previously contracted with Fair & Impartial Policing®, LLC (FIP), an implicit bias awareness training program used in numerous other police agencies across the U.S. The FIP curricula for various audiences (patrol, supervisors, and command staff) were adapted to the NYPD context and vetted by Jennifer Eberhardt, a member of the NYPD Monitor's team and preeminent expert on implicit bias. FIP trainers were responsible for delivering the training. During the rollout of the training, NYPD Commissioner O'Neill asserted in a mass email to NYPD personnel: "Understanding how perceptions can impact performance has relevance to a wide range of vocations, including the business world, the legal and medical professions, and more. But we're cops, and we owe it to ourselves and to the people we serve to keep current with the latest training available."[40]

Any evaluation of a single agency's implicit bias training will, in effect, evaluate the impacts of the training curriculum that the agency uses, and we formulated the evaluation design with the FIP training content in mind. Furthermore, the effects of any police training will depend on the extent to which the training content is compatible with and reinforced by the organizational and operational context, so to some (unknown) degree, the findings from an evaluation in any one agency are of limited generalizability. Nevertheless, we conducted the evaluation not only to inform the NYPD and its stakeholders, but also as a contribution to a body of evidence on the impacts of implicit bias training for police. Accordingly, in Chapter 2 we describe the FIP curriculum and five other implicit bias training curricula for police. We also discuss the generic outcomes in terms of which such training can be assessed, and draw on a broader body of social psychological inquiry to form a firmer base on which hypotheses about training impacts can be formulated.

In Chapter 3, we then describe the NYPD context, which is in several respects unique: the city is the largest in the U.S. and is surely among the most heterogeneous. The department is the largest police department in the country and diverse on several dimensions. Moreover, the current operational environment is marked by a far-reaching

[39] The recruit curriculum also includes content on implicit bias. In 2018, the NYPD held train-the-trainer sessions on FIP for its academy instructors, and they began delivering the training to recruits in August, 2018. See Peter L. Zimroth, *Tenth Report of the Independent Monitor* (New York: Arnold & Porter Kaye Scholer LLP, 2019; corrected January 20, 2020), p. 47.

[40] Peter L. Zimroth, *Ninth Report of the Independent Monitor* (New York: Arnold & Porter Kaye Scholer LLP, 2019), p. 23.

set of reforms, many of which were mandated by the federal district court and are overseen by a court-appointed monitor.

Chapter 4 summarizes the evaluation design, which includes multiple components and provides for two distinct randomized controlled experiments. We describe data collection methodologies and execution, the nature of the experimental controls, and analytic strategies.

In Chapter 5, we report findings on the immediate effects of the training on officers' beliefs and attitudes, including their awareness of and knowledge about implicit bias, and their attitudes concerning discrimination as a social problem and their motivation to act without bias. We also discuss officers' assessments of the utility of the training and, upon the conclusion of the training, their projections of the likelihood of its application in their work. These survey-based findings are supplemented by findings based on semi-structured, open-ended interviews.

Chapter 6 addresses a key form of organizational reinforcement for the training content: immediate supervisors. The training that supervisors receive instruct them in their role in monitoring officers' performance for signs of potential bias, and in intervening when they detect such signs. Supervisors can also reinforce the training with periodic reminders to be mindful of the potential impacts of implicit bias and the need to manage it.

In Chapter 7, we analyze officers' enforcement behaviors, pre- and post-training, to estimate the effects of the training on disparities in the outputs of policing. We examine stops and post-stop outcomes, arrests and the use of force in arrests, and the issuance of summonses.

Finally, in Chapter 8, we summarize the evaluative findings across the range of outcomes examined, and we consider the implications for policing and for policing research.

Chapter 2
Implicit Bias Training for Police

In general, there is a dearth of information on the implicit bias training curricula and methods that police agencies have put to use. The aforementioned CBS News survey shed some light on the prevalence of implicit bias training in American policing. The survey also revealed the frequency with which implicit bias training is delivered, and its duration. Among the police departments surveyed by CBS News, most reportedly offer implicit bias training to their officers once a year, although the frequency ranges from multiple times per year to only once every five years.[41] The duration also ranges across departments; some provide as many as forty hours of training and others as little as thirty minutes, but the most common length is eight hours.

Unfortunately, the CBS News report did not examine the specific content of the implicit bias training that police officers are receiving, so the extent to which such training mirrors the FIP program remains unknown. Whereas some departments appear to be relying on psychological and/or legal experts (e.g., Jennifer Eberhardt, Kimberly Papillon) to deliver lectures about the potential for unconscious processes to affect officers' decision-making, others have joined forces with organizations that have developed systematic programs.[42] FIP is one such program, and the NYPD is one of many agencies that have received the FIP training.

We describe the FIP curriculum here. To place FIP in the context of implicit bias training for police, we also describe five other implicit-bias training programs. We then discuss the outcomes on which implicit bias training for police can be expected to have effects, and we review social psychological research in terms of FIP features.

Training Curricula

Fair and Impartial Policing

The FIP curriculum was designed specifically for a police audience with support from the Office of Community Oriented Policing Services.[43] Introductions to the FIP

[41] CBS News, *We asked 155 police departments about their racial bias training. Here's what they told us*.
[42] https://thebettermind.com/
[43] Our description of the FIP training is based on FIP curriculum materials provided to us by Lorie Fridell and several other sources: the FIP website (http://www.fairimpartialpolicing.com/) and materials accessible through the site; Lorie Fridell and Sandra Brown, "Fair and Impartial Policing: A Science-Based Approach," *The Police Chief* (2015): 20-25; Dwyer Gunn, "Internal Affairs," *Psychology Today* (2017); 66-70, 79. Also see Matt Zapotosky, "In Push to Reform Police Work, Officers Examine Their Own Biases," *Washington Post*, January 6, 2016; Al Baker, "Confronting Implicit Bias in the New York Police Department, *The New York Times*, July 15, 2018.

program clearly establish the tenor of the training as a circumspect, science-based curriculum intended to be a benefit to the law enforcement community and not an additional burden. This training is designed to inform officers that bias influences everyone, while providing actionable instruction for minimizing its effect in the discharge of their daily duties. Issues of racial profiling by the police have long been misrepresented in the media as manifestations of rampant overt bias, which both alienates the law enforcement community and fosters a skepticism about civilian-prescribed antidotes to police bias. Officers naturally resist the broadly oversimplified reasoning that police simply need to "stop being prejudiced." The FIP training is delivered in a format that mitigates this defensiveness, and is crafted with cognizance of police sensibilities. Trainers, all but one of whom was either active or retired law enforcement personnel in 2018, are instructed to emphasize the non-accusatory tone and format. They stress the importance of bias-awareness in the context of community *and* officer safety, and highlight key points of the program: that biased policing is "unsafe, ineffective, and unjust."

The program's primary goals for all training levels include educating personnel about the science of implicit bias and its potential implications for police work, outlining the consequences that could possibly result from biased policing, and describing a number of specific skills sworn personnel can employ to manage their biases. Trainers instruct officers that bias as it exists in the modern world is thought to be overall less explicit, more implicit, and fairly ubiquitous. Trainers also concede that bias, in the context of a task that demands quick decision-making, vast discretion, and the authority to restrict people's liberty, can be especially calamitous. This introduction provides for officers a rationale for why they should care about FIP, which is communicated from organizational, historical, personal and philosophical perspectives.

Drawing on the science of implicit bias, the core lessons of the FIP curriculum first establish that the most pervasive biases are often unconscious, and held by people who would expressly disavow the stereotypes from which implicit biases are derived. The program cites various scientific studies on the subject, illustrating that time constraints and ambiguity creates a strain on people's ability to make rational, unbiased decisions. The ramifications of biased policing are described as not only rendering innocent civilians more vulnerable to unwarranted police attention and suspicion, but also creating blind spots where guilty people can avoid detection or pose threats to officers' safety with unearned inconspicuousness.

Trainers make the final unifying argument for increasing bias-awareness by explaining that bias reduction in law enforcement is instrumental to increasing levels of procedural justice and police legitimacy. By engaging in the FIP curriculum and making efforts to address issues of implicit bias, the police organization is making an effort to increase their levels of procedural justice, and thereby, increasing perceived legitimacy.

Without police legitimacy, or community trust and buy-in to police authority, police organizations cannot be as effective; citizens who perceive law enforcement as less legitimate are less likely to call the police for help, aid the police in an investigation, or stand by the agency as a source of community protection. The cornerstones of procedural justice organically overlap with the formula of FIP training – namely, to treat people with respect, to undertake their duties with "worthy intentions," to allow people to express their point of view, and, finally, to act without bias.

These topics are featured for all eligible levels of training: patrol, first-line supervisors, middle-management, and command staff. All levels of law enforcement train for one full workday (eight hours), except for command-level staff, whose training lasts one and a half days. Command staff may choose to include other high-level officers in their training, or, as recommended, include community members to participate in their training. The training is partially lecture-based, but features numerous activities, exercises, and role-play lessons which engage the participants in envisioning citizen encounters whose outcomes could be affected by bias. Trainers discuss the science of implicit bias by considering a number of studies that demonstrate the specific relevance of the issue to law enforcement. The curriculum cites research that details the benefits of use-of-force training as it specifically applies to implicit bias in policing, such as "shoot/don't shoot" simulations. Training for command staff encourages them to incorporate a consideration of implicit bias into agency use-of-force trainings to bring more conscious attention to otherwise hidden decision-making stimuli.

The FIP training construction for academy recruits, patrol officers, and first-line supervisors is designed to endow trainees with six specific skills in order to neutralize the effect of implicit biases. The first is simply to "reduce your biases." In presenting this goal, trainers recognize that biases are internalized over lifetimes and are therefore difficult to exterminate. The program embraces Contact Theory as a basis for this effort, which posits that positive contacts with people of other groups, or with counter-stereotypical members of these groups, can perceptibly break down long-held assumptions and biases. In this connection, trainers recommend that officers attempt to "blur lines," which encompasses a number of social exercises to help people look for similarities rather than differences between themselves and others.

Other skills prepare an officer to "manage your biases." With increased awareness of one's biases, officers are equipped with tools to identify and address the effects of bias on their actions and behavior. Trainers instruct officers to conduct self-checks on their decisions by asking questions like: "Would I still be stopping this person if they were White?" The skills required to check biases are not confined to self-policing, and the third skill warns officers to be aware of possible colleague and/or community bias. Trainers instruct officers to "avoid profiling by proxy" by utilizing other self-check skills.

If an officer is called to a scene because of reported suspicious behavior (as in the widely reported incident at a Philadelphia Starbucks franchise in 2018), he or she should not rely on the caller's evaluation of the situation (thereby, possibly profiling by proxy). Officers are instructed to assess the scene with fresh eyes and neutrality, mindful not to allow others' bias to influence their decisions.

Acknowledging that police officers' jobs often require quick decision-making under ambiguous circumstances, skill four encourages officers to "slow it down" when possible. This can also be thought of as consciously switching one's thought processes from "System 1" to "System 2" in order to "reduce ambiguity" by more carefully, deliberately, and objectively assessing a situation or an individual. These skills, combined with a thorough knowledge of the agency's biased policing policy (skill five), finally collectively contribute to skill six: analyze your options with a FIP perspective. In order to practice this skill, trainees form small groups that engage in role-playing policing scenarios, which present opportunities for officers to call upon various FIP tools. After these exercises, trainees reconvene to discuss their experiences, as well as their individual and collective reflections in a wider discussion format.

The basic contours of the FIP curriculum resemble those of Patricia Devine's implicit bias curriculum, noted in Chapter 1. FIP is designed to make trainees aware of their biases, provide them with motivation to change, and equip them with strategies to implement that change.

As noted above, the FIP curricula also cater to audiences of different ranks, with a curriculum for commanders, as well as one for middle-managers (which is a hybrid of first-line supervisors and command-level personnel), which we describe further in Chapter 3. The FIP training for supervisors parallels the content of the patrol officer curriculum, and it provides in addition distinct instruction for the supervisor's unique role and perspective. Role playing exercises for supervisors challenge them to construct responses to officers who have been involved in incidents possibly impacted by implicit bias, conceding that candidates for these conversations may be tricky to recognize: bias is difficult to surmise and still more difficult to prove. FIP trainers demonstrate how to structure such a conversation and engage in an inquiry into why the officer may have acted the way that he or she did, underlining tactics that diminish blame. Supervisors are reminded that as "role models, mentors, and representatives of the department," their own biases may carry more weight than they realize – whether it manifests in operational or managerial decisions, or simply is absorbed by officers under their command. Thus, education on the science of bias as it specifically pertains to the supervisor role is not only an important element to seeking out evidence of implicitly biased policing among subordinates, it is also instrumental for supervisors to reflect on their own behavior. An additional goal for supervisors is to help officers avoid the "over-control" response, which may result from a hyper-awareness of bias or concern that any

action would be interpreted as evidence of bias. Officers need to be confident that command staff will stand by them in their decisions to employ force, for example, and not fear "dire consequences" when they are compelled to use force. The skills section of the supervisor FIP module largely overlap with patrol, however supervisors are trained in an additional skill: "Communicate effectively, internally and externally, about bias." This skill encourages supervisors to take advantage of "teaching moments" as mechanisms for continuing the dialogue fostered by the FIP.

Tools for Tolerance®

The Museum of Tolerance in Los Angeles, California, developed Tools for Tolerance®, a series of programs designed for professionals in a variety of fields, a subset of which focus on enhancing law enforcement officers' skills for confronting diversity-related issues.[44] This stands in contrast to FIP, which was developed explicitly for police training. Similar to FIP, however, the Tools for Tolerance® programs are administered by civilian facilitators as well as trained law enforcement. The programs emphasize that awareness of personal values and responsibility can improve officers' ability to serve as "problem solver, mediator, and overall change agent" within their communities. In this way, Tools for Tolerance® stands apart from the FIP approach, which stresses that implicit bias may occur even when it does not align with one's consciously-held personal values and highlights not only community context but officer safety as well.

Programs range from one to four days in length, with several exceeding the length of FIP sessions. Like FIP, however, Tools for Tolerance® programs are geared toward small groups of uniformed and civilian personnel at every level, from recruit to commander, and involve a variety of pedagogical tools, including video presentations, small group activities, discussion, personal testimonies, and role-play and teach-back exercises. Unique to Tools for Tolerance®, however, is its utilization of interactive technology (i.e., anonymous polling equipment which allows participants to share their honest opinions, which are then graphically displayed) and visits to exhibits at the Museum of Tolerance or the New York Tolerance Center (described as "social laboratories designed to challenge visitors to confront personal biases and prejudice, and to promote awareness about tolerance issues"). Another difference between Tools for Tolerance® and FIP is that, although they both rely on research, the former set of programs draws on materials from the Simon Wiesenthal Center, whereas FIP relies on

[44] Tools for Tolerance® for Law Enforcement and Criminal Justice. (2019). Retrieved from http://www.museumoftolerance.com/for-professionals/programs-workshops/tools-for-tolerance-for-law-enforcement-and-criminal-justice/

scientific literature regarding contact theory, thinking systems, and implicit bias broadly and in the context of policing and use-of-force decision-making.

One 8-hour program tackles the issue of cultural diversity broadly. This program, limited to agencies in California and certified by the State of California Commission on Peace Officer Standards and Training (POST), evaluates the fact that our society is becoming increasingly diverse. It also considers the implications of that change for law enforcement. Although this program focuses more on sociological issues related to demography than psychological science related to bias, the emphasis on consequences for police and performance appears consistent with the FIP framework.

Five other Tools for Tolerance® programs center on the problem of racial profiling specifically. The first, POST-certified and mandated for officers within California, is a three-day training that defines racial profiling and addresses pertinent legal issues, all while reviewing relevant historical context and community concerns. The second program in this series is a mandatory four-hour session designed to update participants with the latest developments related to racial profiling. A third program, nationally available but not POST-certified, is entitled "Perspectives on Profiling." Participants are taught to distinguish between useful criminal profiling as a policing tool, on the one hand, and racial profiling that manifests as a result of bias and racism, on the other. Training emphasizes that police should continue to rely on their skilled intuition (i.e., experience-based ability to make quick, accurate judgments based on situational cues) and engage with potential suspects as appropriate; however, it highlights the importance of managing law enforcement challenges while also being mindful of community perception. Delivered via CD-ROM, the interactive video training presents real-life situations in which participants have multiple opportunities to make choices and receive feedback regarding the impact of their choices on the trajectory of the simulated situations. These programs are like FIP in that they distinguish between appropriate and inappropriate use of race in policing while simultaneously recognizing that officer decision-making may be affected by situational constraints. In contrast, whereas both Tools for Tolerance® and FIP recognize that racial profiling or bias-based policing may put innocent individuals at risk for unwarranted police attention, FIP further trains officers about the risks created when bias allows guilty or dangerous individuals to evade attention.

Similar to FIP, Tools for Tolerance® implements a train-the-trainer model, which is represented by the fourth and fifth programs on racial profiling. These focus on providing the trainers with "the information and skills necessary to ensure a successful presentation" of the curricula. One is specific to the California training and lasts five hours, and the other corresponds to the interactive video training. It is three days long and addresses the substantive content of the racial profiling training while also training

trainers in skills to enhance adult learning and ensure proficiency in the technology-facilitated delivery of the training.

Finally, consistent with FIP's approach of providing additional training for command staff, Tools for Tolerance® offers three leadership development programs designed for this level. Two of these are limited to California and are POST-certified. One program aims to provide command staff with strategies to support personnel who have completed other programs in cultural diversity and/or racial profiling. Central themes relate to the changing level of diversity in society, contemporary factors that are instigating or impeding this change, and challenges associated with intolerance as well as tolerance. The program emphasizes the nature and role of trust in active leadership and encourages participants to reimagine themselves as educators. Graduates may go on to complete the second program, which provides more in-depth, active leadership training. This advanced three-day program emphasizes leading by modeling personal values rather than from the status associated with one's position alone, negotiating difficult conversations productively, and building trust, similar to FIP. The last program is for those in the rank of lieutenant or above and is national in scope. For three days, command staff from around the United States meet for ethics-based discussions on challenges they are interested in addressing. For example, the meeting may center on problems command staff are facing with building trust with minority communities. Unlike FIP, however, topics may address other issues unrelated to bias. Command-level staff in the Tools for Tolerance® program explore both past and current leadership models and focus on translating theory into practice.

The Tools for Tolerance® programs and FIP differ in a few further noteworthy regards. First, Tools for Tolerance® does not seem to explicitly address procedural justice, although issues related to community perceptions of police legitimacy are recognized. Also, unlike FIP, Tools for Tolerance® does not include a community training component. However, the Museum of Tolerance and the New York Tolerance Center aim to achieve the goal of community education around issues pertaining to diversity and tolerance. Finally, whereas FIP offers participants a number of actionable strategies for reducing and managing implicit bias (i.e., increasing positive contacts with counter-stereotypical group members, conducting self-checks, reducing ambiguity), it is unclear whether Tools for Tolerance® educates participants about such tactics.

Tactical Perception

The National Initiative provided a three-part training program that includes training focused explicitly on implicit bias. Although FIP is a stand-alone program, like FIP, the National Initiative's "Tactical Perception" program was developed by a collaborative group of academic researchers and law enforcement agents specifically for

use with police officer participants. It shares in common with FIP the aim of teaching police officers about implicit bias stemming from not only race but also other social identities (e.g., gender, sexuality); how bias could affect officers' perceptions and decision-making; and how officers can manage bias. Likewise, central themes include the importance of ensuring positive interactions with the community as well as enhancing officer safety and effective performance. As the third part of the National Initiative's training, the Tactical Perception training followed two trainings that concentrated on procedural justice.

More specifically, across four modules administered over the course of eight hours, facilitators emphasize that when bias occurs, it is usually not a product of officers' explicit prejudicial attitudes but rather of historical and situational pressures. In the first module, facilitators introduce the concept of implicit bias and distinguish it from (a) stereotyping, i.e., overgeneralizing associations between specific characteristics and social group members, (b) prejudice, i.e., endorsement of stereotypes or negative feelings toward a group, (c) discrimination, meaning the differential treatment of group members, and (d) various "-isms" (e.g., racism, sexism) that entail systemic disadvantaging of some groups relative to others. The emphasis within this module is on the idea that situations may trigger automatic associations and mental short-cuts that can impact behavior, but conscious awareness of these phenomena will enable officers to react to civilians and situations in ways that are more in line with their personal beliefs and professional goals. The second module examines situational constraints that may produce "fast traps," in which officers' decision-making may be influenced by implicit bias that occurs quickly, automatically, and outside of conscious awareness. The third module dissects the problem of "slow traps" which are generally conscious negative responses to threats to one's identity (related, e.g., to a stereotype, masculinity, or authority) and may be experienced by either civilians or police officers.

In the fourth and final module, facilitators emphasize the importance of relying on training, experience, knowledge, and instincts while also recommending a number of specific strategies to "defuse traps" and thereby reduce bias and discriminatory behavior. These strategies include (a) changing situations to reduce risk of biased policing; (b) taking time to reflect on situational pressures that may affect perceptions and decision-making; (c) viewing others as individuals rather than stereotypical group members; (d) gaining exposure to counter-stereotypical group members and replacing stereotypes with non-stereotypical responses; (e) using procedurally just policing based on voice, neutrality, respect, and trustworthiness to increase perceived legitimacy; (f) calling for back-up when authority is threatened to avoid experiencing a slow trap; (g) holding oneself accountable by anticipating situations and planning, evaluating whether perceptions and decisions are being influenced by bias, examining consistency and patterns in decision making, and maintaining transparency by explaining decisions and

behavior; and (h) contributing to a culture of accountability within one's department. Thus, in line with FIP, the Tactical Perception program encourages officers to (a) reduce their biases, using Contact Theory as a guiding framework for doing so; (b) manage bias through self-accountability; (c) recognize situational constraints (e.g., time pressures, ambiguity) that increase the risk of bias; and (d) maintain awareness and make informed decisions.

Tactical Perception program facilitators are all police officers, which is in line with the model employed by FIP. They are trained via six lessons ("Traps Academy") that are delivered via web or downloadable PowerPoint files with accompanying audio recordings. These officers also are provided access to a listserv where they may ask questions or provide feedback regarding the trainer training. Facilitators then receive a guide and slides with which to deliver the eight-hour training session.

The Tactical Perception training program, much like FIP, presents scientific evidence regarding implicit bias. It further highlights the relevance of the material to officers through guided discussions, videos, activities for groups of four to six officers, facilitators' contribution of anecdotes and viewpoints, and participants' reports of experiences and reflections. Facilitators are encouraged through training and supporting materials to model procedural justice by giving participants voice while also correcting misperceptions and ensuring the discussion remains positive and productive. FIP also encourages the use of procedurally just principles, but Tactical Perception uses more explicit language to this point. In addition, however, Tactical Perception trainees first complete a race implicit association test, discussing their reactions while completing it and subsequent to receiving feedback regarding their implicit attitudes. They then learn about the science underlying the test and the relevance of the activity to the training session

National Training Institute on Race and Equity

The National Training Institute on Race and Equity (NTIRE) also offers relevant training through its "Implicit Bias and Cultural Competence" program. Although NTIRE's training is unique in that its educational approach is based in part on the teachings of Martin Luther King, Jr., like FIP, it is grounded in data and social science research emphasizing the universality and deep-seated, seemingly immutable nature of implicit bias. The primary focus of NTIRE is on educating participants about implicit *racial* bias; where other social identity biases are considered, they are examined only in the context of their intersections with race. The goal is to reduce the likelihood that bias will translate into discriminatory behavior among training participants.

Central components of NTIRE's training program define implicit bias, explain its root causes, and teach participants to understand how to assess implicit bias and

identify it in daily practice. This maps well to FIP's effort to impart skills related to managing one's own bias and being aware of bias within one's colleagues or community. Additional components of NTIRE's program address the possibility that implicit bias can lead perceivers to rely on stereotypes, develop prejudicial attitudes, and/or engage in discriminatory behavior. It also examines how targets may be affected by bias and its outcomes. Of importance, like FIP, NTIRE seeks to train individuals to understand how to manage bias both personally and within their institutions. Whereas FIP prescribes remedies based on having positive or counter-stereotypical contacts with racial minority individuals or performing self-checks and deliberate analysis, NTIRE seeks to reduce the effects of bias by teaching participants the "CUE Model," which emphasizes the need for participants to Communicate with, Understand, and Empathize with those who are different from themselves. Thus, where FIP includes training elements that are designed to reduce levels of implicit bias in the first place, NTIRE directs attention at perspective-taking across social divides to generate respect and inclusion, even in the event that diverse views are not reconciled.

Training is delivered to police officers, chiefs, and executives and other law enforcement, courts, correctional, and government professionals throughout the United States via conference presentations as well as half-day or full-day sessions. Both the half-day and full-day training incorporate national data and research findings; interactive exercises; small group discussions centered on "working with males of color," videos, images, and other relevant media; and anonymous audience polling and sharing of results. The longer full-day training tackles topics in greater depth and additionally includes participant completion of an implicit bias measure and discussion of results, as well as discussion of potential remedies and challenges and solutions related to implementing those remedies. The training is facilitated by an academic principal trainer, but it is unclear whether police officers may also serve as trainers as with FIP. The optimal size of training groups is 30 to 50 participants, but the training can be delivered to as few as 20 or as many as 200 participants.

Counter-Bias Training Simulation

Counter-Bias Training Simulation**,** also known as CBT Sim, was developed by Assistant Professor Lois James and Assistant Research Professor Stephen James at Washington State University to increase police officers' awareness about factors that may affect their decision-making in life-threatening encounters they may have with civilians. CBT Sim entails using a portable simulator to repeatedly expose police officers to realistic scenarios with "suspects" whose demographic characteristics (i.e., age, gender, race, and socioeconomic status) are not reliably related to the actual level of threat that they present (i.e., whether they are armed or unarmed). The simulator

projects these scenarios onto a screen and officers are tasked with deciding whether or not to shoot the suspects with guns that have been modified for the task. Following the simulation, officers participate in debriefing and self-evaluation sessions in which they learn to identify and understand the factors that shaped their decisions. The goal is to reduce the likelihood that implicit bias related to various social identities will influence officers' ability to detect threat in dangerous real-world circumstances, which may teach broader lessons that extend to a wider range of police officers' experiences.

The researchers who developed CBT Sim are currently conducting a randomized controlled trial, funded by the National Institute of Justice, to ascertain the efficacy and effectiveness of this simulation-based implicit bias training. Specifically, in spring 2019, they began training a sample of 400 police officers in Cleveland, Ohio, who were randomly assigned to one of four experimental groups. The first group will receive no training; the second group will participate in CBT Sim; the third group will receive classroom training on implicit bias; and the fourth group will receive both CBT Sim and classroom training. This design will enable the researchers to determine whether, compared to others, officers who participate in CBT Sim engage in more equitable decision-making, are perceived as less biased by citizens and arrestees, and perceive the training to be more effective over time. The researchers hypothesize that the simulation-based training method will be more effective in achieving these outcomes relative to the classroom-based training method, but it is unclear what the classroom-based training entails. Thus, we are unable to compare their classroom-based training directly to FIP.

Interaction and Perception

Spokane County Sheriff's Office provides its officers with implicit bias training via a program they refer to as Interaction and Perception.[45] The program integrates implicit bias training with the Defense Advanced Research Projects Agency's Strategic Social Interaction Modules (DARPA'S SSIM), which were developed to provide military personnel with skills for assessing and interacting with culturally different people and situations. As such, it aims to improve officers' ability to appreciate and identify biases and their potential effects on decision making to ensure these processes do not interfere with their work. It is unclear to what extent the program focuses specifically on race-based bias, but the pairing with DARPA's SSIM suggests that it may take a broader aim at sensitizing officers to many social identity biases. Like FIP, Interaction and Perception acknowledges that bias is the product of basic psychological mechanisms common to all human beings but which may be exacerbated by personal experiences.

[45] Details retrieved from https://www.spokanecounty.org/CivicAlerts.aspx?AID=566&ARC=919 and https://www.firstforward.com/Marketplace/Detail/3777b186-a69d-11e6-b170-001b219f8cb3.

Overall, the program aims to use bias training to ensure that officers approach civilians armed with knowledge about bias and prepared to manage their encounters as they relate to policing, conflict resolution, de-escalation, surveillance, and interviewing and interrogating. As such, it encourages officers to evaluate those civilian encounters after the fact to consider what they did well to address bias as well as what they might do better in the future to improve the accuracy of their decision making, ensure fair interactions with all social groups, and enhance community trust in police.

As with CBT Sim, only a broad overview of the program is publicly available, precluding a detailed comparison of the pedagogy of Interaction and Perception in relation to FIP. Interaction and Perception is delivered over the course of 12 hours, which is equal to the FIP training received by command-level staff but 50% longer than the FIP training other officers receive. The training is optimized for 16 to 22 students who may have had any level of prior training or relevant experience, with three instructors who, like facilitators of FIP and Tools for Tolerance®, are required to have completed a train-the-trainer module. It is described as being a participatory course that involves classroom training as well as opportunity for practice in interactions in the community—this latter facet stands out as particularly unique from FIP.

Comparisons

Whereas FIP and the five comparison programs reviewed herein similarly emphasize the social justice issues caused by biased policing, and they all use critical discussion as a central curricular feature, they vary in a number of respects:

- their theoretical frameworks (e.g., FIP relies on contact theory whereas NTIRE uses the CUE Model);
- the extent to which they focus on bias related to race versus other social identities (e.g., gender in the Tactical Perception program); and
- their attention to personal values and responsibility (more in Tools for Tolerance®) and skilled intuition (more in Tools for Tolerance® and the National Initiative's Tactical Perception program) versus unconscious processes, perceived legitimacy of police, and the safety of both community members and officers (more in FIP and the Tactical Perception).

The other programs incorporate some training techniques that the FIP training does not: Tools for Tolerance additionally uses museum exhibits; the National Initiative's Tactical Perception program and NTIRE's program use the implicit attitude test as an educational intervention; CBT Sim allows officers to practice decision making in a specially designed simulator, and Interaction and Perception offers officers the opportunity to practice their newly gained skills in community interactions and

emphasizes probing for bias in encounters after they occur.[46] Other structural formatting issues that vary across programs relate to other training that may be offered in tandem with the bias-focused training component (e.g., FIP, Tactical Perception, and Interaction and Perception each address procedural justice, whether implicitly or explicitly, whereas NTIRE and CBT do not).

It is unknown whether these differences in theoretical or philosophical approaches, content, and pedagogical tools are differentially successful in garnering police officers' receptivity and responsiveness to training. For example, FIP, Tactical Perception, NTIRE, and Interaction and Perception emphasize how officers may exhibit bias unconsciously even if they do not personally endorse prejudicial beliefs. On the one hand, this approach may alleviate the perception that officers are being blamed for engaging in bias. On the other hand, however, it could lead officers to feel as though they cannot and should not be held responsible for automatic psychological processes that happen outside of their conscious awareness. In contrast, the Tools for Tolerance® values-based approach may not resonate well with police officers, as it suggests that they may have misguided values. These are hypotheses that could be empirically tested with future research that disaggregates the features of these training programs. Such work will be needed to determine how to design implicit bias training to be maximally effective.

As it stands, evaluation findings are currently publicly available only for Tactical Perception, and those findings bear only on the immediate effects of the training on officers' knowledge of the concepts. There is, then, no empirical basis on which to consider whether – and if so, in what respects – any of these programs are better than others in enabling officers to manage their unconscious biases or in reducing biased behavior. Therefore, we are currently able only to compare how FIP approaches this goal compared to a few other similarly targeted programs, as well as whether FIP uses interventions that have proven to be effective in the scholarly literature.

Outcomes of Implicit Bias Training

The most proximate outcomes of the FIP training, we presume, are officers' beliefs about implicit bias and procedural justice, and their implications for police work – that is, trainees' grasp of the substantive content of the training. This knowledge and

[46] The rationale for having trainees complete the Implicit Association Test is that it demonstrates the pervasiveness of bias even in people who think they do not have them. If taking the IAT is a sufficiently powerful teaching tool to result in greater shifts in knowledge, motivation, concern, etc., relative to people to who do not complete the IAT, then one would expect to see greater efficacy in recognizing and managing bias among trainees who take the IAT. Lorie Fridell reports that "FIP considered and piloted having trainees take an implicit association test as part of its program, but abandoned the technique after finding it did not enhance the training." Personal communication, May 8, 2020.

awareness may be a necessary, but not sufficient, condition for changes in officers' enforcement practices. Officers' attitudes toward bias and discrimination may also play an important part in shaping their behavior, and these attitudes might also be affected by the training. Officers' views about racial discrimination as a social problem, and their personal motivation to act without prejudice, might amplify the effects of the training; moreover, the training may raise the salience of the issue for trainees and thus affect these attitudes directly. We describe these outcomes in conceptual terms here; we describe our measures of these outcomes as part of the evaluation design.

As we noted in the introduction, disparities in the outputs of policing are commonly found in the U.S., and implicit bias is one of the factors thought to contribute to such disparities. Training in implicit bias may be offered with a view toward reducing enforcement disparities, and we may therefore regard officers' enforcement behavior as an important potential outcome of implicit bias training. Police behavior is shaped by many forces, however, and the impacts of an 8-hour training should be considered in the context of those influences. We discuss these considerations.

Attitudes

Motivation and concern are foundational concepts that have been established in the literature as theoretically instrumental. We briefly review that literature.

Motivation to Respond without Prejudice

In characterizing the forces that regulate an individual's biases, many researchers have referred to "motivation" as one behavioral crux. This construct has been variously conceived as a factor reliant upon situational contexts (such as an audience), a factor vulnerable to normative pressures, or a factor derived entirely from within the individual. In other words, motivation can be construed as internal or external – sprung from one's own personal standards or morals, or dependent on one's perceptions of others' standards and morals.

One of the first treatments of "motivation" as a key mechanism in the attitude and behavior equation was Russell Fazio's creation of The Motivation and Opportunity model, or MODE model, in 1990.[47] For his conceptualization, Fazio built to some degree on the "fear of invalidity," or fear of saying the wrong thing, as a primary motivator for judgement and decision-making. Kruglanski and Freund determined that this "fear" mitigated individuals' reliance upon stereotypes and initial impressions when making

[47] Russell Fazio, "Multiple Processes by Which Attitudes Guide Behavior: The MODE Model as an Integrative Framework," *Advances in Experimental Social Psychology* 23 (1990): 75-109.

probability judgements, instead spurring them to more carefully process information.[48] Thus, in constructing the MODE model, Fazio cast "motivation" as a factor largely dependent upon an audience, whether present or theoretical. He examined both deliberative and spontaneous processes in efforts to disentangle and delineate the relationship between attitudes and behavior, concluding that "because the perceived costliness of the potential behavior motivates the individual to exert cognitive effort, the degree to which the individual's attitude toward the object is capable of automatic activation from memory becomes irrelevant to the behavior decision process."[49] In the following year, Devine, Monteith, Zuwerink, and Elliot distinguished the two types of motivation, describing motivation driven by external pressures, or "standards derived from others' expectation for how one should respond," as a weaker stand-in for "well-internalized personal standards."[50] The authors posited and provided evidence for an inverse relationship between prejudice and measures of personal values.

Dunton and Fazio devised the Motivation to Control Prejudiced Reactions (MCPR) scale in 1997, basing their theory in large part on Fazio's 1990 MODE model.[51] This scale sought to determine the degree to which individuals exerted effort to control and check prejudicial displays. The scale consists of 17 items that query motives for behavior, such as "I feel it's important to behave according to society's standards," and "I think it is important to speak one's mind rather than worry about offending someone."[52] In creating items for the scale, the authors considered a number of factors that might influence behavior, such as concern for self-presentation, norm conformity, sensitivity to others' feelings, and "internalized personal standards."[53] As Plant and Devine later noted, the inclusion of this final construct ("internalized personal standards") in creating survey items may have muddied subsequent analyses by conflating internal and external operations. In a footnote, Plant and Devine state that "the ambiguity of the items … may

[48] Arie W. Kruglanski and Tallie Freund, "The Freezing and Unfreezing of Lay-Inferences: Effects on Impressional Primacy, Ethnic Stereotyping, and Numerical Anchoring," *Journal of Experimental Social Psychology* 19 (1983): 448-468.
[49] Fazio, "Multiple Processes," 93.
[50] Patricia G. Devine, Margo J. Monteith, Julia R. Zuwerink, and Andrew J. Elliot, "Prejudice With and Without Compunction," *Journal of Personality and Social Psychology* 60 (1991): 817-830, p. 824. Also see Ashby E. Plant and Patricia G. Devine, "Internal and External Motivation to Respond without Prejudice," *Journal of Personality and Social Psychology* 75 (1998): 811-832. Scale Items for these personal standards more closely resemble what Plant and Devine classify as Internal Motivation, such as "how important it is to you to respond to gays in ways that are consistent with your personal standards?" or "How committed are you to trying to respond consistently with your own personal standards?"
[51] Bridget C. Dunton and Russell H. Fazio, "An Individual Difference Measure of Motivation to Control Prejudiced Reactions," *Personality and Social Psychology Bulletin* 23 (1997): 316-326.
[52] Ibid., 319.
[53] Ibid., 318.

have obscured real differences between internal and external sources of motivation to respond without prejudice."[54]

Plant and Devine's 1998 research demonstrated the predictive validity of both Internal and External Motivation measures, thereby establishing the two types of motivation as discrete. The External Motivation Scale (EMS) captured high scores for respondents whose motivations were driven by external approval or to satisfy norm expectation (e.g., "It is important to me that other people not think I'm prejudiced").[55] The Internal Motivation Scale (IMS) captured high scores for those whose motivations stemmed from deep personal values concerning race and discrimination (e.g., "I get angry with myself when I have a thought or feeling that might be considered prejudiced").[56] The authors expanded on the importance of such a distinction in 2009:

> Distinguishing between the intention to hide prejudice versus the intention to be free of prejudice is important because the existence of these alternative intentions is the fundamental reason that White's nonprejudiced self-reports are often met with suspicion and that the assessment of prejudice remains a difficult endeavor…. For minority group members, clarifying the intentions underlying nonprejudiced responses is essential for issues of trust and knowing what to expect from outgroup members when not under public scrutiny.[57]

Plant and Devine elaborated on their operationalization of internal motivation, describing it as a motivation "that gives rise to the intention to be free from prejudice altogether."[58] Separate internal and external motivation scales have since been employed in numerous studies.[59]

---

[54] Plant and Devine, "Internal and External Motivation to Respond without Prejudice," p. 812.
[55] Ibid.
[56] Ibid.
[57] E. Ashby Plant and Patricia G. Devine, "The Active Control of Prejudice: Unpacking the Intentions Guiding Control Efforts," *Journal of Personality and Social Psychology* 96 (2009): 640-652, p. 641.
[58] Ibid.
[59] See: Patricia G. Devine, Patrick S. Forscher, Anthony J. Austin, and William T. L. Fox, "Long-term Reduction in Implicit Race Bias: A Prejudice Habit-Breaking Intervention," *Journal of Experimental Social Psychology* 48 (2012): 1267-1278; Joshua Correll, Bernd Wittenbrink, Charles M. Judd, Bernadette Park, Melody S. Sadler and Tracie Keesee, "Across the Thin Blue Line: Police Officers and Racial Bias in the Decision to Shoot," *The Journal of Personality and Social Psychology* 92 (2007): 1006-1023; Calvin K. Lai, Allison L. Skinner, Erin Cooley, Sohad Murrar, Markus Brauer, Thierry Devos, Jimm Calanchini… Brian A. Nosek, "Reducing Implicit Racial Preferences: II. Intervention Effectiveness Across Time," *Journal of Experimental Psychology: General* 145 (2016): 1001-1016.

Concern about Discrimination

Fickle characterizations of the term "concern" in extant work prompts some conceptual clarification for the purposes of this evaluation: if motivation is a mechanism by which attitudes are conveyed, concern is the infrastructure upon which attitudes are raised.[60] As an exemplar of concern measurement, a number of researchers utilize self-reported *should* and *would* response discrepancies as a measure of conflict between hypothetical and actual behavior. For example, when provided with the scenario: "Imagine that a Black person boarded the bus and sat next to you," a person who stated that they *should* "feel comfortable" but that they *would* "feel uncomfortable" would generate a discrepancy in their *should-would* response. Devine et al. employed this type of measurement, and followed this analysis with an evaluation of the "affective consequences" that resulted from learning of their discrepancies, looking at both "diffuse and/or qualitatively distinct affects."[61] Upon the individuals' review of their *should-would* discrepancies, the authors measured global discomfort (uncomfortableness, anxiety) and more specific feelings of compunction (guilt, self-criticism). The researchers concluded that "personal standards," which we may interpret as "concern," are correlated to prejudice, and that the affective consequences are contingent on these values.

Perugini, O'Gorman, and Prestwich, in an evaluation of the construct validity of theories which support the Implicit Association Test (IAT), interpret *concern* as "strongly related to egalitarianism and implies a particular concern toward negative biases against historically disadvantaged groups such as blacks."[62] This politically rooted supposition is common amongst attitudinal research that allows for a consideration of "skepticism" as a subcategory of concern. One such early estimation of concern is John McConahay's Modern Racism scale (MRS).[63] The premise for this scale reflected a growing sentiment that "discrimination is a thing of the past, blacks are pushing too hard, they are getting

[60] Dunton and Fazio use the denomination "concern with acting prejudiced" when defining their measure of "motivation," describing it as being measured by "items that reflect being concerned about appearing prejudiced to others, ones that reflect a more private concern with observing oneself having prejudiced thoughts of feelings, and ones that reflect a personal standard regarding the avoidance of prejudiced and offensive expressions" ("An Individual Difference Measure of Motivation to Control Prejudiced Reactions," pp. 320-321). These items clearly include both internal and external measures, but also draw attention to the significant conceptual overlap between "internal motivation" and broader "concern." By defining motivation as being influenced beyond an assiduousness for other's or personal standards, and including a broader consideration of personal value systems, we see one example in which the conceptualizations of "concern" and "motivation" are tangled.

[61] Devine, Monteith, Zuwerink, and Elliot, "Prejudice With and Without Compunction," p. 818.

[62] Marco Perugini, Andrew Prestwich, and Rock O'Gorman, "An Ontological Test of the IAT: Self-Activation Can Increase Predictive Validity," *Experimental Psychology* 54 (2007): 134-147, p. 142.

[63] John B. McConahay, "Self-Interest Versus Racial Attitudes as Correlates of Anti-Busing Attitudes in Louisville: Is it the Buses or the Blacks?" *The Journal of Politics* 44 (1982): 692-720.

too much attention and sympathy from the nation's elites and that black's gains and demands are no longer justified."[64] Each item prompts respondents to indicate the degree to which they agree with the statement that emulated generalized lack of concern for discrimination in society, such as "Over the past few years Blacks have gotten more economically than they deserve."[65] Similar skepticism measures have been employed elsewhere as an auxiliary evaluation of Concern, such as in John Brigham's 1993 Attitudes Toward Blacks Scale ("I enjoy a funny racial joke, even if some people might find it offensive"), and select items of Devine et al.'s "Concern" scale ("People need to stop focusing so much time and energy worrying about racial discrimination").[66]

Devine, Forscher, Austin and Cox utilized a measure of concern in their 2012 development of a prejudice habit-breaking tool, predicated on the idea that unconscious bias forms as a habit, and can therefore be addressed like other habits.[67] One key component of this experiment was to educate the intervention group about bias in society. Therefore, it was necessary for researchers to administer a survey which included items querying respondent's concern over bias in society, proposing that "people must be *aware* of their biases and, second, they must be *concerned* about the consequences of their biases before they will be motivated to exert effort and eliminate them."[68] The authors urged for further research on concern as a foundational construct, and caution that "education may play a specialized role in increasing awareness and concern, but both education and training may be necessary to produce changes in implicit bias."[69] In a replication of this intervention, Forscher, Mitamura, Dix, Cox, and Devine defined concern as "the extent to which a person believes discrimination toward Black people is a serious problem in society."[70] The authors conclude that any movement of this construct is related to "a broad range of psychological processes related to one's orientation toward oneself and the social environment."[71]

---

[64] Ibid, p. 707.

[65] Ibid, p. 708.

[66] John C. Brigham, "College Student's Racial Attitudes," *Journal of Applied Social Psychology* 23 (1993), p. 1940; Patricia G. Devine, Patrick S. Forscher, Anthony J. Austin, and William T. L. Cox, "Long-Term Reduction in Implicit Race Bias: A Prejudice Habit-Breaking Intervention," *Journal of Experimental Social Psychology* 48 (2012), p. 1277.

[67] Devine et al., "Long-Term Reduction in Implicit Race Bias."

[68] Ibid., p. 1268. The NYPD evaluation utilized the following item from Devine et al.'s item in the scale for concern: "I consider racial/ethnic discrimination to be a serious social problem," and their following item in a scale of skepticism: "People need to stop focusing so much time and energy worrying about racial/ethnic discrimination."

[69] Ibid., p. 1277.

[70] Patrick S. Forscher, Chelsea Mitamura, Emily L. Dix, William T. L. Cox and Patricia G. Devine, "Breaking the Prejudice Habit: Mechanisms, Timecourse, and Longevity," *Journal of Experimental Social Psychology* 72 (2017): 133-146.

[71] Ibid.

Beliefs

For many of the training programs reviewed above, officers' knowledge about and awareness of implicit bias, as well as their understanding of its implications for police work, are arguably the fulcrum on which other training effects rest. If training fails to enhance trainees' comprehension of the science of bias, and their recognition of how their enforcement judgments can be influenced by unconscious bias, then it is surely unlikely to affect their performance. These outcomes are "beliefs" – outlooks that are largely or exclusively cognitive in nature – as opposed to the more affective attitudes of, e.g., concern about discrimination and motivation to respond without prejudice. The latter involve value judgments, which are not inherently correct or incorrect; beliefs relate directly to facts.

For the FIP training, the facts in question concern the science of bias and its application to policing:

- Unconscious biases stem from exposure to social stereotypes, such that anyone exposed to the stereotypes is susceptible to holding them, even if they consciously reject the stereotypes.
- Consequently, well-intentioned people have such biases, and that they have unconscious bias does not mean that they are racists.
- However, unconscious bias can affect one's perceptions and behavior, even without one's awareness.
- For police, that implies that officers could take – or fail to take – actions based on implicit biases: they could be over- or under-vigilant, with unwanted effects on the community and the officers.
- Unconscious biases cannot be easily changed, but they can be managed and controlled to reduce their effects.

The FIP training also covers procedural justice, citing it as a "cornerstone to fairness and impartiality." The tenets of procedural justice naturally complement FIP training objectives, and enrich the body of knowledge from which members may pull.

FIP is a knowledge-based training – that is, one that addresses subconscious processes by arming the trainee with relevant skills and information. Forscher and Devine cited the efficacy of such an approach:

> The optimal process to target may thus be one that is moderately central to how people view themselves. One possible candidate is knowledge. People become invested in certain views of themselves and the world … and yet knowledge does change in response to evidence.[72]

[72] Patrick Forscher and Patricia G. Devine, "Knowledge-Based Interventions Are More Likely to Reduce Legal Disparities Than Are Implicit Bias Interventions," in Sarah E. Redfield (ed.), *Enhancing Justice: Reducing Bias* (Chicago: American Bar Association, 2017), pp. 303-316.

Behavior

In the U.S., and especially in its racially and ethnically diverse cities, people of color are disproportionately represented among those whom police stop and investigate, and against whom police invoke the law. They are also disproportionately those against whom lethal force is used, which was catapulted onto front pages in the last five years. To some (unknown) degree, disparities in enforcement outputs stem from disparities in the behavior of people with whom police interact. However, the magnitude and consistency of disparities, and the inability to rule out bias as one explanation for disparities, leads many to attribute disparities partly to modern racism – unconscious bias – as we discussed above, and to prescribe implicit bias training as a partial antidote.

Insofar as implicit bias leads officers to form suspicion based on the actions of, e.g., Blacks more than the same actions of Whites (as Judge Scheindlin opined), then effective training could be expected to reduce disparities in stops, at the margin. Insofar as implicit biases inflate the perceived threat posed by Blacks or Hispanics, then effective training could be expected to reduce disparities in frisks and in the use of force. Insofar as people of color are unconsciously associated with various types of crime, then effective training could be expected to reduce disparities in searches, at the margin. Proportionately fewer stops of people of color would likely eventuate in reduced disparities in arrests and summonses as well. Moreover, insofar as officers' judgments about future offending influence their decisions to invoke the law, especially for low-level offenses, and are influenced by unconscious bias, then effective training could be expected to reduce disparities in summonses and (low-level) arrests.

All of these behaviors can be measured in terms of administrative records of enforcement "outputs" (though underreporting is a source of measurement error), and disparities in these enforcement outputs can be assessed. Bias, as such, need not be estimated.

Though neither the FIP executives nor the NYPD set behavioral change as a training objective, we believe that a thorough assessment of implicit bias training for police would include enforcement behaviors among the outcomes on which training impacts are estimated. On this we agree with Cynthia Lum and her colleagues, whose assessment of the evidence bearing on the recommendations of the President's Task Force on 21st Century Policing concluded that, "Rigorous evaluations of implicit bias training are needed. These studies ideally would use randomized designs and follow-up assessments to examine the long-term impacts of training on officers' knowledge, attitudes, and behavior."[73] However, we recognize that training is but one influence on

[73] Cynthia Lum, Christopher S. Koper, Charlotte Gill, Julie Hibdon, Cody Telep, and Laurie Robinson, *An Evidence-Assessment of the Recommendations of the President's Task Force on 21st Century Policing —*

how police officers do their jobs. The impacts of training are likely to turn to a large degree not only on the design and delivery of the training, but also on the extent to which the training is reinforced by other organizational influences, and not contradicted by some organizational forces.

Organizational Context

In *Producing Bias-Free Policing: A Science-Based Approach,* Lorie Fridell emphasizes that the FIP training is not expected to affect some trainees, notably: officers without motivation to act without bias, and officers who exhibit explicit biases. Therefore, a broader and more suffused organizational approach to promoting, sustaining, and institutionalizing the lessons of FIP training in police agencies is necessary to ensure both (a) the organizational commitment to FIP is established and formalized, and (b) the agency commitment to FIP extends beyond the officers who are impacted by FIP training.[74] One way in which these lessons are thoroughly integrated into the agency context is through supervisory training, which provides supervisors with tools to detect and skillfully respond to officers who may be engaged in biased policing, as well as routinely reinforce FIP lessons to officers under his or her command beyond the training. Another way to insinuate the lessons of FIP training into the agency environment is to integrate them into academy and in-service training curricula. Fridell outlines several ways in which the science of bias can be incorporated into use-of-force training, and illustrates how integrating FIP lessons within other training contexts can deepen the organizational impact.

Importantly, Fridell addresses the relative advantages and shortcomings associated with efforts to measure biased policing with activity data, such as vehicle stops, detentions, and searches. She further stresses the importance of carefully planning and vetting the processes by which agencies review and analyze these data, noting "Some … disparity may be produced by biased policing; some of that disparity may reflect other, legitimate, factors."[75] Properly and thoroughly vetted accountability measures available to agencies (such as body-worn cameras, early intervention systems,

---

Implementation and Research Priorities (Fairfax, VA: Center for Evidence-Based Crime Policy, George Mason University; Alexandria, VA: International Association of Chiefs of Police, 2016), p. 38. Also see Renée J. Mitchell and Lois James, "Addressing the Elephant in the Room: The Need to Evaluate Implicit Bias Training Effectiveness for Improving Fairness in Police Officer Decision-Making," *The Police Chief* (2018), https://www.policechiefmagazine.org/addressing-the-elephant-in-the-room/. They pose, as "the critical question": "Can implicit bias training reduce police officer bias, improve officer fairness in behavior, and ultimately promote public trust in police legitimacy?"

[74] Lorie Fridell, *Producing Bias-Free Policing: A Science-Based Approach* (Springer Briefs in Criminology, 2017), pp. 31-94.

[75] Ibid., p. 66.

complaint review systems, supervisory monitoring, and employee evaluations) may be utilized to check individual and collective fidelity to the policies in line with FIP objectives. Agencies should, Fridell maintains, also develop corrective measures that can be administered as needed. Fridell further discusses the importance of the agency's operational compliance with the bias-free policing agenda, specifically noting high-discretion crime control operations as uniquely susceptible to the effects of unconscious bias. She concludes: "… police need to adopt focused, information-led strategies that target behaviors and not populations, and those strategies should be adopted with the cooperation and support of community members …."[76] These strategies should also be considered with regards to how they might be perceived within the community.

One crucial element of the FIP agenda is executive and command-level training, in which community members may be optionally included. Like the training for patrol, first-line supervisors, and mid-management, executive and command-level training discusses the science of implicit bias and its significance to policing, however differs from the other trainings in a few key respects. Executive and command-level trainings importantly feature discussions of "assessment and reform in the realms of policy, recruitment and hiring, training, supervision and accountability, leadership, measurement, and outreach to diverse communities."[77] As part of this broader organizational undertaking, command-level training is designed to give commanders tools "to implement various strategies that promote bias-free policing."[78] To this end, the FIP founder describes several actions that leadership can take. Fridell first discusses the role of departmental leadership in curating an executive tone for the messaging of bias-free policing initiatives to both officers and citizens. She encourages leadership to put forth a "nuanced message" that does not invalidate the community's perceptions, nor alienate or impugn the character of their officers. Department leaders should, she explains, frame discussions around the legitimate Constitutional mandate for their officers to commit to bias-free enforcement, and assure their communities of the seriousness with which the agency is undertaking that charge. She notes that this type of format "neither affirms nor denies the existence of biased policing."[79] She further recommends folding in accolades and reviews for bias-free policing in personnel evaluations in an effort to institutionalize the training, additionally evaluating supervisors' "attentiveness to, and handling of, biased policing concerns."[80]

Fridell next discusses agencies' policies regarding bias in policing. Though most agencies have some policy regarding bias-free policing, many are outdated and vague.

[76] Ibid., p. 85.
[77] Ibid., p. 6.
[78] Ibid., p. 30.
[79] Ibid., p. 32.
[80] Ibid., p. 33.

An emphasis on current and specific directives regarding the use of demographics in policing is not only important for educating officers, she maintains, but also vital to fixing the imprimatur of leadership. Some suggestions for policy models that outline parameters for the appropriate use of demographic factors in policing decisions are delineated as follows:

- Using information that is *locally relevant*
- Linking specific types of crimes to certain demographics
- Using information that is from a credible source
- Using information that is temporally relevant (i.e., "not stale")
- Using demographic variables as they relate to the totality of the circumstances.

Agencies are also advised to refine their policies and practices regarding recruitment, hiring, and promotion in order to ensure that they accomplish their stated goals (in both creating a diverse police force and thoroughly vetting prospective officers). In order to avoid bias in the processes of hiring, promotion, or review of officers, Fridell cites the recommendations of Fiske and Kreiger as guiding principles:

- Use data to monitor diversity of the agency
- Ensure that the agency's diversity is made a priority
- Ensure that those making hiring and promotional decisions are likewise trained in implicit bias awareness, and
- Outline clear parameters for hiring and promotional decisions.[81]

Departments are also directed to encourage officers to engage in programs that aim to bridge gaps between officers and communities. Fridell notes a number of programs that involve more foot patrol and invite community interaction with officers, such as "Park, Walk, and Talk" in St. Petersburg, "Cops Out of Cars" in Minneapolis, or "Coffee with a Cop," adopted by numerous agencies. More target-specific programs are also encouraged, such as those that promote outreach to the LGBTQ communities or to at-risk youth.

Extant Evaluative Evidence

To date, only two studies of the impacts of implicit bias training for police have been completed. One study was undertaken as part of the Urban Institute's evaluation of the National Initiative to Build Community Trust and Justice (hereafter the National Initiative).[82] Implicit bias training – Tactical Perception, described above – was one of

[81] Susan T. Fiske and Linda H. Krieger, *Policy Implications of Unexamined Discrimination: Gender Bias In Employment as a Case Study* (Princeton, NJ: Princeton University Press, 2013).

[82] Jesse Jannetta, Sino Esthappan, Jocelyn Fontaine, Mathew Lynch, and Nancy LaVigne, *Learning to Build Police-Community Trust* (Washington, DC: Urban Institute, 2019). Results varied from one department to the next, but the differences were only noted and not discussed further.

three trainings delivered in the six sites. "Learning assessment surveys" were administered before and after the implicit bias training, eight items on which were designed to measure officers' understanding of "core concepts." The items were confined to matters of training content, that is, what we characterize above as beliefs; the survey did not extend to, e.g., concern about discrimination or motivation to respond without prejudice. Across more than 7,000 respondents, researchers found statistically significant mean differences, before and after training, in seven of the eight items. The largest difference on the 1-to-5 response scale, however, was 0.31 (increasing from 3.78 to 4.09). Other differences that achieved statistical significance ranged from 0.06 to 0.29 along the 1-to-5 scale. The conclusions to be reached on that basis were admittedly limited; "survey results cannot speak to how lasting any changes in attitudes and knowledge were or how they may have manifested in officer behavior."[83]

A second study was an evaluation of Fair and Impartial Policing training in an eastern Canadian police service.[84] Pre- and post-training questionnaires were administered to trainees. One pattern that emerged from the analysis was that self-assessed knowledge about implicit bias concepts increased while enthusiasm about the training and its applicability to policing decreased. For example, in the post-training questionnaire, trainees were more likely to report that they fully understood the concept of implicit bias, and could analyze and recognize the potential impact of bias in unique situations. They were less likely to agree that the training would help them be more effective at their job. An analysis of "knowledge-check items" – assessing trainees' understanding of FIP concepts – showed that trainees' comprehension of the training content was limited, with a mean score slightly less than 50 percent.

Insights from Social Psychology

Beyond evaluations of implicit bias training for police, some clues to the promise of FIP for enhancing police officers' awareness of implicit bias, improving their ability to prevent implicit bias from affecting their behavior, or even reducing implicit bias can be found in extant research; we review that evidence next. We consider research not only on interventions designed to address implicit bias and/or its behavioral consequences,

[83] Ibid., p. 22.

[84] Chris Giacomantonio, "Fair and Impartial Policing at Halifax Regional Police: Evaluation of Impact on Attitudes and Knowledge," presented at the conference of the American Society of Evidence-Based Policing, Cincinnati, May, 2019. That the training was delivered by the agency's own personnel, trained by FIP in its Training-of-Trainers (TOT) Program, may be a relevant consideration (Lorie Fridell, personal communication, May 8, 2020). In any evaluation of a training, one should treat the fidelity with which the trainers deliver the curriculum as an important feature of implementation. We are aware of no evidence about the fidelity with which this training was delivered.

but also on a somewhat broader and diverse range of training methods and programs, which we believe may provide some insight into the virtues and limitations of the FIP training. We acknowledge, and readers should note, that some of the training programs that we consider – such as diversity training programs – have a wide variety of goals and content. We consider them because many of the distinct program characteristics overlap with those of FIP, and findings on their effects may be instructive in understanding the effects of implicit bias awareness training programs.[85]

Bias Awareness and Management

One of the general aims of FIP is to educate officers about the science of implicit bias and how the phenomenon can negatively impact community and officer safety as well as community perceptions of fairness and legitimacy. Bezrukova, Spell, Perry, and Jehn's meta-analysis of diversity training evaluations revealed that diversity training is associated with immediate and stable increases in knowledge related to cultural diversity.[86] The same is likely true with regard to training focused specifically on implicit bias, although empirical assessment is needed, to be sure.

However, the question remains as to how knowledge about implicit bias subsequently impacts officers' decision-making and behavior. If we presume that FIP serves to warn officers about the deleterious effects of implicit bias, we can perhaps draw from social psychological research to speculate about the answer to that question. For instance, Axt and Lai recently tested whether a bias warning could reduce discriminatory decision-making.[87] Specifically, they alerted participants that decision-makers are biased in favor of physically attractive people and asked them to try to avoid exhibiting that bias while deciding whether to accept applicants who varied in attractiveness into a hypothetical academic honor society. The researchers found that participants who received the warning exhibited significantly less bias toward attractive applicants in their acceptance decisions relative to participants who received no such warning.

[85] Diversity training is focused on teaching the importance of collaborating with diverse voices for organizational functions (e.g., increased creativity) as well as the importance of being inclusive in activities like hiring, promoting, etc. Though the goals and content of diversity training and implicit bias training differ in a number of respects, both forms of training are directed toward reducing bias and discrimination in organizational contexts, so some of the lessons learned from studying diversity training might be relevant to understanding the potential effects of implicit bias training.

[86] Katerina Bezrukova, Chester S. Spell, Jamie L. Perry, and Karen A. Jehn, "A Meta-Analytical Integration of Over 40 Years of Research on Diversity Training Evaluation," *Psychological Bulletin* 142 (2016): 1227-1274.

[87] Jordan R. Axt, and Calvin K. Lai, "Reducing Discrimination: A Bias versus Noise Perspective," *Journal of Personality and Social Psychology*, advance online publication, 2019, study 5. Participants were drawn from the Project Implicit research pool.

Axt, Ebersole, and Nosek had previously used the same decision-making paradigm to test whether a pro-Black bias in acceptance decisions could be mitigated by instructing participants that academic decision-makers are easier on Black applicants and tougher on White applicants.[88] Again, participants who received the warning exhibited less bias in favor of Black applicants relative to control participants who did not receive the warning. Yet other participants who were advised to avoid anti-White bias or bias in general continued to discriminate in favor of Black applicants. These findings are consistent with other research showing that participants who received training that focused on rejecting negative stereotypical associations (e.g., Black–violent) actually had stronger automatic stereotype activation and more negative evaluations than they did prior to training.[89] Together, these studies suggest that even if training is effective at enhancing bias awareness, the extent to which that awareness subsequently affects behavior may depend on whether warnings are tailored specifically to the types of bias that officers are expected to guard against. For example, rather than explicitly instructing officers to avoid acting with anti-Black bias or bias generally, it may be more effective to caution them about managing a pro-White bias. (Though training content delivered at one point in time is not necessarily equivalent to warnings delivered in the immediate context of choice.) Given that FIP takes the former approach, it is unclear whether it will be effective at preventing bias from translating into discriminatory decision-making or behavior.

In addition to educating officers about implicit bias, FIP seeks to instill in officers the ability to consciously manage it. In particular, FIP encourages officers to evaluate their decision-making for bias. The promise of this tactic is supported by a meta-analysis examining relations between implicit and explicit measures of intergroup attitudes and stereotyping across 126 studies. Specifically, the more cognitive effort people exert to control their responses, the less likely their implicit bias is to manifest in self-reported racial bias.[90] Also, Forscher et al.'s meta-analysis revealed that interventions designed to affect participants' goals to weaken bias (e.g., by making anti-prejudiced norms salient before administering an implicit racial bias measure) reduced bias exhibited on implicit

[88] Jordan R. Axt, Charles R. Ebersole, and Brian A. Nosek, "An Unintentional, Robust, and Replicable Pro-Black Bias in Social Judgment," *Social Cognition* 34 (2016): 1-39.

[89] Bertram Gawronski, Roland Deutsch, Sawsan Mbirkou, Beate Seibt, and Fritz Strack, "When 'Just Say No' is Not Enough: Affirmation versus Negation Training and the Reduction of Automatic Stereotype Activation," *Journal of Experimental Social Psychology* 44 (2008) 370-377.

[90] Wilhelm Hofmann, Bertram Gawronski, Tobias Gschwendner, Huy Le, and Manfred Schmitt, "A Meta-Analysis on the Correlation between the Implicit Association Test and Explicit Self-Report Measures," *Personality and Social Psychology Bulletin* 31 (2005): 1369-1385.

and explicit measures.[91] If FIP effectively trains officers to control biased responding, it is likely to curb the effects of implicit bias on decision-making, as dismissing irrelevant demographic information reduces the frequency with which errors disproportionately favor one group over another.[92]

FIP also advises that officers use careful and deliberate assessments to disambiguate situations and prevent bias from filling in for uncertainty and affecting judgments. This strategy is likely to reduce the deleterious effects of implicit bias on officers' decision-making and behavior, too. Axt and Lai conducted a series of studies on this point.[93] Generally, their findings indicate that having additional time to focus on outcome-relevant information can reduce noise in decision-making by decreasing the total number of errors made. For example, in their Study 2a, the researchers found that as the length of time that participants had to complete a first-person shooter task increased, there was a corresponding linear decrease in the rate at which participants made errors. In fact, all participants exhibited a racial bias by (a) erroneously shooting unarmed suspects more often when they were Black versus White, and (b) correctly not shooting armed suspects more often when they were White. But reducing time pressure decreased this discriminatory behavior by decreasing the total incidence of errors. This study's findings are particularly notable considering that the difference between high and low time pressure was operationalized as a mere 160 milliseconds. The same pattern was found across several studies. The researchers also determined that noise in decision-making could be reduced by imposing a delay before allowing participants to record their decision about whether to accept applicants who varied in attractiveness into a hypothetical academic honor society (Study 3) and by instructing participants to engage in systematic rather than heuristic thinking (i.e., telling them to think hard and deliberately versus not to overthink and go with their "gut" response) (Study 4), although bias remained constant across conditions. In fact, discrimination was diminished by reducing both noise *and* bias only when participants on the academic decision-making task experienced a delay prior to registering their decision, which decreased the total error rate, *and* received the warning about bias described in detail earlier, which decreased the rate at which errors disproportionately favored attractive people over others. Thus, to the extent that FIP warns officers of the risks associated with implicit bias and trains officers to take the time to pay attention to objective behavioral and environmental cues while making judgments, it is likely to succeed in reducing the incidence of biased decision-making and behavior.

---

[91] Patrick S. Forscher, Calvin K. Lai, Jordan R. Axt, Charles R. Ebersole, Michelle Herman, Patricia G. Devine, and Brian A. Nosek, "A Meta-Analysis of Procedures to Change Implicit Measures," *Journal of Personality and Social Psychology* (2019).

[92] Axt and Lai, "Reducing Discrimination: A Bias versus Noise Perspective," studies 1a and 1b.

[93] Ibid.

Training Requirement

The volatility of the effects of diversity training programs has been clearly demonstrated by varying results across a number of studies, and the considerations in applying any type of diversity training are delicate and consequential. Frank Dobbin and Alexandra Kalev examined at length the application and outcomes of diversity training programs that have been employed in American companies over a 30 year period, finding that "companies do a better job of increasing diversity when they forgo the control tactics and frame their efforts more positively."[94] Research evaluating the outcomes of these programs found that in many cases diversity in management positions tended to decrease, especially when the training was mandatory, or "undertaken mainly with an eye to avoiding liability in discrimination lawsuits."[95] Though the study's focus primarily concerned corporate managerial positions and their hiring and promotional practices, their findings have broad relevance and salient parallels to the policing profession, in both bureaucratic structure and power dynamics. Dobbin and Kalev's findings support the utilization of certain tactics intended to enhance organizational racial and ethnic representation, such as framing the training message in terms of personal responsibility for increasing diversity. But they determined that the most effective trainings are optional. The authors note that providing employees with the option to partake in diversity training minimizes "backlash," a phenomenon that has been hypothesized as a cause for negative downstream training impacts, such as persisting or worsening ethnic and racial disparities in management positions.[96]

Though voluntary participation in diversity training appears to mitigate "backlash" effects, some research has concluded that the overall efficacy of these trainings falls short of their potential. In a meta-analysis examining the effects of intergroup contact on prejudice, Pettrigrew and Tropp found that "no-choice" programs, or mandatory programs, yielded more robust effect sizes than those in the "choice," or voluntary, samples.[97] Bezrukova et al. interpreted this finding to mean that voluntary programs are only reaching individuals already predisposed to the message, and

[94] Jena McGregor, "To Improve Diversity, Don't Make People Go to Diversity Training. Really," *The Washington Post*, July 1, 2016. See Alexandra Kalev, Frank Dobbin, and Erin Kelly, "Best Practices or Best Guesses? Assessing the Efficacy of Corporate Affirmative Action and Diversity Policies," *American Sociological Review* 71 (2006): 589-617; Frank Dobbin and Alexandra Kalev, "Why Diversity Programs Fail," *The Harvard Business Review* 94 (2016): 52-60.

[95] Ibid.

[96] Jessica Nordell, "Does Starbucks Understand the Science of Racial Bias?" *The Atlantic*, May 2, 2018.

[97] T F. Pettigrew and L R. Tropp, "A Meta-Analytic Test of Intergroup Contact Theory," *Journal of Personality and Social Psychology* 90 (2006): 751-783.

therefore, any behavioral or attitudinal change may be stunted by a diminished potential for variation.[98]

In the NYPD, FIP is mandatory for all personnel, and may therefore be vulnerable to some consequential backlash among trainees. The FIP training curriculum for patrol officers is designed to most effectually reach officers who are motivated to act without bias, and it is constructed to sidestep the possible causes of backlash - presumably by other officers who are required to attend – in stressing its foundations in scientific evidence and by curating a non-accusatory tone. Though its founder does not purport that the FIP training will have impacts on trainees' behavior, the potential for such unintended and adverse effects cannot be neglected.

Reducing Implicit Bias

To reduce implicit bias, FIP encourages participants to seek positive contact with outgroup members, engage in contact with counter-stereotypical outgroup members, and identify ways in which they are similar to outgroup members. There is reason to be optimistic that such tactics would have the intended impact. Pettigrew and Tropp's meta-analysis of 515 studies indicated that contact with outgroup members is inversely related to prejudice.[99] Longitudinal research in Belgium, England, and Germany verified that this association is causal in nature: secondary school students who had more (versus fewer) intergroup contacts reported less negative intergroup emotions and desire for social distance approximately six months later.[100] Such contacts increase familiarity with outgroup members, reduce the likelihood that they will be perceived as threatening, reduce intergroup anxiety, and facilitate physiological recovery following stressful intergroup interactions.[101]

---

[98] Bezrukova et al., "A Meta-Analytical Integration of Over 40 Years of Research on Diversity Training Evaluation," p. 43.

[99] Pettigrew and Tropp, "A Meta-Analytic Test of Intergroup Contact Theory." Also see Christopher L. Aberson, "Indirect Effects of Threat on the Contact-Prejudice Relationship: A Meta-Analysis," *Social Psychology,* advance online publication, 2019.

[100] Jens Binder, Hanna Zagefka, Rupert Brown, Friedrich Funke, Thomas Kessler, Amelie Mummendey, Annemie Maquil, Stephanie Demoulin, and Jacques-Phillippe Leyens, "Does Contact Reduce Prejudice or Does Prejudice Reduce Contact? A Longitudinal Test of the Contact Hypothesis among Majority and Minority Groups in Three European Countries," *Journal of Personality and Social Psychology* 96 (2009): 843-856.

[101] On increasing familiarity, see Pettigrew and Tropp, "A Meta-Analytic Test of Intergroup Contact Theory." On perceived threat, see Aberson, "Indirect Effects of Threat on the Contact-Prejudice Relationship." On intergroup anxiety, see Binder et al., "Does Contact Reduce Prejudice or Does Prejudice Reduce Contact?" On physiological recovery, see E Page-Gould, W B. Mendes, and B Major, "Intergroup Contact Facilitates Physiological Recovery Following Stressful Intergroup Interactions," *Journal of Experimental Social Psychology* 46 (2010): 854-858.

Research also suggests that countering stereotypes via a variety of methods can weaken bias.[102] As noted previously, Gawronski et al. found that negation training led participants to have automatic associations and evaluations that were more in line with negative stereotypes.[103] However, exposure to affirmation training, which instructed participants to respond in a supportive manner to counter-stereotypical associations (e.g., Black–smart) limited the extent to which stereotype activation and stereotype-consistent evaluations occurred in a subsequent task. FIP training includes elements of both: the management of negative biases and exposure to counter-stereotypes. Yet other research suggests that exposure to counter-stereotypes, regardless of one's endorsement of them, can confer similar benefits. For instance, participants who imagined counter-stereotypical individuals (e.g., a strong female) were less likely to exhibit implicit stereotypical associations on a subsequent task relative to participants who imagined stereotypical or neutral content or did not engage in imagery at all.[104] Of importance for understanding the relevance of exposure to counter-stereotypes for police officers, Park and Glaser found that undergraduate participants who were exposed to a greater number of counter-stereotypic targets (i.e., unarmed Blacks and armed Whites) on a shooter task exhibited reduced racial bias on a subsequent shooter task relative to those who were exposed to approximately equal numbers of counter-stereotypic and stereotypic targets.[105]

Evidence suggests that countering stereotypes in other ways also can be effective for reducing implicit bias and its effects. Stewart and Payne manipulated undergraduate participants' goals by instructing them to either associate Blacks with safe interactions or to make accurate or quick responses while completing a weapon identification task.[106] They found that participants who formed counter-stereotypical safety-related thoughts were less likely than others to mistakenly identify neutral objects as weapons following exposure to Black faces. In other work, Sim, Correll, and Sadler found that

---

[102] See, e.g., Calvin K. Lai, Maddalena Marini, Steven A. Lehr, Carlo Cerruti, Jiyun-Elizabeth L. Shin, Jennifer A. Joy-Gaba, ... and Brian A. Nosek, "Reducing Implicit Racial Preferences: I. A Comparative Investigation of 17 Interventions," *Journal of Experimental Psychology: General* 143 (2014): 1765-1785.

[103] Gawronski et al., "When 'Just Say No' is Not Enough."

[104] Irene V. Blair, Jennifer E. Ma, and Alison P. Lenton, "Imagining Stereotypes Away: The Moderation of Implicit Stereotypes through Mental Imagery," *Journal of Personality and Social Psychology* 81 (2001): 828-841.

[105] S H. Park and J Glaser, "Implicit Motivation to Control Prejudice and Exposure to Counterstereotypic Instances Reduce Spontaneous Discriminatory Behavior," *Korean Journal of Social and Personality Psychology* 25 (2011): 107-120. Also see Jessica J. Sim, Joshua Correll, and Melody S. Sadler, "Understanding Police and Expert Performance: When Training Attenuates (vs Exacerbates) Stereotypic Bias in the Decision to Shoot," *Personality and Social Psychology Bulletin* 39 (2013): 291-304.

[106] Brandon D. Stewart and B. Keith Payne, "Bringing Automatic Stereotyping Under Control: Implementation Intentions as Efficient Means of Thought Control," *Personality and Social Psychology Bulletin* 34 (2008): 1332-1345.

undergraduate participants exhibited less racial bias on the shooter task after reading a newspaper article that described a White as opposed to Black criminal, but police participants' performance was unaffected, and neither was that of undergraduates who practiced the task.[107] Taken together, these findings suggest that a single superficial exposure to a counter-stereotype or exposure that does not provide feedback about decision accuracy may not be sufficient to mitigate the effect of racial bias on officers' shooting decisions. This is supported by evidence that practice on a simulated shooter task in which race is merely nondiagnostic for correct decision-making (i.e., Black and White suspects are equally likely to be displayed with a gun or a neutral object) results in inhibited activation of racial concepts and fewer racially biased errors among undergraduates and police officer participants.[108]

With regard to the goal of enhancing feelings of empathy and perceptions of similarity, there is evidence this may reduce implicit bias, too. For example, perspective taking during brief interactions improved White Americans' attitudes toward Mexican immigrants and Israelis' attitudes toward Palestinians.[109] In further support, Forscher et al.'s recent meta-analysis of 492 studies tested the effectiveness of various procedures at reducing implicit bias specifically.[110] Results indicated that, compared to participants in neutral control conditions, those exposed to interventions that sought to weaken implicit bias directly (e.g., by showing participants pictures of admired Black people) or indirectly (e.g., by asking participants to adopt the perspective of a Black person) exhibited less bias on both implicit and explicit measures. These effects were small but significant.

There is also a small but compelling body of research that suggests that any bias interventions, or interventions specifically oriented to suppress individuals' biases, may actually evoke the opposite effect. Macrae, Bodenhausen, Milne, and Jetten found this to be true of interventions that encourage people to actively inhibit stereotypic thoughts.[111] The authors depict this type of effort as often counterproductive and, paradoxically, sometimes resulting in more ingrained impulses. One example of this effect is when someone on a diet tries not to think about fast food. The authors note that, especially in cognitively demanding contexts, "formerly unwanted thoughts

[107] Sim et al., "Understanding Police and Expert Performance."

[108] E. Ashby Plant, B. Michelle Peruche, and David A. Butz, "Eliminating Automatic Racial Bias: Making Race Non-Diagnostic for Responses to Criminal Suspects," *Journal of Experimental Social Psychology* 41 (2005): 141-156; E. Ashby Plant and B. Michelle Peruche, "The Consequences of Race for Police Officers' Responses to Criminal Suspects," *Psychological Science* 16 (2005): 180-183.

[109] Emile G. Bruneau and Rebecca Saxe, "The Power of Being Heard: The Benefits of 'Perspective-Giving' in the Context of Intergroup Conflict," *Journal of experimental social psychology* 48 (2012): 855-866.

[110] Forscher et al, "A Meta-Analysis of Procedures to Change Implicit Measures."

[111] C. Neil Macrae, Galen V. Bodenhausen, Alan B. Milne, and Jolanda Jetten, "Out of Mind But Back in Sight: Stereotypes on the Rebound," *Journal of Personality and Social Psychology* 67 (1994): 808-817.

become what is termed *hyperaccessible* for perceivers."[112] In three experiments, participants were shown an image of a stereotypical skinhead, and were asked to write a passage describing his average day. Half of participants were instructed to suppress their biases, and the other half were given no instruction with regards to prejudice. The authors tested the "rebound effect" of prejudicial attitudes in subsequent testing of the participants who were told to suppress their biases, finding in all three experiments that when removing the suppression constraints, those in the treatment groups displayed greater stereotypicality. The authors concluded that their findings provide evidence of attitudinal volatility, and that "there may be a range of ironic side effects associated with the seemingly functional and adaptive process of stereotype inhibition."[113] Though the FIP curriculum does not instruct trainees to *suppress* their biases, the line between suppression and management may be a thin one, especially in the minds of trainees (as opposed to academic discourse). In any case, this research is, we believe, properly included in a thorough account of the research on bias-related training.

A Synthesis

The distinct elements of FIP work together. One review of the literature on implicit bias training concluded that more comprehensive training and interactive workshops—like FIP—can be successful at raising participants' awareness of bias.[114] A meta-analysis revealed that diversity training yielded more benefits when it was accompanied by initiatives aimed at enhancing issue awareness and developing pertinent skills and strategies.[115] It is helpful to review the prejudice habit-breaking intervention established and evaluated by Devine, Forscher, Austin, and Cox.[116]

First, undergraduate participants completed explicit measures of racial attitudes, motivation to control prejudice, prejudice-relevant discrepancies, and concern about discrimination in society. Four weeks later, participants came to the researchers' laboratory to complete a baseline measure of implicit racial bias. Participants received feedback regarding the extent to which they exhibited implicit bias in favor of White or

[112] Ibid., p. 809.
[113] Ibid., p. 813.
[114] Doyin Atewologun, Tinu Cornish, and Fatima Tresh, *Unconscious bias training: An assessment of the evidence for effectiveness* (Equality and Human Rights Commission Research Reports Series, 2018). Retrieved from https://www. equalityhumanrights. com/en/publication-download/unconscious-biastraining-assessment-evidence-effectiveness.
[115] Bezrukova et al., "A Meta-Analytical Integration of Over 40 Years of Research on Diversity Training Evaluation."
[116] Patricia G. Devine, Patrick S. Forscher, Anthony J. Austin, and William T.L. Cox, "Long-Term Reduction in Implicit Race Bias: A Prejudice Habit-Breaking Intervention," *Journal of Experimental Social Psychology* 48 (2012): 1267-1278.

Black people and then were randomly assigned to either a control or intervention condition. Those in the control condition were dismissed but intervention participants then viewed a 45-minute narrated and interactive slideshow that educated them about implicit bias and trained them in five strategies for reducing implicit race bias: (1) increasing contact with outgroup members; (2) imagining counter-stereotypical others; (3) taking the perspective of an outgroup member; (4) recognizing when their responses are based on stereotypes and replacing the response with an unbiased response; and (5) observing specific information about individuals to prevent making inferences based on stereotypes. (These strategies largely overlap with those advocated for in FIP.) Intervention participants then completed measures assessing their reactions to the strategies. All participants subsequently returned to the lab to complete the implicit racial attitude test and completed explicit measures online at multiple times. Intervention participants additionally completed measures assessing their use of the strategies on which they had been trained. Final assessments were taken 8 weeks after the intervention. Results revealed that intervention participants had lower implicit racial bias than control participants at both the 4- and 8-week follow-up assessments. The only effects on explicit measures were that, compared to control participants, intervention participants exhibited increasingly more concern about discrimination over the course of the study and greater discrepancy between how they believed they should versus would think, feel, and act in intergroup interpersonal reactions. Of importance, the intervention did not change participants' standards for how they believed they should act, but it did lead participants to be more aware of their biases and the potential consequences for their behavior. Related to this point, intervention participants who perceived themselves as being more likely to use the strategies they had learned experienced greater reductions in implicit racial bias over the course of the study.

Devine et al.'s results suggested that enhancing awareness, motivation, and effort could generate long-term reductions in implicit bias. However, a subsequent replication study using a larger sample, shorter study period, and more follow-ups assessed more frequently yielded more nuanced results. Specifically, Forscher, Mitamura, Dix, Cox, and Devine found that both intervention and control participants exhibited less implicit racial bias after two weeks, although the effect was stronger in the intervention condition.[117] As in the original study, compared to control participants, intervention participants reported greater concern about discrimination and greater discrepancies between how they believed they should versus would think, feel, and act in intergroup interpersonal reactions. Yet, whereas concern about discrimination persisted over time, their should-

[117] Patrick S. Forscher, Chelsea Mitamura, Emily L. Dix, William T.L. Cox, and Patricia G. Devine, "Breaking the Prejudice Habit: Mechanisms, Timecourse, and Longevity," *Journal of Experimental Social Psychology* 72 (2017): 133-146.

would discrepancies did not. Forscher et al. additionally examined a number of behavioral outcomes. Strategy use did not generate change in implicit bias, nor did any particular strategies have consistent associations with concern about discrimination or should-would discrepancies. At follow-up assessments, participants also reported the number of race-related thoughts and conversations and interracial interactions they had had. Although intervention and control participants reported having these experiences at a similar rate, the quality of the experiences differed. Intervention participants were more likely to have (a) identified that a person was acting with bias, (b) labeled the action as bias, and (c) had interracial interactions with people they did not know well. Intervention participants used each of the strategies they had been trained in less than two times on average during the two-week observation period, and use declined over time. Of importance, observing others acting with bias and being cognizant of racial inequality in race-related conversations related to greater concern about discrimination, and having more interracial interactions was associated with lower expectations of exhibiting bias in those interactions, providing indirect evidence for intervention efficacy.

Further, two years later, Forscher et al. invited participants to partake in a separate, supposedly unrelated study on student engagement. Participants read a student newspaper essay that advocated for the use of racial stereotypes and then were given an opportunity to respond privately and publicly in an online post. They were also invited to donate to four charities, one of which aimed at reducing racial discrimination. To begin, there were no differences between intervention and control participants on implicit bias, concern about discrimination, or should-would discrepancies. Moreover, the two groups did not differ in the level of disagreement with the essay that they reported privately or the amount they donated to the social inclusion charity. However, intervention participants were significantly more likely to post a public comment objecting to the essay endorsing racial stereotypes. Overall, these results offer modest evidence of the prejudice habit-breaking intervention's effectiveness, and they temper confidence that FIP as a whole can have its intended effects on police officer participants.

Normative Influences

Largely anchored around Sherif and Sherif's Group Norm Theory (1953), numerous researchers have examined the question of group effects on individual prejudicial expression. In 2002, Crandall, Eshelman, and O'Brien conducted a number of experiments, which sought to qualify the correlation between an individual's prejudicial inclinations and society's acceptance or disapproval of those inclinations. They included in their analyses measures of prejudice that society tends to permit or encourage (such

as against rapists). By doing so, Crandall and colleagues were able to construct a measure of "acceptability of discrimination," thereby providing a standard of normative bounds against which individual attitudes can be compared. The authors found broad support for their supposition that "the public unwillingness to express prejudice may be more determined by normative influence than by personal attitudes."[118]

Duguid and Thomas-Hunt also conducted a series of experiments to explore the question of prejudicial malleability as it conforms to broader social normative standards.[119] The authors' hypothesis concerned the notion that, in training people to be less prejudiced by informing them that "most people are prejudiced," we are inadvertently normalizing prejudice. In four separate studies, participants were separated into two groups. One group was informed that *most people stereotype*, and the other that *most people do not stereotype*. In all four experiments, individuals in both groups were asked to respond to a number of surveys that queried their stereotypical perceptions of target groups or scenarios, as well as their hypothetical actions within those scenarios. In all four of these experiments, mean scores for stereotypicality were higher amongst those exposed to the message that *most people stereotype*, suggesting that this knowledge spurred some escalation in prejudicial affection. The authors note that this work demonstrates that "the impact of normative behavior expands beyond just perceptions and beliefs to include actions."[120] A critical conclusion of these works is that individual prejudice, to some degree, may in fact be a reflection of their shared normative values and "awareness of the general pervasiveness of stereotyping behavior does not mitigate stereotypic expression and, in fact, may have the opposite effect in increasing stereotyping."[121]

The literature reviewed thus far offers mixed support for the premise that FIP will be effective at enhancing officers' awareness of implicit bias and giving them skills to manage it. Atewologun et al. warn that there is very little evidence that implicit bias training impacts intergroup behavior.[122] Forscher et al. echoed this concern after finding that weakening implicit associations directly was the only intervention they studied that reduced biased behavior, and even this effect was "trivial."[123]

---

[118] Christian S. Crandall, Amy Eshleman, and Laurie O'Brien, "Social Norms and the Expression of Prejudice: The Struggle for Internalization," *Journal of Personality and Social Psychology* 82 (2002): 359-378, p. 374.
[119] Michelle M. Duguid and Melissa C. Thomas-Hunt, "Condoning Stereotyping? How Awareness of Stereotyping Prevalence Impacts Expression of Stereotypes," *Journal of Applied Psychology* 100 (2015): 343-359. Duguid and Thomas-Hunt did not analyze racial prejudice, and instead focused on the target groups of women, the elderly, and the obese.
[120] Ibid., p. 354.
[121] Ibid.
[122] Atewologun et al., *Unconscious bias training*.
[123] Forscher et al., "A Meta-Analysis of Procedures to Change Implicit Measures," p. 14.

Another important concern relates to the durability of training effects. Even when bias interventions have demonstrated effects on implicit measures, the effects have generally been short-lived.[124] For instance, the mitigating effect of perspective taking on Israelis' negative attitudes toward Palestinians did not persist for even one week following their dyadic interactions.[125] Moreover, Lai et al. found that despite having initial success in reducing implicit bias, none of the nine interventions they studied had effects that persisted over the course of one day.[126] Although Devine et al. showed that a prejudice habit-breaking intervention had lasting effects over an 8-week period, Forscher et al. found that even this comprehensive training strategy had limited effects over two weeks.[127] The durability issue is especially important considering that police officers' local environments could reinforce cultural stereotypes and implicit associations.[128] For example, Sim et al. found that special unit officers who dealt primarily with gangs and street crime exhibited more racial bias than patrol officers on a first-person shooter task.[129] Thus, without sustained efforts to maintain them, any gains made by FIP may be lost once officers go back to work.

It is important to keep in mind, however, that even trivial, short-term effects can be meaningful. Multiple meta-analyses indicate that implicit bias is correlated with discrimination-relevant behavior.[130] Even though the association is generally small,

---

[124] See, e.g., Devine et al., "Long-Term Reduction in Implicit Race Bias"; Forscher et al., "Breaking the Prejudice Habit"; Calvin K. Lai, Allison L. Skinner, Erin Cooley, Sohad Murrar, Markus Brauer, Thierry Devos, Jimmy Calanchini, Y. Jenny Xiao, Christina Pedram, Christopher K. Marshburn, Stefanie Simon, John C. Blanchar, Jennifer A. Joy-Gaba, John Conway, Liz Redford, Rick A. Klein, Gina Roussos, Fabian M. H. Schellhaas, Mason Burns, Xiaoqing Hu, Meghan C. McLean, Jordan R. Axt, Shaki Asgari, Kathleen Schmidt, Rachel Rubinstein, Maddalena Marini, Sandro Rubichi, Jiyun-Elizabeth L. Shin, and Brian A. Nosek, "Reducing Implicit Racial Preferences: II. Intervention Effectiveness Across Time," *Journal of Experimental Psychology: General* 145 (2016): 1001-1016. Also see Dale T. Miller, Jennifer E. Dannals, and Julian J. Zlatev, "Behavioral Processes in Long-Lag Intervention Studies," Perspectives on Psychological Science 12 (2017): 454-467.

[125] Bruneau and Saxe, "The Power of Being Heard."

[126] Lai et al., "Reducing Implicit Racial Preferences."

[127] Devine et al., "Long-Term Reduction in Implicit Race Bias"; Forscher et al., "Breaking the Prejudice Habit."

[128] See Nilanjana Dasgupta, "Implicit Attitudes and Beliefs Adapt to Situations: A Decade of Research on the Malleability of Implicit Prejudice, Stereotypes, and the Self-Concept," in Patricia Devine and A. Plant (eds.), *Advances in Experimental Social Psychology*. Vol. 47 (San Diego, CA: Academic Press), pp. 233-275.

[129] Sim et al., "Understanding Police and Expert Performance."

[130] Anthony G. Greenwald, T. Andrew Poehlman, Eric Luis Uhlmann, and Mahzarin R. Banaji, "Understanding and Using the Implicit Association Test: III. Meta-Analysis of Predictive Validity," *Journal of Personality and Social Psychology* 97 (2009): 17-41; Benedek Kurdi, Allison E. Seitchik, Jordan R. Axt, Timothy J. Carroll, Arpi Karapetyan, Neela Kaushik, Diana Tomezsko, Anthony G. Greenwald, and Mahzarin R. Banaji, "Relationship between the Implicit Association Test and Intergroup Behavior: A Meta-Analysis," *American Psychologist* 74 (2019): 569-586; Frederick L. Oswald, Gregory Mitchell, Hart Blanton, James

considering that nearly one million law enforcement officers in the United States[131] make a number of decisions on a daily basis that are potentially impacted by implicit bias, it can have societally significant consequences.[132] Moreover, it is encouraging that Forscher et al. found that a comprehensive intervention changed what might arguably be the most difficult kind of behavior—public objection to another's display of bias—as much as two years later.

One issue to monitor, however, is whether FIP has the ironic, unintended effects of increasing implicit bias or exacerbating its influence on officers' judgments and behavior. Atewologun, Cornish, and Tresh noted that training that leads individuals to believe that implicit bias is unchangeable can backfire.[133] As reviewed by Axt and Lai, people must not only have the skills needed to reduce discriminatory outcomes, but they must also have the motivation.[134] The backfire effect may occur if officers lack the motivation to address implicit bias because they believe it is out of their control. This is concerning because FIP intentionally adopts a non-accusatory framework that emphasizes the ubiquity of unconscious bias that is even found in people who explicitly endorse egalitarian norms.

FIP in a Nutshell

To recap: The goals of the Fair and Impartial Policing training are to inform police personnel about the scientific evidence of implicit bias and the various forms in which it may impact their day-to-day work. The primary points of patrol and first-line supervisor training include:

- Bias tends to be unconscious and ubiquitous
- Even well-intentioned people have biases
- Implicit biases can impact perceptions and behaviors, even outside of conscious awareness
- Policing based on biases and stereotypes can make policing unsafe, ineffective and unjust

---

Jaccard, and Philip E. Tetlock, "Predicting Ethnic and Racial Discrimination: A Meta-Analysis of IAT Criterion Studies," *Journal of Personality and Social Psychology* 105 (2013): 171-192.

[131] Federal Bureau of Investigation, *Crime in the United States, 2017* (Author, September, 2018). Retrieved from https://ucr.fbi.gov/crime-in-the-u.s/2017/crime-in-the-u.s.-2017/tables/table-74

[132] See Robert P. Abelson, "A Variance Explanation Paradox: When a Little is a Lot," *Psychological Bulletin* 97 (1985): 129-133; Anthony G. Greenwald, Mahzarin R. Banaji and Brian A. Nosek, "Statistically Small Effects of the Implicit Association Test Can Have Societally Large Effects," *Journal of Personality and Social Psychology* 108 (2015): 553-561.

[133] Atewologun et al., *Unconscious bias training.*

[134] Axt, and Lai, "Reducing Discrimination."

The FIP curriculum also imparts to front-line officers and supervisors strategies intended to help manage their biases. Briefly, these strategies are to:

1. Reduce biases by having positive contact with people who are different from you;
2. Manage biases by conducting self-checks on your behavior and practicing "controlled responses";
3. Avoid profiling by proxy by increasing your awareness of others' biases and ensuring that your actions are unaffected by them;
4. "Slow down" your responses and assess the situation more thoughtfully in order to reduce ambiguity;
5. Possess a strong grasp on the agency's biased policing policy; and
6. Assess the situation with all knowledge acquired through FIP training and with all of these tools at your disposal.

Some outcomes we might expect to see from FIP training would be an expanded understanding of implicit bias and the FIP training content, and changes in attitudes with regard to discrimination (for example, an increased concern for bias). In addition, effective training in strategies for managing implicit bias could be expected to lead to diminished disparities in enforcement – stops, frisks, searches, arrests, summonses, or the use of force – provided that such disparities were previously generated at least in part from implicit biases.

Chapter 3
The NYPD Context

New York City is arguably like no other city in the U.S., and the NYPD is arguably like no other police agency in the country. Extrapolating from the findings of evaluation research conducted in the NYPD to other police departments is fraught even under ordinary circumstances. And the current circumstances of the NYPD are not ordinary, as it is undergoing a series of reforms, a number of which are mandated by a federal district court and overseen by a court-appointed monitor. Here we briefly describe features of the City and the department that make them unique, and we describe at greater length the reform environment.

New York City

The City of New York is the most populous city in the United States with approximately 8.4 million residents. It is a very diverse community, with the largest foreign-born population of any city in the world, including more than 3.2 million residents born outside of the U.S.[135] The City spans a geographic area of 302 square miles and encompasses five boroughs, which vary in size and the composition of their populations (see Table 3-1).

Though New York City was widely regarded as a high-crime city in the early 1990s, the "crime drop" that began in the 90s in cities across the country was steeper and lasted longer in New York City than elsewhere.[136] Generally, crime has been decreasing for the past several years in the City of New York. Murder declined 87% from 1990 to 2018, including a 54% reduction from 2001 to 2018. Major crimes decreased 82% between 1990 and 2018, with some variation across New York's boroughs (from a 28.4% decrease in the Bronx to a 52.5% decrease in North Queens).[137]

[135] More foreign-born immigrants live in NYC than there are people in Chicago. Retrieved from https://www.huffpost.com/entry/new-york-city-immigrants_n_4475197.

[136] Franlin E. Zimring, *The City That Became Safe: New York's Lessons for Urban Crime and its Control* (New York: Oxford University Press, 2012).

[137] Major Crimes as defined by the NYPD differ slightly from UCR Part I crimes as defined by the Federal Bureau of Investigation. Major violent crimes include murder, rape, robbery, aggravated assault. Major property crimes include burglary, grand larceny and motor vehicle theft. Borough-specific crime information is from NYPD, *Borough and Precinct Crime Statistics* (2019). Retrieved from https://www1.nyc.gov/site/nypd/stats/crime-statistics/borough-and-precinct-crime-stats.page

Table 3-1. Population Characteristics of New York City and Its Boroughs.

	New York City	Brooklyn	Queens	Manhattan	The Bronx	Staten Island
Residential population	8.4 million	2.6 million	2.3 million	1.6 million	1.4 million	476,000
Median income	$57,782	$52,782	$62,008	$79,781	$36,593	$76,244
White	42.8%	49.5%	47.9%	64.5%	44.9%	75.2%
Black	24.3%	34.1%	20.7%	17.9%	43.6%	11.7%
Asian	14.0%	12.7%	26.8%	12.8%	4.5%	10.2%
Other	19.0%	3.7%	4.5%	4.8%	7.0%	2.8%
Hispanic	29.1%*	19.1%*	28.1%*	25.9%*	56.4%*	18.7%*

Source: 2017 Census estimates, US Census Bureau. Quick Facts. Retrieved from: https://www.census.gov/quickfacts/fact/table/newyorkcitynewyork,bronxcountybronxboroughnewyork,kingscountybrooklynboroughnewyork,newyorkcountymanhattanboroughnewyork,queenscountyqueensboroughnewyork,richmondcountystatenislandboroughnewyork/PST045218

*According to the US Census Race and Hispanic Origin definitions, because Hispanics may be of any race, percentages of race categories should not be combined with percentages of Hispanic (therefore, percentages will not add up to 100).

The NYPD

The City of New York Police Department (NYPD) was established in 1845 and is one of the oldest police departments in the United States. The NYPD is unlike any other police department, employing 36,000 sworn officers – more than twice as many officers as the second largest department in the U.S. – along with 18,000 civilians.[138] The NYPD provides a variety of services such as anti-terrorism, emergency services, and protection of the public transit systems, among other specialized services.

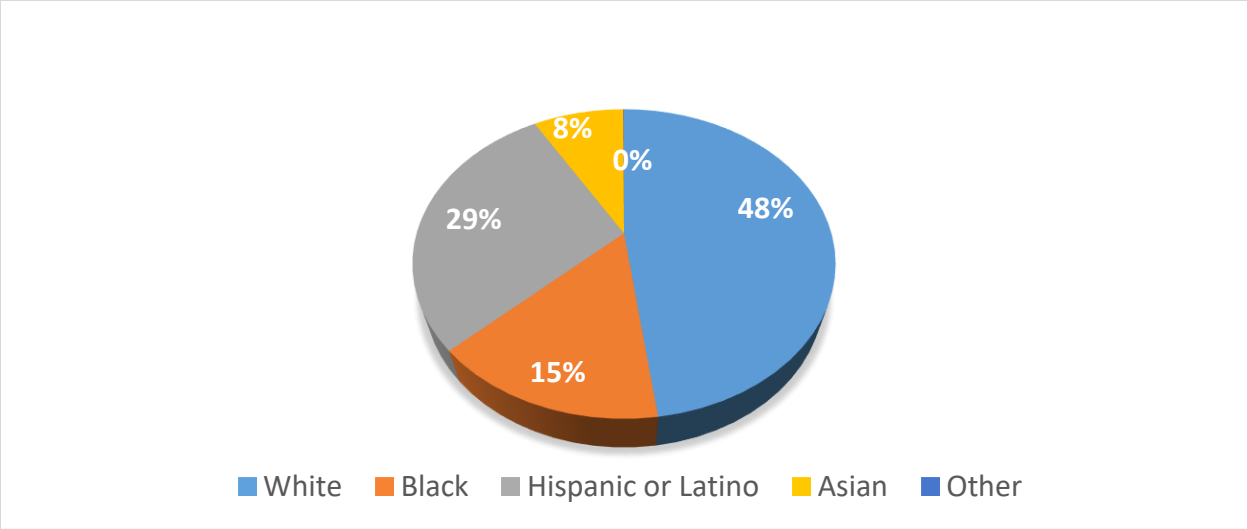
The NYPD is led by a police commissioner, a civilian administrator appointed by the city's mayor. A civilian first deputy commissioner and a number of other civilian deputy commissioners report directly to the commissioner, as does the chief of department – the highest-ranking sworn member of the NYPD – and several bureau chiefs, who are sworn police commanders.

The NYPD is itself racially and ethnically diverse. Overall, 52% of the NYPD's sworn personnel are members of racial or ethnic minorities, compared to the City's minority population share of 67.5%, a difference of -15.5%. Since the national average

[138] www.nyc.gov/site/nypd/about; Brian A. Reeves, *Local Police Departments, 2013: Personnel, Policies, and Practices* (Washington, DC: Bureau of Justice Statistics, 2015).

difference is -24.5%, the NYPD is more representative of the population it serves than the typical police agency.[139]

Figure 3-1. Race/Ethnicity of NYPD Sworn Personnel (2019)



Source: NYPD, *Data Transparency Initiative* (2019). Retrieved from: https://www1.nyc.gov/site/ccrb/policy/data-transparency-initiative-mos.page

Our evaluation focused on three bureaus of the NYPD: Patrol Services, Transit, and Housing. The Patrol Services Bureau is divided into eight borough commands, which are further divided into 77 police precincts. The number of residents in each precinct varies, but the typical precinct ranges from 70,000 to 150,000 residents. The Transit and Housing Bureaus police the subway and public housing developments in New York City, respectively. Prior to 1995, these two Bureaus were independent police departments, but later merged with the NYPD to enhance efficiency and coordination. The Transit Bureau is comprised of 12 transit districts and the Housing Bureau is comprised of 9 police service areas and serves approximately one half million citizens.

The NYPD formally prohibited racial profiling and bias-based policing in a policy issued in June of 2016.[140] The policy states that:

- Race, color, ethnicity, or national origin may not be used as a motivating factor for initiating police enforcement action. When an officer's decision to initiate enforcement action against a person is motivated even in part by a person's actual or perceived race, color, ethnicity or national origin, that enforcement action violates

[139] Police Department Race and Ethnicity Demographic Data: New York City Police Department. Available from https://www.governing.com/gov-data/safety-justice/police-department-officer-demographics-minorityrepresentation.html.

[140] NYPD Patrol Guide, Procedure No. 203-25, Department Policy Prohibiting Racial Profiling and Bias-Based Policing.

Department policy unless the officer's decision is based on a specific and reliable suspect description that includes not just race, age, and gender, but other identifying characteristics or information.

- Individuals may not be targeted for any enforcement action, including stops, because they are members of a racial or ethnic group that appears more frequently in local crime suspect data.
- The Administrative Code and Department policy prohibit the Department and individual officers from intentionally engaging in bias-based profiling, which is defined as "an act of a member of the force of the police department or other law enforcement officer that relies on actual or perceived race, national origin, color, creed, age, alienage or citizenship status, gender, sexual orientation, disability, or housing status as the determinative factor in initiating law enforcement action against an individual, rather than an individual's behavior or other information or circumstances that links a person or persons to suspected unlawful activity."

Other FIP Training

Prior to the delivery of FIP training to patrol officers, the NYPD began FIP training for command staff and for middle-managers (personnel above the rank of first-line supervisors but below command staff) in February 2018. Like FIP training for patrol, training for mid-management and command staff covers the science of bias and provides trainees with strategies individuals can use to manage their own biases. Their respective curricula, however, are more narrowly tailored to their specific roles. The 8-hour training for mid-managers focuses on their potential to identify and address bias exhibited by subordinates, instructs trainees on how to discuss bias with agency personnel and external groups (such as the media), and includes some topics more thoroughly covered in command-level training, such as hiring practices and accountability. Training for command staff outlines a more global approach to promoting fair and impartial policing within the organization, and covers topics such as agency policy, recruitment, data collection and measurement, and community outreach. At the time of the 10th Monitor's report in early January 2020, over 2,500 sworn personnel in middle-management and command-level positions had been trained, as well as the Chief of Staff, 1st Deputy, and Police Commissioner. We did not request or review training materials for the middle-management or command-level, and are therefore unable to provide a more detailed account of the curricula.

Contemporaneous Reforms

Implicit bias awareness training in the NYPD was delivered in the context of a broader set of reforms, many of them mandated and overseen by a federal court. Though implicit bias training was not a mandated reform, this context is important to consider and understand, for it might either amplify or attenuate the impacts of the training that we evaluate. The broader reform environment might reinforce the training in shaping officers' outlooks and performance. Alternatively, the multiple reforms might represent change on a scale to which officers are unable to readily adapt. Furthermore, external scrutiny associated with these reforms provides additional information about their implementation, of which we should take account. We describe the reforms that have the greatest relevance to understanding the effects of implicit bias training, and their implementation to date.

Court-Ordered Reforms

In 2013, the district court ruled in *Floyd v. City of New York* that the NYPD's practice of stop, question, and frisk was unconstitutional, violating the Fourth Amendment prohibition of unreasonable searches and seizures, and Fourteenth Amendment prohibition of discrimination based on race.[141] Other litigation against the NYPD, *Ligon v. City of New York*, and *Davis v. City of New York,* held that NYPD practices of criminal trespass enforcement in and around multiple-dwelling buildings enrolled in the Trespass Affidavit Program (TAP) and in New York City Housing Authority (NYCHA) buildings was unlawful.[142] The resolution of these cases eventuated in a set of remedial measures, the implementation of which was to be overseen by a court-appointed monitor. The remedial measures include as "immediate reforms" changes in NYPD's policies and procedures, supervision, performance evaluation, training, auditing, complaint processing, and discipline, as well as a pilot program to equip officers with body-worn cameras. The Independent Monitor began its work in late-2014, and it has issued a number of reports on the NYPD's implementation of the reforms.

In its ruling in the *Floyd* case, the court noted the potential influence of unconscious bias on officers' judgments and behavior:

> … recent psychological research has shown that unconscious racial bias continues to play an objectively measurable role in many people's decision processes. It would not be surprising if many police officers share the latent biases that pervade our society. If so, such biases could provide a further source of unreliability in officers' rapid, intuitive impressions of whether an individual's movements are furtive and

[141] *Floyd et al. v. City of New York, 08-CV-1034*,

[142] *Ligon et al. v. City of New York et al., 12-CV-2274*; *Davis et al. v. City of New York et al., 10-CV-00669.*

> indicate criminality. Unconscious bias could help explain the otherwise puzzling fact that NYPD officers check "Furtive Movements" in 48% of the stops of blacks and 45% of the stops of Hispanics, but only 40% of the stops of whites. There is no evidence that black people's movements are objectively more furtive than the movements of white people.[143]

Training in implicit bias was not ordered by the court, however. The NYPD elected to provide such training. Even so, the training is among the initiatives overseen by the Independent Monitor. In the first monitor's report, he explained that

> Training on implicit bias and procedural justice are two important areas of education. The Department has recognized that police officers will be much more effective and safer if they are aware of their own unconscious biases as well as those of others with whom they interact—e.g., community residents, witnesses and complainants, prosecutors, lawyers and judges. Incorporating training on implicit bias into the NYPD training curriculum will inform new recruits and officers about how stereotypes and unconscious attitudes (some developed during police work) can sometimes improperly influence their perceptions.[144]

Members of the Monitor's team reviewed the training materials and made recommendations for revision. Along with community leaders and plaintiffs' counsel, members of the Monitor's team also observed the training as it was delivered to NYPD senior executives.[145]

Policies & Procedures

The NYPD revised its patrol guide to include a definition of what a stop is and explain the circumstances under which stops can be made legally. The revised patrol guide clarifies definitions for: the reasonable suspicion that is the minimum basis for a stop; the separate suspicion – that a person is armed and dangerous – that allows police to frisk a person; and the legal requirements for conducting a search. New procedures also provide for a revised stop form to document stops, the bases for stops, and associated actions. The new stop form includes space for narrative explanations of the reasons for the stop, and for a frisk or search. Policies governing stops in or outside of TAP buildings and in NYCHA buildings have also been revised. Finally, revised procedures require supervisory review of the legality of stops and frisks (further discussed below).

[143] *Floyd et al. v. City of New York, 08-CV-1034*, pp. 44-45.
[144] Peter L. Zimroth, *First Report of the Independent Monitor* (New York: Arnold & Porter Kaye Scholer LLP, 2015), p. 38.
[145] Zimroth, *Ninth Report of the Independent Monitor*, p. 22.

Auditing

The Independent Monitor noted that "With respect to NYPD stop and frisk practices, the court found that the Department's monitoring and audit procedures were inadequate. Supervisors, ICOs [precinct-based Integrity Control Officers] and QAD [the Quality Assurance Division] reviewed the paperwork to assess how the stop report forms were filled out but did not conduct reviews to ensure that the stops were lawful."[146] New QAD auditing procedures include assessments of the lawfulness of stops and of frisks, as well as of supervisors' reviews of stop forms.

In addition, QAD now has procedures to conduct broader audits of stop documentation, running keyword searches of the new computer-aided dispatch system (ICAD) to identify incidents that may have been stops, reviewing the corresponding radio transmissions, and checking NYPD records to determine whether the stops were documented in a stop form. If no stop form was completed, QAD checks further to confirm that incidents were (or were not) stops.[147] QAD also audits police-initiated enforcement – arrests in which the People of the State of New York are the complainants – to determine whether stop reports should have but were not completed. In the monitor's ninth report, he pointed to the underreporting of stops as an unresolved issue: "The underreporting of stops has been acknowledged by the Department and by officers and supervisors in focus groups conducted by the monitor, and explicitly identified in audits."[148] In the monitor's tenth report, he reported that audits showed that in the first quarter of 2019, few stops were unreported.[149] If QAD and other audits were successful in improving the fidelity with which stops are documented, then we would expect to see increases in reported stops across the time period of our evaluation. If stops of Blacks and/or Hispanics were subject to greater underreporting, then we would expect to see increases that vary by the race and ethnicity of the people stopped.

Supervision

As noted by the Independent Monitor in 2015,

> One of the significant changes to NYPD procedures is that front-line supervisors will be responsible for reviewing the legality of stops and trespass arrests. … Front-line supervision was emphasized by the court because there is consensus among police agencies that these supervisors play the most important role affecting the culture of the organization. … The revised stop report form … requires supervisors to document whether the stop was supported by reasonable suspicion and, separately,

[146] Zimroth, *First Report of the Independent Monitor*, p. 46.
[147] Zimroth, *First Report of the Independent Monitor*, p. 49.
[148] Zimroth, *Ninth Report of the Independent Monitor*, p. 5.
[149] Zimroth, *Tenth Report of the Independent Monitor*, pp. 64-65.

> whether the frisk, if conducted, was supported by reasonable suspicion. If not, the supervisor must then document what action, if any, was taken: whether the report was corrected or the officer was instructed, referred for training or disciplined.[150]

Supervisors are required not only to review the stop report, but also to discuss the facts of the stop with the officer.[151]

By the time of the monitor's tenth report in December, 2019, however, such reviews by supervisors left room for improvement. In stop reports reviewed by QAD during the first three calendar quarters of 2018, about one-fifth did not adequately articulate reasonable suspicion, but less than 4 percent of those had been noted as deficient by supervisors.[152] According to the NYPD, the monitor explained, supervisors had in as many as 28 percent of the cases sent stop reports back to officers for corrections.[153]

To address the problems of underreporting by officers and under reviewing by supervisors, the NYPD initiated a Compstat-like process to focus the attention and efforts of senior officers on their subordinates' compliance with policies – policies regarding reporting, reviewing, and other matters (such as the activation of body-worn cameras) relating to risk.[154] Dubbed Remediation of Identified Situations Key to Success (RISKS), review meetings began in December, 2018 (a bit more than halfway through our evaluation period), following a "roll-out" in which commands were briefed on the performance indicators that would be tracked and the expectations for their involvement. Each command is subject to review at least twice each year; commands that exhibit compliance issues will be reviewed more frequently. By October of 2019 all commands had experienced their first RISKS review.[155]

Supervisors play a vital role in organizational efforts to ensure impartial policing as well as lawful policing. Insofar as supervisors are unable to fully adjust to the additional expectations that NYPD reforms place on them, and do so within the time frame of our evaluation, then the estimated impacts of the training would turn on only the training itself.

---

[150] Zimroth, *First Report of the Independent Monitor*, pp. 42, 43, 45.
[151] Zimroth, *Ninth Report of the Independent Monitor*, p. 4.
[152] Zimroth, *Tenth Report of the Independent Monitor*, p. 32.
[153] Ibid.
[154] Introduced in the NYPD in 1994 and widely emulated in other agencies, Compstat is a system of performance measurement and management accountability. In the NYPD, Compstat assigns managerial responsibility and authority to precinct commanders and holds them accountable for outcomes in their commands. A prominent feature is the Compstat meeting, at which commanders' performance is scrutinized. See, e.g., Eli B. Silverman, *NYPD Battles Crime: Innovative Strategies in Policing* (Boston: Northeastern University Press, 2001); Mark H. Moore and Anthony A. Braga, "Measuring and Improving Police Performance: The Lessons of Compstat and its Progeny," *Policing* 26 (2003): 439-453.
[155] Zimroth, *Tenth Report of the Independent Monitor*, pp. 12-14. We learned of the RISKS reviews too late to incorporate them into our evaluation.

Performance Evaluation

Lest the criteria and procedures for evaluating officers' performance generate administrative pressure to conduct enforcement actions that are not otherwise well-founded and prudent, the NYPD revised its performance evaluation process. Under the new system, an "officer profile report" is generated automatically each month, compiling data from various department data sources (not to include a count of stops). On the supervisor feedback form, supervisors can make note of an officer's accomplishments as well as areas of performance that need improvement. On the officer self-report form, officers may document noteworthy achievements and actions. These three sources form a partial basis for a quarterly evaluation, for which officers are assessed on each of twelve dimensions, such as problem identification/solving, judgment, community interaction, and initiative. The monitor surmised that the new evaluation system does not prompt supervisors to encourage officers to conduct stops without regard to their lawfulness, but focus groups with supervisors and a review of evaluations in a sample precinct identified some challenges. Supervisors were confused about the types of conduct encompassed by the various evaluation dimensions and about how composite scores were formed. It was also evident that supervisors were reluctant to use the "needs improvement" rating.[156] We would add that the new evaluation system adds further to the adjustments required of supervisors in the reform process.

Training

In-service training courses on stop and frisk and racial profiling, for officers and supervisors, respectively, were overhauled. These courses cover fundamental principles and the law of investigative encounters and interior patrols (i.e., in housing developments), as well as applicable NYPD procedures. Coverage extends to the proper documentation of stops and trespass arrests, including the circumstances under which the completion of the stop form is required, and how to complete the form's narrative section. Supervisors' training further encompasses their role in ensuring proper documentation, which includes discussing stop reports with their officers. The revised courses were rolled out beginning in February of 2018.[157] Revised training for *newly promoted* sergeants and lieutenants began in spring of 2018, likewise providing for greater attention to their responsibilities for supervising investigative encounters.

Though the in-service FIP training was delivered separately, the content of the training on stop and frisk and racial profiling partially overlaps with that of FIP. The training includes, at several junctures, opportunities for the discussion of the role of race in investigative encounters. "The materials describe the difference between the

[156] Zimroth, *Tenth Report of the Independent Monitor*, pp. 68-71.

[157] See Peter L. Zimroth, *Seventh Report of the Independent Monitor* (New York: Arnold & Porter Kaye Scholer LLP, 2017), pp. 15-19; and Zimroth, *Tenth Report of the Independent Monitor*, pp. 38-39.

constitutionally permissible use of race based on a specific, reliable suspect description and the constitutionally impermissible targeting of racially defined groups for stops."[158] The stop and frisk training for patrol officers began prior to the FIP training for patrol officers and is projected to finish in the summer of 2020, well after the FIP training for the patrol services, transit, and housing bureaus was completed in April of 2019.[159]

Body-Worn Cameras

The NYPD undertook a court-mandated pilot program of the use of body-worn cameras (BWC) in 2017. "The goal of the pilot program is to assess the costs and benefits of deploying cameras and whether deployment results in reducing unconstitutional stops and frisks."[160] The execution of the pilot program was organized around an evaluation design that provided for a randomized controlled trial. Forty matched pairs of precincts were identified based on complaint counts and other criteria, and one of the precincts in each pair was randomly selected as a treatment precinct in which body-worn cameras would be deployed for one year. In each of the treatment precincts, the design specified that body-worn cameras would be worn by 45 officers in each precinct's 3rd platoon and 15 other officers ("plainclothes officers from the specialized anti-crime units and traffic enforcement officers").[161] In all, 1,200 officers would participate in the treatment group, and 1,200 officers from the other twenty precincts would form the control group. In its 8th report, the monitor reported that "The implementation of the BWC cluster randomized experiment began on April 24, 2017 in the 34 Precinct. The remaining treatment precincts received the BWC technology in a staggered manner with the final treatment precinct, Precinct 121, receiving BWCs on November 14, 2017."[162] Each pair of precincts was to remain in the designated experimental condition – BWCs on 3rd platoon officers in treatment precincts, no BWCs in control precincts – for one full year from the initial date of deployment. The NYPD was to deploy BWCs to all officers (including 1st and 2nd platoons) in precincts other than the 40 pilot precincts beginning in December, 2017 and concluding in December, 2018.[163] Deployment to all uniformed officers was completed in February 2019.[164] Thus the deployment of BWC to pilot and other precincts overlapped with the FIP training.

[158] Zimroth, *Tenth Report of the Independent Monitor*, p. 37.
[159] Zimroth, *Tenth Report of the Independent Monitor*, p. 39.
[160] Peter L. Zimroth, *Sixth Report of the Independent Monitor* (New York: Arnold & Porter Kaye Scholer LLP, 2017), p. 1.
[161] Zimroth, *Sixth Report of the Independent Monitor*, p. 5.
[162] Peter L. Zimroth, *Eighth Report of the Independent Monitor* (New York: Arnold & Porter Kaye Scholer LLP, 2018), p. 3.
[163] Zimroth, *Eighth Report of the Independent Monitor*, p. 8
[164] Zimroth, *Tenth Report of the Independent Monitor*, p. 52.

Cameras were deployed to officers in Housing PSAs between February and December 2018, a period that also overlapped with FIP training.[165]

We note also that the deployment of body-worn cameras came with new requirements for supervisors:

> Sergeants must review five BWC videos each month. After reviewing these videos, sergeants must complete a BWC self-inspection worksheet for each video; the sergeant's platoon commanders or lieutenant must then review two of the videos and complete the self-inspection worksheet; and the command's executive officer must review and approve the BWC self-inspection worksheet.[166]

Other Reforms

Four additional reforms of which our evaluation must be cognizant were likely to affect enforcement practices in the NYPD. First, New York State's Raise the Age (RTA) law, passed in April of 2017, raised the age of criminal responsibility from 16 to 18 years of age. Prior to its passage, New York State was only one of two states in the country that treated all 16 and 17 year-olds as adults in the criminal justice system without consideration of the crime committed. RTA legislation provided for a two-year phase-in, with the age of criminal responsibility becoming 17 on October 1, 2018, and 18 on October 1, 2019. The initial phase became effective several months after FIP training for patrol officers commenced.

Assessment of arrest trends in across New York State indicated successful implementation of both Phase 1 and Phase 2 of the law. Felony arrests of 16 year-olds decreased 36 percent, from 244 per month in 2017 to 155 after the law took effect in October 2018. County-level analysis revealed variation across counties, with 10 of 62 counties reporting no felony arrests of 16 year-olds in the 6 months immediately following RTA implementation.[167]

Second, as announced in June 2018 and effective September 2018, the NYPD relaxed marijuana enforcement, issuing summonses rather than making arrests for marijuana violations. Specifically, on June 19, 2018, New York City Mayor Bill de Blasio and NYPD Police Commissioner James O'Neill announced that by September 1, 2018:

> The majority of New Yorkers found smoking marijuana in public will face criminal summonses instead of arrest – continuing a significant shift in overall marijuana enforcement to better balance fairness with public safety and quality of life concerns.

---

[165] Zimroth, *Tenth Report of the Independent Monitor*, pp. 51-52.

[166] Zimroth, *Tenth Report of the Independent Monitor*, p. 50.

[167] New York State Raise the Age Implementation Task Force, *Raising the Age of Criminal Responsibility*. First Annual Report, August 2019. https://www.ny.gov/sites/ny.gov/files/atoms/files/NYS_RTA_Task_Force_First_Report.pdf

> The change is the result of the final report and recommendations produced by the 30-Day Working Group on Marijuana Enforcement that convened in May, and according to NYPD projections, will likely reduce overall marijuana arrests in New York City by about 10,000 per year based on 2017 arrest records and patterns.[168]

This policy change came on the heels of a May 2018, *The New York Times* article that examined racial/ethnic disparities in arrests for marijuana drug offenses made by NYPD Officers. Mueller reported that after considering the number of citizen calls for service regarding marijuana violations, higher arrest rates remained in precincts with larger non-White residential populations.[169] The resulting call for more scrutiny regarding the NYPD's policy and practice of enforcement of marijuana laws was echoed by others.[170] The 30-day Working Group on Marijuana also reported that about 40 percent of people arrested by the NYPD for smoking marijuana had no prior arrest history. A follow-up article reiterated the NYPD's analyses that demonstrated a consistent reduction in the number of marijuana arrests over the past decade from 53,000 marijuana arrests in 2010 to 19,000 in 2017, however added that 9 out of 10 marijuana arrests by the NYPD typically involved a non-White (i.e., Black and/or Hispanic) suspect.[171] Likewise, *The New York Times* article found that the first three months of 2018 resulted in approximately 4,000 people arrested for marijuana possession, and 89 percent of those arrested were Black or Hispanic.[172]

In a press release, Mayor de Blasio and Commissioner O'Neill indicated that the purpose of this policy change was to "strike a balance on marijuana enforcement between fairness and safety" and to "help reduce unnecessary arrests, while making our City fairer and safer."[173] Underlying this stated purpose was the expectation that this policy change would significantly reduce the racial disparities in low-level drug offenses.

The Working Group also found that in 2017, nearly 52,000 complaints regarding public smoking of marijuana were received; while addressing these complaints is important, the NYPD noted that officers have discretion in how to exercise their enforcement powers. Over the course of the summer, NYPD officers received instructions regarding this policy change, with full implementation by September 1,

---

[168] New York City Police Department, "Mayor De Blasio, Commissioner O'Neill Unveil New Policy to Reduce Unnecessary Marijuana Arrests," June 19, 2018 press release. Retrieved from https://www1.nyc.gov/site/nypd/news/pr0619/mayor-de-blasio-commissioner-o-neill-new-policy-reduce-unnecessary-marijuana-arrests

[169] Benjamin Mueller, "Mayor and Some Prosecutors Move to Curb Marijuana Arrests," *New York Times*, May 15, 2018. Retrieved from www.nytimes.com.

[170] Clodagh McGowan, "The NYPD is Officially Done Arresting Most Marijuana Smokers in the City," *Spectrum News New York One*, September 1, 2018. Retrieved from www.ny1.com.

[171] Ibid.

[172] Mueller, "Mayor and Some Prosecutors Move to Curb Marijuana Arrests."

[173] New York City Police Department, "Mayor De Blasio, Commissioner O'Neill Unveil New Policy to Reduce Unnecessary Marijuana Arrests."

2018. This policy change allowed for marijuana arrests if suspects were on probation or parole, had criminal warrants, could not provide identification, had a recent documented history of violence, or there was an immediate public safety risk as associated with smoking (e.g., operating a motor vehicle).

This change in the NYPD marijuana enforcement policy coincided with the FIP training. Our analyses of officers' enforcement behavior, especially the frequency and racial disparities in arrests and uses of force, must account for changes in policy and practice that are unrelated to the implicit bias training. As Mayor de Blasio suggested, "we believe that this will result in thousands of fewer arrests … In fact next year, we think at least 10,000 fewer New Yorkers will be arrested under this new policy." Complicating our study design for examining the impact of implicit bias training, it was also anticipated that reductions in arrest corresponding to the changes in marijuana enforcement would significantly impact non-White suspects more than Whites.

Third, Neighborhood Policing was initiated in some precincts in 2015 and had been implemented in every precinct by October of 2018.[174] Neighborhood Policing is NYPD's 21st century version of community policing. Each precinct is divided into three to five sectors, whose boundaries are congruent (as nearly as feasible) with those of recognized neighborhoods. Assignments of officers to sectors and shifts are intendedly stable over time, so that officers and residents can become familiar with one another. Dispatch practices are designed to maintain "sector integrity," such that requests for assistance originating in the sector are handled by officers assigned to that sector. Staffing is sufficient to allow officers unassigned time during which they can engage with the community and practice problem-solving. Two officers in each sector are neighborhood coordination officers (NCOs), who attend community meetings, visit schools, follow up on incidents, and work with precinct detectives to develop leads in criminal investigations. NCOs also host "Build the Block" meetings, which are devoted to collaborative problem-solving. The roll-out of NCOs was completed in all but 13 precincts prior to the start of FIP training in May of 2018, and each of those 13 precincts had implemented Neighborhood Policing prior to their FIP training. However, between May and October, 2018, those 13 precincts were included in control training blocks as Neighborhood Policing was implemented.

Fourth, The Right to Know Act, which became effective in October of 2018, consists of two laws, the consent to search law and the NYPD ID law. The former requires that, lacking other legal bases to search a person, his/her vehicle, or his/her home, an officer may search with a person's consent provided that the officer not only asks for consent, but also informs the person that a search will not be conducted if

[174] See https://www1.nyc.gov/site/nypd/bureaus/patrol/neighborhood-coordination-officers.page, and https://www1.nyc.gov/site/nypd/news/pr1022/neighborhood-policing-now-every-neighborhood-new-york-city#/0, accessed June 15, 2020.

consent is withheld and ensures that the person understands. If the person's English language proficiency is limited, then the officer must use appropriate interpretation services. The officer must document the requests for consent to search. If the interaction is recorded by the officer's body-worn camera, the officer must provide the person with information on how a copy of the recording can be obtained. If the officer asks for consent to search, then the officer must provide a business card regardless of whether consent is granted, unless a summons is issued or an arrest made.

The NYPD ID law applies to instances in which a person is stopped (including roadblocks and checkpoints), frisked, searched, or otherwise suspected of criminal activity, and those in which a person is questioned as a survivor or witness. Under these circumstances, officers are required to identify themselves by name, rank, command, and shield number, and provide a business card that contains the same information. Under other circumstances, officers must provide a business card upon request. [175]

Summary

The NYPD is a distinctive police organization, from which generalization to American policing is subject to more than the ordinary caveats. It is sufficiently large to permit a randomized controlled trial to estimate the effects of training on behavior, because training thousands of patrol officers cannot be accomplished in a few weeks' time, such that the post-training behavior of trained officers can be compared to that of officers who have yet to be trained. If agency size represents a contingency in the effects of training, however, then extrapolating from findings in the NYPD must be done with great care. Furthermore, the city that the NYPD serves is not only large but diverse, with a multiplicity of races, ethnicities, religions, and nationalities.

Moreover, implicit bias training in the NYPD was delivered in a larger context of numerous reforms, many of which have been overseen by a court-appointed monitor. Some changes in enforcement patterns preceded the implementation of reforms, as we discuss below, and those changes alter the baseline against which post-training patterns are compared. The breadth and depth of these reforms could amplify or attenuate the effects of the training. It is also possible that the effects of one or more reforms – such as the revised policy on marijuana enforcement – on officers' enforcement activity could be confounded with the effects of the training. Of these facts our evaluation must be cognizant.

[175] https://www1.nyc.gov/site/ccrb/complaints/right-to-know-act.page, and https://www.changethenypd.org/about-right-know-act, accessed June 15, 2020.

Chapter 4
Evaluation Design

The evaluation concentrates on outcomes that could be expected to follow from the FIP training. FIP trainees are told that, at the conclusion of the training, they will be able to:

- recognize their own human biases;
- articulate how implicit biases can affect their perceptions and behavior;
- discuss how biased policing impacts community members and the law enforcement organization;
- describe how FIP supports procedural justice and police legitimacy; and
- demonstrate skills and tactics to reduce the influence of bias on police practice and allow them to be safe, effective and just police professionals.

If the training is effective, its most immediate impact would be on officers' beliefs and attitudes, as officers become better informed of facts about implicit bias, based on the science of implicit bias, and the potential implications for policing. In addition, insofar as the training points to the detrimental impacts of biased policing on officers' work environment, it could be expected to affect their attitudes about the salience of bias and discrimination as a social problem and the importance of policing without prejudice.[176] Our evaluation therefore, includes an analysis of officers' beliefs about implicit bias and their attitudes toward discrimination. Since the training instructs officers about strategies to manage their unconscious biases, it should also elevate officers' cognizance of the potential for managing unconscious bias. To the extent that officers learn and later apply the strategies that the training presents, we could expect corresponding reductions in enforcement disparities – in stops, frisks, searches, arrests, summonses, or the use of force – insofar as disparities in these outcomes stem at least in part from implicit biases. Thus, the evaluation also extends to an analysis of enforcement actions on the street.

The evaluation design includes several components. First, the "training-day survey," administered on the day of FIP training, forms the basis for estimates of the most immediate effects of the FIP training on officers' beliefs and attitudes. Trainees were surveyed either prior to or following the training, allowing for a comparison of pre- and post-training responses. Second, since the impacts of the training are likely to be greater when it is supported by other organizational forces, a survey of immediate supervisors was administered between March and May, 2019, to assess the degree to which supervisory practices reinforce the FIP training. Third, a follow-up survey of trained officers was administered between July and August of 2019, allowing an

[176] The FIP training is concerned with implicit biases of many types. However, our analyses of beliefs, attitudes, and behavior focus on racial and ethnic biases and disparities.

assessment of officers' actual (self-reported) use of the FIP bias-management skills, and its associations with the beliefs and attitudes on which FIP might have effects, as well as an assessment of decay in the immediate effects of the training on officers' knowledge about implicit bias. Fourth, brief, open-ended interviews with a small sample of 70 officers and supervisors were conducted and analyzed qualitatively to provide context for the interpretation of quantitative findings based on the surveys.

Finally, the evaluation includes an analysis of officers' enforcement behavior, as it is captured in reports of arrests, summonses, stops, frisks, searches, and uses of force, to estimate the impacts of the training on behavior. For this purpose, officers were randomly assigned to treatment and control groups. Commands were organized into five pairs of experimental training blocks; within each pair, one was randomly selected as a "treatment block." The five treatment blocks were scheduled for training first, followed by the control blocks, such that the post-training enforcement behavior of treatment blocks can be compared with the pre-training enforcement behavior of the control blocks. The details of this design are elaborated below.

Training-Day Survey

The training-day survey took two forms: pre-training and post-training versions. The content of the two survey instruments was for the most part identical. The heart of both versions of the survey consisted of 27 items about "beliefs and attitudes relating to stereotypes and prejudice," as it was explained to respondents on the survey instrument.

Twenty of the items were statements rooted in the training content, with which respondents could agree or disagree. Depending on whether the statement represented or misrepresented the training, a respondent's agreement or disagreement with a statement can be construed as consistent or inconsistent with the content (i.e., correct or incorrect). Several of these items were adopted from survey instruments used in other studies: the Urban Institute's evaluation of implicit bias training in the six National Initiative cities, from which four items were adopted, and an evaluation of FIP training in a Canadian city, from which four items were adopted. Based on our reading of the FIP materials, we created twelve additional items, which were reviewed by FIP executives. Of the twenty items, eighteen concerned implicit bias and two concerned procedural justice.

The other seven items measure outlooks that are partly cognitive and partly affective and value-laden – i.e., attitudes. Two such items were drawn from social psychological inquiry on people's motivation to respond without prejudice in social situations, and two other items were drawn from studies of people's concern about discrimination; the research in which these items are rooted was discussed above. Jennifer Eberhardt, a social psychologist and one of the experts on the Independent

Monitor's team, contributed three items about officers' concerns relating to police-community relations.

The post-training survey included, in addition, a set of six items that prompted officers to assess what they learned from the training (two of which were adopted from the Urban Institute's survey instrument), and a set of items on the likelihood that they would use the strategies for managing unconscious bias to which the training introduced them.

Trainees were asked to complete one but not both versions of the survey, and survey administration was randomized: on even-numbered days, officers were invited – and given time – to complete the pre-training survey at the beginning of the training session; on odd-numbered days, officers were invited to complete the post-training survey at the end of the day. The training for patrol officers commenced on May 14, 2018, as plans for the survey were being finalized. The survey was first administered on July 9, 2018, such that seven weeks of training attended by 2,333 trainees preceded the initiation of surveying. The survey was administered each day of training through the conclusion of training for patrol officers on April 16, 2019, excepting the classes between January 7 and March 13, during which 3,249 trainees attended the training. A misunderstanding by the FIP trainers about the duration of the surveying led to the survey's discontinuation at that time, until it was discovered by the research team on March 13. Surveying was resumed on March 14 and continued through the completion of FIP training for patrol officers in the three bureaus.[177]

In all, more than 15,000 trainees attended the training. Excepting the 2,333 who attended prior to the initiation of the survey, and the 3,249 who attended during the suspension of surveying, 9,981 trainees attended on days during which surveys were administered. Of those, 7,540 trainees participated by completing at least the first section of the survey (a 75.5 percent response rate). Participation in the post-training survey was higher (85.6 percent) than that in the pre-training survey (65.7 percent). We suspect that the difference stems from late-arriving trainees, as even tardiness of only five minutes would likely affect their ability to participate.

The plan for randomized survey administration – with pre-training surveys completed on even-numbered dates and post-training surveys on odd-numbered dates – was followed with few exceptions. Of the 282 occasions on which the survey was

[177] Given the number of respondents who had already completed the survey, our analysis of immediate training impacts on trainees' beliefs about implicit bias and attitudes toward discrimination should be unimpaired. The sample size provides for ample statistical power in detecting training impacts Compared with the officers who were surveyed (before and after the survey hiatus), officers who were not surveyed were not very different: disproportionately White (48.2% versus 44.2%) rather than Black (13.4% versus 15.3%), Hispanic (29.8% versus 31.0%), or Asian (8.6% versus 9.5%); assigned to the patrol bureau (82.3% versus 79.2%) or transit (10.7% versus 9.8%) rather than housing (7.0% versus 11.0%); and slightly younger (33.48 versus 33.81).

administered (i.e., two training sessions on each of 141 days), only 15 departures from the planned odd/even scheduling occurred, and the departures themselves appear to be random.[178] Successful randomization would leave only one reason to expect that the officers who complete the pre-training survey differ in relevant respects from the officers who complete the post-training survey: a systematic difference in rates of survey participation.

The success of the random administration can be evaluated in terms of identifiable differences between pre-training and post-training survey respondents (i.e., their length of service, rank, sex, age, race, education, and military experience). We found only small differences between the pre- and post-training respondents, and although a few of the differences (in respondents' sex and race) achieved statistical significance given the large sample, the differences were substantively minor.[179] The details of these comparisons are shown in Appendix A. We can therefore draw inferences about immediate impacts of the training on officers' beliefs and attitudes from straightforward comparisons of post- and pre-training responses. We need not rely entirely on the assumption that randomization suffices to control for all third factors, however, as our analysis includes statistical controls that supplement the experimental control.

Follow-up Survey

We administered a follow-up survey to all FIP trainees in the three bureaus, about two months following the completion of training by officers in the last training classes. The follow-up survey was intended to serve two purposes: (1) to assess officers' actual (and not anticipated) use of the bias-management strategies on which FIP instructs them; and (2) to assess decay in training impacts on officers' beliefs about implicit bias. The follow-up survey instrument included the same set of belief and attitude items included in the training-day survey, and in addition, items on self-reported use of the various FIP strategies. Trainees in the various experimental training blocks were

[178] Survey administration within weeks generally alternated between pre- and post-training administration. The departures include: July 27, when the pre-training survey was mistakenly administered to the morning training session, and March 18, when the post-training survey was mistakenly administered to the morning training session; August 30 and March 18, when the post-training survey was mistakenly administered to the afternoon training session; the week of July 30, when survey administration alternated from one day to the next even though the post-training survey should have been administered back-to-back, July 31 and August 1; the holiday weeks of September 3 and January 1, when survey administration alternated from one day to the next; November 1, when the post-training survey should have been administered on the second consecutive day, and November 2, when survey administration alternated.

[179] Note, however, that post-training respondents were more likely to decline to report their length of service and their rank.

surveyed separately, so that we could distinguish among respondents in terms of when they received the training. The NYPD's Strategic Initiatives Bureau sent initial invitations June 21 and 24-26; several reminders were sent in July and August.

Among the 15,693 trainees still employed when the follow-up survey was administered, 3,121 opened the survey but of those, 1,568 answered questions, a response rate of 10 percent. This response rate was certainly lower than that to which we aspired, but it was not low relative to what is achieved when personnel in large police agencies are asked to participate in a web-based survey. Consider the survey of police officers conducted by the Pew Research Center in 2016. Beginning with the agency sample for the National Police Research Platform, officers in 54 agencies were surveyed online. Overall, the response rate was 13.9 percent, and in the larger agencies (i.e., departments with more than 1,600 sworn), the response rate was *under* 10 percent.[180] Moreover, response rates like those obtained in the follow-up survey do not imply that the findings based on the survey are invalid or uninterpretable. That "low" response rates – i.e., rates under 50 or 60 percent – are unacceptable, yielding data that are unrepresentative, has been characterized as a "methodological myth" based on numerous analyses.[181] The response rate on the follow-up survey appropriately qualifies generalizations that we can draw and comparisons that we can make with the training-day survey, but does not invalidate the findings.

A priori, we might expect that the completion of the follow-up survey would be more likely among those who are either (1) especially concerned about discrimination and motivated to respond without prejudice, or (2) especially skeptical that bias in policing is a serious issue. The data offer some support for both expectations. First, follow-up survey respondents were older and disproportionately non-White compared with the population and the training-day survey respondents (although about one-third of the follow-up survey respondents declined to complete the background items); in the training-day survey, these demographic groups were more likely to express concern about discrimination and a motivation to respond without prejudice (and to answer implicit bias survey items correctly). Second, follow-up survey respondents' attitudes tended to be more crystalized and intense, with proportionately fewer "neither agree nor disagree" responses and more responses of strong agreement or disagreement. Since we would expect that attitudes are less malleable, and the training-day survey

[180] Pew Research Center, *Behind the Badge* (Author, 2017), p. 92. https://www.pewsocialtrends.org/2017/01/11/behind-the-badge/

[181] See Justin T. Pickett, "Methodological Myths and the Role of Appeals in Criminal Justice Journals: The Case of Response Rates," *ACJS Today* 42 (2017): 61-69. Also see Justin Nix, Justin T. Pickett, Hyunin Baek, and Geoffrey P. Alpert, "Police Research, Officer Surveys, and Response Rates," *Policing & Society* 29 (2019): 530-550; and Justin T. Pickett, Francis T. Cullen, Shawn D. Bushway, Ted Chiricos, and Geoffrey Alpert, "The Response Rate Test: Nonresponse Bias and the Future of Survey Research in Criminology and Criminal Justice," *The Criminologist* 43 (2018): 7-11.

data indicate that the training had fairly small effects on officers' attitudes, the differences between training-day and follow-up survey responses on attitudinal items is likely attributable to different patterns of non-response rather than change over time. Overall, the follow-up survey respondents appear to overrepresent those who were more skeptical about the problem that racial/ethnic discrimination constitutes.

Supervisor Survey

The FIP training for supervisors is designed both to better inform trainees about the science of implicit bias and to prepare them to play a critical role in promoting unbiased policing. Supervisors may respond to potential bias that they detect; they can also take a proactive approach to preventing biased policing on the street. In these and other ways, supervisors can reinforce the FIP training and, presumably, deepen and sustain the effects of the training on officers' will and skill in managing their unconscious biases. A one-day training, however well-conceived and delivered, cannot be expected to carry the entire organizational burden of minimizing the effects of unconscious biases. Other organizational supports for fair and impartial policing are necessary, and given the key role that immediate supervisors play in police patrol work generally, the NYPD's sergeants might be expected to play an important part in this process.

To better understand the role of supervision in enhancing fair and impartial policing, we surveyed sergeants. The survey instrument encompassed five topics: (1) police supervision generally; (2) implicit bias and discrimination (27 items that also appeared on the training-day survey); (3) hypothetical scenarios of subordinate officers whose performance might call for supervisory intervention; (4) the FIP training and supervisors' applications of the strategies for managing and supervising implicit bias; and (5) respondents' backgrounds.

The NYPD's Strategic Initiatives Bureau sent a link to the web-based survey to all of the department's sergeants on March 20, 2019; reminders were sent on April 2 and April 9, and the survey closed on May 6. Sergeants had already received the FIP training about a year prior to survey administration. Of the 4,622 sergeants who were invited to participate, 1,474 (31.9 percent) opened the survey, 1,011 (21.9 percent) responded to at least one section of the survey, and 536 respondents (11.6 percent) reached the end of the survey, though they may have skipped individual items.[182] Thus the number of usable survey responses varies from section to section.

[182] With respect to the response rate on the supervisor survey, we refer to our remarks above concerning response rates.

Interviews

The patrol and supervisor interviews were designed to provide additional insight into the training. Interviews are an especially helpful method for providing context to better understand patterns found in the quantitative data collected through the training day and supervisor surveys, for example. Two different interview instruments were developed – one for officers and one for supervisors. The officer interview protocol prompted respondents to discuss their views on: (1) the need/appropriateness for them personally and the department as an organization to direct attention to the issue of implicit bias; (2) the extent to which implicit bias affects officer behavior and citizen behavior; (3) the forms in which these biases play out in citizen-police interactions; (4) the likelihood that raising awareness of implicit biases could lead to behavioral changes on the street; and (5) whether and how sergeants address issues related to biased policing. The supervisor interview protocol gathered supervisors' perspectives on: (1) the need to monitor subordinates' performance for potential bias; (2) the extent to which supervisors do so; (3) signs/factors that could signal the possibility of biased policing in a subordinate; and (4) whether and how they have intervened with officers exhibiting potentially biased behaviors. Both instruments tapped into one shared topic: the application of strategies for managing implicit bias.

We identified in advance a set of commands from among patrol, transit, and housing and identified, for each, the number of officers and sergeants with whom we would like to speak. We passed that information along to the Strategic Initiatives Bureau, which shared the list with the Office of the Chief of Department, which then tasked command supervisors with identifying the specific individuals to be interviewed. All interviewees worked either the day or evening tour. Interviews were conducted at NYPD Headquarters at two points, the first during the week of May 6-10 and the second during the week of August 12-16. At each point, the same member of the research team conducted one-on-one interviews lasting, on average, 15-20 minutes with 70 interviewees (41 patrol officers and 29 sergeants).

Enforcement

The training is designed to raise officers' awareness of the detrimental effects of implicit bias and to teach them techniques with which they can manage their implicit biases. Post-training, therefore, we would expect officers to perceive and act on suspects' race/ethnicity more circumspectly. Even if the officers are not entirely successful in managing their implicit biases, we hypothesize that they will exhibit less pronounced disparities in enforcement post-training compared to pre-training. "'The

key to this training is your behavior,' Dr. Fridell said. 'We need to make sure that your behavior is not biased.'"[183]

We can assess the change in enforcement disparities using several analytic strategies. If the findings all point in the same direction, we can be confident that our conclusions are not sensitive to the form of analysis. We include below descriptions of our analytical approaches. First, we describe the data. Second, we describe the application of the stepped-wedge design, a form of randomized controlled trial, which was executed by the NYPD in scheduling and delivering the training. We also address the challenge that is presented by changes in some officers' assignments. Third and finally, we describe our analytic approaches to estimating pre-/post-training changes in disparity.

Data

We organized the patrol precincts, housing PSAs, and transit districts into 10 training blocks or clusters. As we explain further below, the blocks were randomized and scheduled for training accordingly. At that time, in late-April of 2018, 16,789 patrol officers and detectives were assigned to the patrol, housing, and transit bureaus. Ultimately, the FIP training was delivered to 11,759 police officers assigned to patrol precincts, 1,449 assigned to transit districts, and 1,533 assigned to housing PSAs – 14,741 officers in all.[184]

For each officer, the NYPD provided event-level data files on enforcement behaviors for the period January 1, 2017 through June 30, 2019:

- stops;
- frisks in stops;
- searches in stops;
- force in stops;
- arrests;
- force in arrests; and
- summonses.

All of these data files included the race/ethnicity, sex, and age (in ordinal categories) of the citizen, in addition to the date, the precinct, the officer's (anonymized) tax ID and command code, as well as the geographic coordinates of the street block on which the event transpired. Personnel data on each officer included his/her rank, assigned command, length of service, age, sex, and race/ethnicity, as well as the date of

[183] Baker, "Confronting Implicit Bias in the New York Police Department."

[184] The NYPD's personnel file includes records on 15,924 sworn personnel who attended the FIP training between May of 2018 and April of 2019, including 293 detectives and 107 sergeants. Of those, 12,248 were in the PSB, 1,612 in the housing bureau, and 1,475 in the transit bureau. Not all of those personnel were assigned to a patrol precinct, housing PSA, or transit district, however.

his/her FIP training, and the command (and training block) to which s/he was assigned at that time.[185] Data on arrests and summonses included the charge, such that we can analyze all such events and subsets thereof (e.g., misdemeanor arrests).

In addition, we collected precinct-level counts of citizen complaints, by month, so that the evaluation would treat complaints as another outcome on which the training has hypothetical effects. Though a complaint is the product of both an officer's behavior, as it is perceived and assessed by a citizen, and a citizen's behavior in choosing to file a complaint, it is an outcome in which make stakeholders are interested.

The Stepped-Wedge Design

The examination of behavioral outcomes relies upon a stepped-wedge randomized controlled trial (RCT) design. The stepped-wedge cluster RCT is a type of crossover design in which clusters of subjects all begin as no-intervention controls, and thereafter cross over permanently from the control group to the intervention group in sequence at randomized, pre-specified points in time.[186] Thus, at the end of the trial, all clusters have eventually crossed over to the intervention group. In this case, precincts, PSAs, and transit districts – hereafter "patrol commands" generically – formed the clusters, which were scheduled for training in order. The design allows for an experimental comparison between patrol commands in clusters receiving the intervention to clusters receiving "treatment as usual" awaiting crossover to the treatment group. This approach is advantageous in that it allows for the implementation of a randomized experimental design under circumstances in which randomization would otherwise be impractical, and because no operational units are permanently consigned to a control group for the duration of the study.

Using this stepped-wedge approach, the research team divided the patrol commands into 10 distinct training clusters, endeavoring to distribute commands in the three bureaus across the clusters and also equalize as much as possible the numbers of officers across clusters. We also examined arrest rates in 2017 (i.e., the number of arrests relative to the number of uniformed officers in each cluster) and formed the clusters to ensure: (a) a randomized process of intervention consistent with the stepped-wedge

[185] We note that our initial design provided for linking enforcement data on individual officers to survey data from the same officers. The plan was to administer a web-based survey to a large sample of officers at a single point in time. It would have required that the NYPD send separate survey links to individual officers, which would have been very resource-intensive. Ultimately, we abandoned this survey plan in favor of a survey administered in class on the day of training, and since all respondents were directed to the same internet address, the survey was anonymous. As we discuss below, it is nevertheless possible to link survey data to enforcement data for a substantial number of officers.

[186] Michael A. Hussey and James P. Hughes, "Design and Analysis of Stepped Wedge Cluster Randomized Trials," *Contemporary Clinical Trials* 28 (2007): 182-191.

design, and (b) that high and low-arrest locations would be included as controls for a sufficiently long period during the stepped-wedge approach so that both high- and low-arrest precincts were included in each training cluster. In short, high- and low-arrest commands (clusters) were included in the early and late onset of treatment for a suitable period of comparison. The 10 training clusters had an average of roughly 1,600 uniformed officers and detectives per cluster.

Figure 4-1, below, illustrates the training schedule (originally) anticipated to include approximately four weeks to train the roughly 1,600 officers in each cluster. Baseline measures (at Step 0) and follow-up measures (at Step 11) add additional parameters to statistical models (and the time period of the baseline and follow-up measures are equivalent to the average duration of the treatment period for the steps in the design). At Step 1, Cluster 1 (Precincts 7, 23, 28, 48, 52, 69, 73, 76, Transit District 4, and PSA 9) moved into the treatment classification (and remain there in all analyses) while all other clusters serve as controls. At Step 2, Cluster 2 (Precincts 6, 10, 26, 43, 79, 81, 83, 88, TD 30, and PSA 2) also became treatment units while all other clusters serve as controls; and so on for each step in the design until all 10 clusters had been trained.

The NYPD largely adhered to the stepped-wedge schedule. Of 14,741 police officers in the training blocks, 13,720 (93.1 percent) were trained with the block to which they were assigned. The principal departures from the training schedule involved (1) a precinct assigned to Treatment block A, 181 of whose officers were trained with Treatment block B; (2) a precinct assigned to Treatment block B, 156 of whose officers were trained with Treatment block A; and (3) 217 officers assigned to 17 different commands scheduled for Treatment blocks A or B and trained in a "make-up" block following Treatment block C. See Figure 4-2, below, for a summary.

Figure 4-1. The Stepped-Wedge Design.

	Step 0		Step 1		Step 2		Step 3		Step 4		Step 5		Step 6		Step 7		Step 8		Step 9		Step 10		Step 11	
Stepped-Wedge Block	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C
Treatment A																								
Treatment B																								
Treatment C																								
Treatment D																								
Treatment E																								
Control A																								
Control B																								
Control C																								
Control D																								
Control E																								

Step 0 = Baseline (pre-training measures) and Step 11 = post-training measures (follow-up period)

Figure 4-2. Execution of the Stepped-Wedge Design: Training as Assigned & Delivered.

	Training as Assigned									
Delivered	T – A	T – B	T – C	T – D	T – E	C – A	C – B	C – C	C – D	C – E
T – A	1201	163	1	3	2	3	2	2	4	1
T – B	186	1189	1	3	2	2	5	0	3	1
T – C	43	22	1414	5	2	4	1	7	3	6
T – A & B	68	149	9	1	2	1	1	0	0	1
T – D	12	3	33	1445	37	0	2	1	6	0
T – E	16	10	6	37	1398	2	0	2	1	2
C – A	3	1	4	2	1	1445	2	0	0	2
C – B	3	2	1	0	0	2	1307	21	1	6
C – C	0	0	9	2	2	0	0	1429	0	1
C – D	1	1	0	0	3	1	0	0	1391	0
C – E	0	0	0	0	0	1	0	0	69	1501

Officer Mobility

One complication in this design is the potential for movement of officers from one command to another between step 1 and step 10 of training delivery. For example, an officer might train with treatment block A, the first cluster to be trained, and later be transferred to and work in a command in a different treatment block. Such movement compromises the analysis of any command as treatment or control in any given step; were this to occur, the pre-training enforcement activity could include activity by trained officers, and/or the post-training enforcement activity could include activity by untrained officers.

We address this issue in two ways. First, from the personnel data we determine the proportion of officers who, for the entire period of the study, were assigned to a command in the same training block with which they were trained. Among the 14,519 officers assigned to commands in our training blocks and still employed with NYPD in June, 2019, more than four-fifths (88.4 percent) had stable assignments: at the conclusion of the FIP training, they had the same assignment that they had at the beginning of the FIP training (or at the time of their initial assignment following appointment).[187]

Second, the enforcement data indicate, for each event, the command to which the officer was assigned at the time of the event. These data enable us to ascertain the activities performed by officers whose reassignments placed them in a different training block:

- officers whose activity while assigned to an untrained command followed their training in a previous command, and

[187] For the 222 officers who had separated from the NYPD by the end of the training, we cannot assess the stability of assignments in this way.

- officers whose activity while assigned to a trained command preceded their training in the command to which they were subsequently assigned.

Activity by the former is "early," performed by trained officers in control commands, while activity by the latter is "late," performed by untrained officers in treatment blocks. See Table 4-1, in which "early" signifies events performed by trained officers prior to the training for that block, and "late" signifies events performed by untrained officers following the training for that block. As the table shows, very small fractions (less than one-half of one percent) of arrests, stops, and summonses were performed by officers who were not trained in that block. Most of the activity associated with reassignment from one command to another was within the training blocks.

Table 4-1. Enforcement Actions Performed by Reassigned Officers Outside of Treatment as Delivered

	Arrests (148,984)		Stops (10,574)		Summonses (62,269)	
Block	Early	Late	Early	Late	Early	Late
T – A	20	0	0	0	34	0
T – B	3	0	0	0	0	0
T – C	24	0	1	0	10	0
T – A&B	1	37	0	3	0	105
T – D	36	9	0	0	14	5
T – E	11	2	1	0	2	0
C – A	10	0	0	0	10	0
C – B	52	4	7	0	22	0
C – C	0	1	0	0	0	1
C – D	0	13	0	0	0	7
C – E	0	15	0	4	0	7

Randomized Balance

We would note that the randomization of the commands in the training blocks eventuated in experimental units that were not entirely equivalent in potentially important respects prior to the training. Table 4-2 shows the racial/ethnic composition of the people arrested, stopped, and issued summonses, respectively, by officers in each training block in the month preceding the FIP training, when all of the commands were in the control condition. Enforcement in the first three treatment blocks involved as subjects people who were disproportionately Black or Hispanic, and in the last three control blocks involved as subjects people who were disproportionately White. As commands in training blocks moved successively into the treatment condition with each step in the design, post-training enforcement by officers in those blocks was compared to enforcement patterns in control blocks that – in the absence of any training effects – tended to involve White subjects.

Table 4-2. Racial/Ethnic Percentages, Step 0.

	Arrests				Stops				Summonses		
	W	B	H		W	B	H		W	B	H
Treatment A	9.8%	62.0%	28.2%		6.9%	64.4%	28.7%		12.1%	65.0%	22.9%
Treatment B	14.5%	61.3%	24.2%		8.1%	75.7%	16.2%		13.1%	63.8%	23.1%
Treatment C	11.0%	62.1%	27.0%		8.4%	72.0%	19.6%		14.2%	62.9%	22.9%
Treatment D	25.3%	54.0%	20.7%		11.5%	74.7%	13.8%		23.2%	61.1%	15.7%
Treatment E	18.1%	38.8%	43.1%		16.4%	54.1%	29.5%		26.7%	43.6%	29.7%
Control A	12.4%	57.7%	30.0%		13.0%	64.1%	22.8%		16.3%	56.0%	27.7%
Control B	14.6%	64.6%	20.8%		11.5%	71.2%	17.3%		20.6%	64.9%	14.4%
Control C	23.4%	59.1%	17.5%		17.0%	67.0%	16.0%		27.9%	55.4%	16.8%
Control D	25.8%	54.6%	19.6%		13.0%	78.3%	8.7%		26.5%	58.5%	15.0%
Control E	32.6%	37.7%	29.7%		18.5%	69.2%	12.3%		35.3%	40.2%	24.6%
Total	18.6%	55.3%	26.1%		12.2%	68.6%	19.2%		21.4%	57.4%	21.2%

In addition, the BWC pilot treatment precincts were not distributed evenly across the blocks of commands formed for the evaluation of the training.[188] One control block included none of the pilot precincts, one treatment block included four pilot precincts, and one control block included five pilot precincts. These precincts are:

Treatment A: 48, 79
Treatment B: 43
Treatment C: 34, 71, 72
Treatment D: 18, 30, 102, 105
Treatment E: 115
Control A: 42
Control B: 47
Control C: 25, 44, 60, 63, 67
Control D: none

[188] Footnote 6 in the Monitor's 8th report included 120 and not 121 among the treatment precincts, though the text of p. 3 of the same report refers to 121 as a treatment precinct.

Control E: 13, 121

Some precincts presumably received BWCs prior to the start of FIP training in May of 2018, while deployment of BWCs to other precincts may have followed (or coincided with) FIP training.

Analytical Plans

We have analyzed enforcement behavior at the aggregate, training block level and also at the level of individual events. We describe each of these analytic approaches.

Block-Level Analysis

We hypothesize that, post-training,

- officers will less frequently take enforcement action against Blacks; and
- officers will less frequently take enforcement action against Hispanics.

The training is not intended to reduce enforcement activity, but insofar as enforcement action against people of color is partly attributable to implicit biases, then the application of bias-management strategies would result in fewer enforcement actions against people of color. The simplest approach to testing these hypotheses is to use the repeated cross-sectional design to analyze pre-/post changes in the counts of the outcomes, e.g., the number of arrests for each training block during the step period, controlling for number of officers. We analyze changes in the counts of events that involve (a) Blacks, (b) Hispanics, (c) people of color (i.e., Blacks or Hispanics), and (d) Whites, respectively. Aggregating all of the data to the training blocks that correspond to the steps yields 120 observations for the period of examination: 12 time periods, including baseline, 10 steps in the design, and the follow-up period, multiplied by 10 blocks, or clusters. The cluster assignment serves as a fixed effect parameter in the repeated cross-sectional framework. Ultimately, this approach allows us to assess whether the average treatment effect corresponds with a change in officer behavior (measured at the cluster level). Separate analyses of counts were conducted for the racial/ethnic categories.

One potential problem with this approach is that each "step" in the design may (and did) encompass time periods of somewhat different lengths, depending on when the training is completed for one cluster and when it begins for the next cluster. While efforts were made to form clusters of equivalent size, we anticipated that day-to-day training schedules would lead to steps of unequal durations. If the time periods are comparable (i.e., only a few days of difference among the steps in the wedged design), this approach will yield the greatest statistical power. Maximum likelihood (negative binomial regressions) using the panel design as well as binomial distributions (the presence of an event or not) were used in the different analytical approaches. We

control for the number of days/weeks in the period by creating a rate per unit of time for each step in the wedge.[189] One strength of this approach is that uneven periods between the steps are smoothed via an exposure (i.e., number of days per unit).

It is important to note that this statistical analysis plan, which corresponds with the randomization plan, is designed to assess whether a change in outcomes is effected immediately following the training and sustained for the duration of the post-training period, and is assumed to be unique only after police commands are trained. It is possible that control commands (yet to be trained) may also experience changes in these outcomes due to the impacts of the broader reforms. In this case, the training effect would likely be attenuated because control conditions may also decline at the same time.[190] Based on this methodological consideration, we conducted (as needed) a series of sensitivity tests, i.e., standard growth curve analyses over time, to assess whether all sites (treatment and control at unique points in time) experienced declines in these outcomes.

Veil-of-Training

For analysis at the level of individual events, we can adapt the logic and analytic strategy of Grogger and Ridgeway's "veil-of-darkness" method of analyzing stops.[191] Grogger and Ridgeway posit that darkness impairs officers' ability to detect the race of drivers, thereby forming a more race-neutral baseline against which the racial composition of daytime stops can be compared. The difference that we would expect to see in the enforcement behavior of unconsciously biased officers between pre- and post-training periods is analogous to the difference that we would expect to see between daylight and darkness in the stops of officers engaged in racial profiling. The training encourages officers to self-impose a veil of neutrality on their perceptions of and interactions with citizens. Just as officers patrolling after darkness falls (in the context of a veil-of-darkness analysis) exhibit less pronounced racial disparities due to a degraded ability to detect motorists' race, so too would trained officers exhibit less pronounced disparities in enforcement due to their efforts to manage their implicit biases. Thus, we might expect to see behavioral manifestations of officers' applications

[189] More specifically, a count per day exposed, to control for differences in the days at risk between the steps in the blocked training schedule.

[190] A similar issue was seen in Kochel and Weisburd's study that attempted to assess the impact of community perceptions on procedural justice implementation in randomized police hot spots. A general decline in citizen complaints was observed in both treatment and control conditions, likely because the St Louis County Police Department mandated an overall change, and thus the hot spots procedural justice implementation could only have so much additional impact on citizen complaints. Tammy Kochel and David Weisburd, "Assessing Community Consequences of Implementing Hot Spots Policing in Residential Areas," *Journal of Experimental Criminology* 13 (2017): 143-170.

[191] Jeffrey Grogger and Greg Ridgeway, "Testing for Racial Profiling in Traffic Stops from Behind a Veil of Darkness," *Journal of the American Statistical Association* 101 (2006): 878-887.

of their FIP skills in any of the forms of enforcement behavior that we examine. We therefore hypothesize that, post-training,

- officers will be less likely to take enforcement action against Blacks; and
- officers will be less likely to take enforcement action against Hispanics.

Stops. Our "veil of training" analysis turns on the estimated likelihood that a Black person or a Hispanic person would be stopped, relative to that of Whites, post-training rather than pre-training. In place of the daylight/darkness variable in Grogger and Ridgeway's model we substitute the pre-/post-training variable, based on the date of the stop and the date on which the officer's cluster completed the training. As Grogger and Ridgeway do, we control statistically for place (precinct) and time (day of week and time of day), as well as other factors.

Summonses & Arrests. Lacking data on incidents in which officers choose not to invoke the law even though they have evidence of offending, we have to approach the analysis of summonses and arrests in much the same way that we approach the analysis of stops. We analyze individual summonses and arrests, respectively, predicting the race/ethnicity of the suspect in terms of officers' training status at the time and other factors (place and time). All analyses of arrests can be repeated on subsets – e.g., misdemeanors and other lower-level offenses – with the expectation that the less serious the offense, the larger the training impact, or the more likely we would detect training impacts.

Frisks & Searches. Frisks and searches could be analyzed using the approaches described above, for stops and arrests, but such an analysis would not reflect the fact that frisks and searches are contingent on stops. If, say, we found that frisks of Hispanics were less frequent or less likely (relative to Whites) post-training, it would not necessarily mean that, once stopped, a Hispanic is less likely post-training to be frisked; it might simply be that they are less likely to be stopped.

Instead, we analyze frisks and searches, respectively, as binary outcomes among all stops. The key predictor is the cross over to treatment by the officer's training cluster. We cannot control for legal factors that justify a frisk or a search; we must count on the stepped randomization to experimentally control for such factors. We can control statistically for place (precinct) and time (day of week and time of day), and we also control for the type of offense that led to the stop, and whether the stop stemmed from a citizen report (communicated directly to the officer at the scene or through dispatch) or was initiated by the officer.

Use of Force. Data on use of force come from two sources – stop forms and arrest reports – in different forms, so we analyze use of force in each context. Moreover, arrest reports include some information on subject resistance, but the stop forms do not; in the analyses of force in the latter, therefore, we must rely on experimental controls for this potentially confounding factor. We analyze use of force as we do frisks

and searches: as binary outcomes among all stops/arrests, with the key predictor the intervention cross-over indicator.

Summary

The evaluation design includes a number of components to better capture the range of outcomes on which implicit bias training may have effects, and the conditions under which those effects are amplified or muted. The design includes two RCTs: one executed through the randomized administration of pre- and post-training surveys on the days of FIP training, and the other a stepped-wedge RCT on the basis of which to estimate behavioral effects. A follow-up survey of trainees was administered, both to assess decay in the effects of the training on officers' implicit bias beliefs and to describe their self-reported use of bias-management strategies. A survey of sergeants was administered to capture an important feature of the organizational context. The NYPD and the FIP trainers cooperated in executing the design, a feat that should not be underestimated.

Chapter 5
Impacts on Officers' Beliefs and Attitudes

If the training has effects, its most immediate effects are on officers' awareness of and knowledge about implicit bias, their motivation to minimize the consequences of implicit bias, and their understanding of bias-management strategies. The training-day survey included items in terms of which we can measure and analyze these outcomes, on which much of this chapter dwells. We first describe the items and the measures that we formed, and then present the results of our analysis of the differences between pre- and post-training responses, from which we infer training impacts. We also analyze officers' post-training assessments of the utility of the training, and their projections of the likelihood that they would use the bias-management strategies that the FIP training covers. We supplement the findings from the analysis of the training-day survey data with findings from semi-structured interviews with officers. Finally, we examine the longer-term impacts of the training – i.e., the potential decay in the immediate training impacts.

Measures of Officers' Beliefs and Attitudes

The survey includes a number of items that tap officers' beliefs about implicit bias and attitudes toward discrimination. All of these items take the form of a statement with which respondents may agree or disagree on a five-point scale: strongly agree; agree somewhat; neither agree nor disagree; disagree somewhat; strongly disagree. A response set of this kind is quite common in survey research, capturing both the valence – agree or disagree – and intensity of belief or attitude. We have formed several indices from these individual items by assigning numerical scores to the responses and summing the scores across multiple items.

One concern about survey responses is the possibility that some participants choose responses that they perceive as more socially desirable, rather than their true beliefs. Although we cannot dismiss this possibility, the likelihood of such a social desirability response is minimized by the use of a web-based, anonymous survey, in which respondents have no personal interaction with an interviewer, and their identities remain unknown. Furthermore, social desirability responses are likely to be randomly distributed across pre- and post-training respondents.

The survey instrument included 27 items on beliefs about implicit bias and attitudes toward discrimination. From these items, we formed six indices, each of which combines multiple items by summing the quantified, valid survey responses to the constituent items.[192] For a few indices, our measure of scale reliability (Cronbach's

[192] Some respondents skipped individual survey items. Generally, 96-97 percent of the respondents who answered any of the items that comprise an index answered every item. Thus, very little of analytical value

alpha) is rather low (i.e., under 0.70). We note that our findings are the same for all but one of the individual survey items. Additional analysis of officers' beliefs concentrates on their responses to individual items, rather than the index scores.

Implicit Bias 1

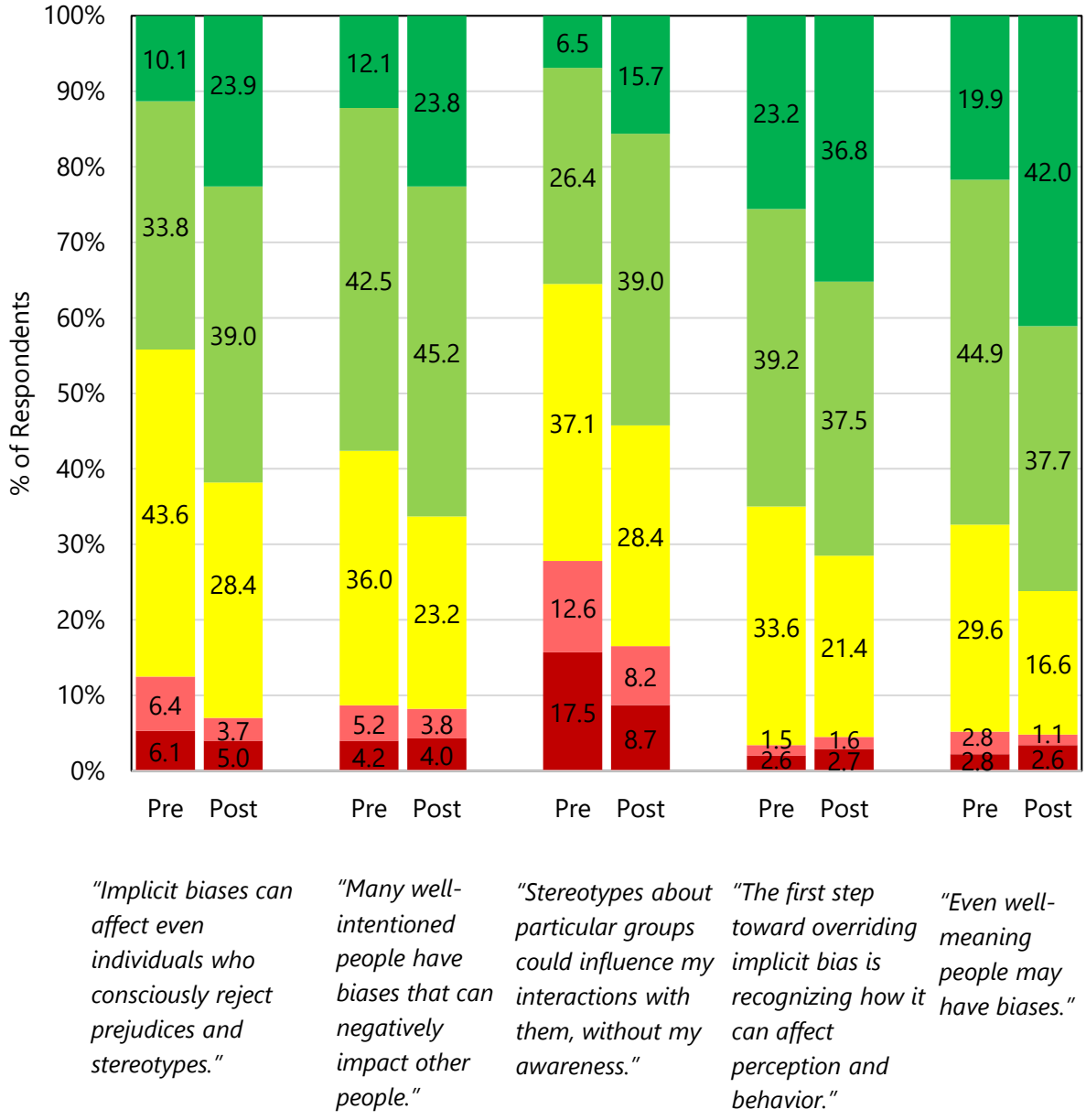
Eight survey items measure beliefs about implicit bias – concerning either the nature or the implications of implicit bias – that relate directly to the FIP training content. Agreement with each of these statements represents a "correct" response – that is, an answer that is consistent with the training content, and a belief that officers would be (hypothetically) more likely to hold following the training. Figures 5-1 and 5-2 summarize the responses for pre- and post-training respondents. Note that in these figures and those to follow in this chapter, shades of red indicate disagreement, yellow indicates neither agreement nor disagreement, and shades of green indicate agreement.

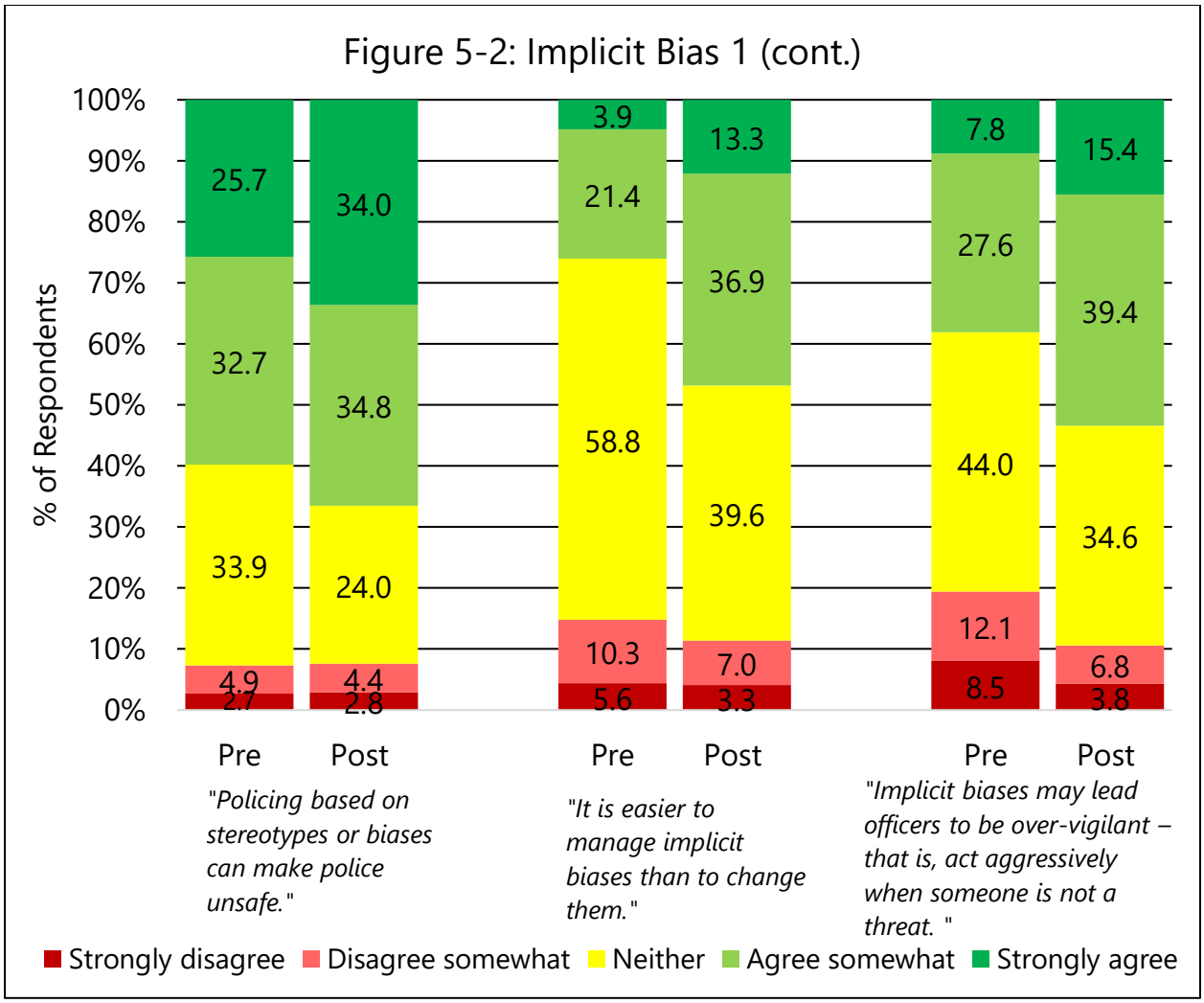
Many officers – one-third to two-thirds – responded in accordance with training content even prior to the training, indicating some baseline level of knowledge. Larger proportions of post-training respondents agree with the statements, consistent with the expectation that the training would increase their knowledge. We combined these eight items to form a single index, labeled "Implicit Bias 1," analysis of which we report below in analyzing training impacts.[193] Higher scores on the index correctly reflect FIP training content, and we hypothesized that post-training scores would be higher than pre-training scores.

can be gained by imputing values to replace the skipped items, and we analyze index scores for only those respondents for whom we have a complete set of responses to items that form an index.

[193] A factor analysis of these items yields one factor with an Eigenvalue greater than 1.0, explaining 42.5% of the variance. They have an alpha of 0.798.

Figure 5-1: Implicit Bias 1

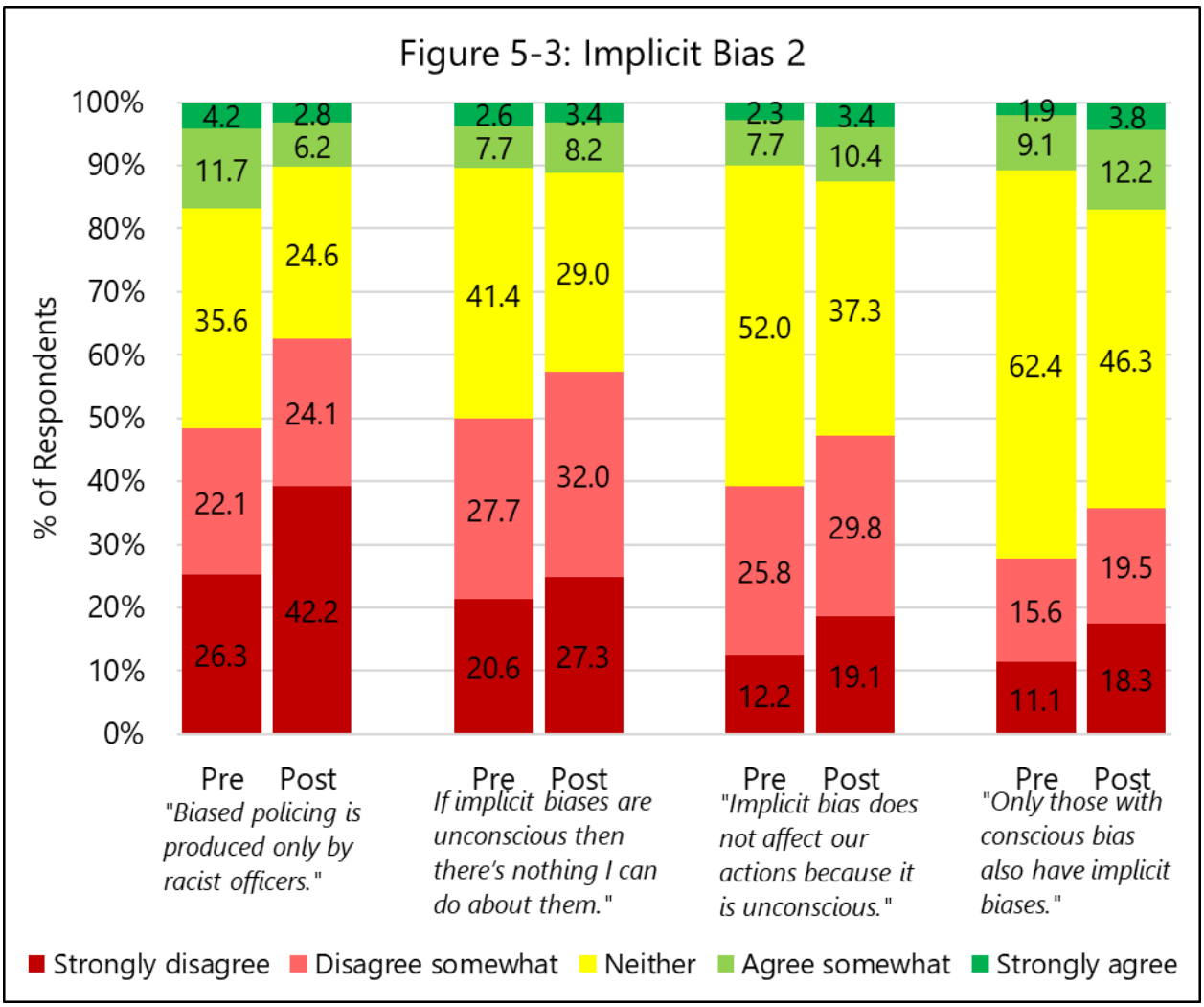




Implicit Bias 2

Four items measure beliefs that are – according to the FIP training – incorrect. We would expect that the training would lead officers to disagree with these statements, and we hypothesized that post-training scores on this index would be *lower* than pre-training scores.[194] As shown in Figure 5-3, while about one-third to half of the pre-training respondents disagreed with the statements, somewhat larger percentages of the post-training respondents disagreed, as hypothesized, though the percentages who agreed are also slightly larger post-training.

[194] A factor analysis of these items yields one factor with an Eigenvalue greater than 1.0, explaining 50.2% of the variance. They have an alpha of 0.658.



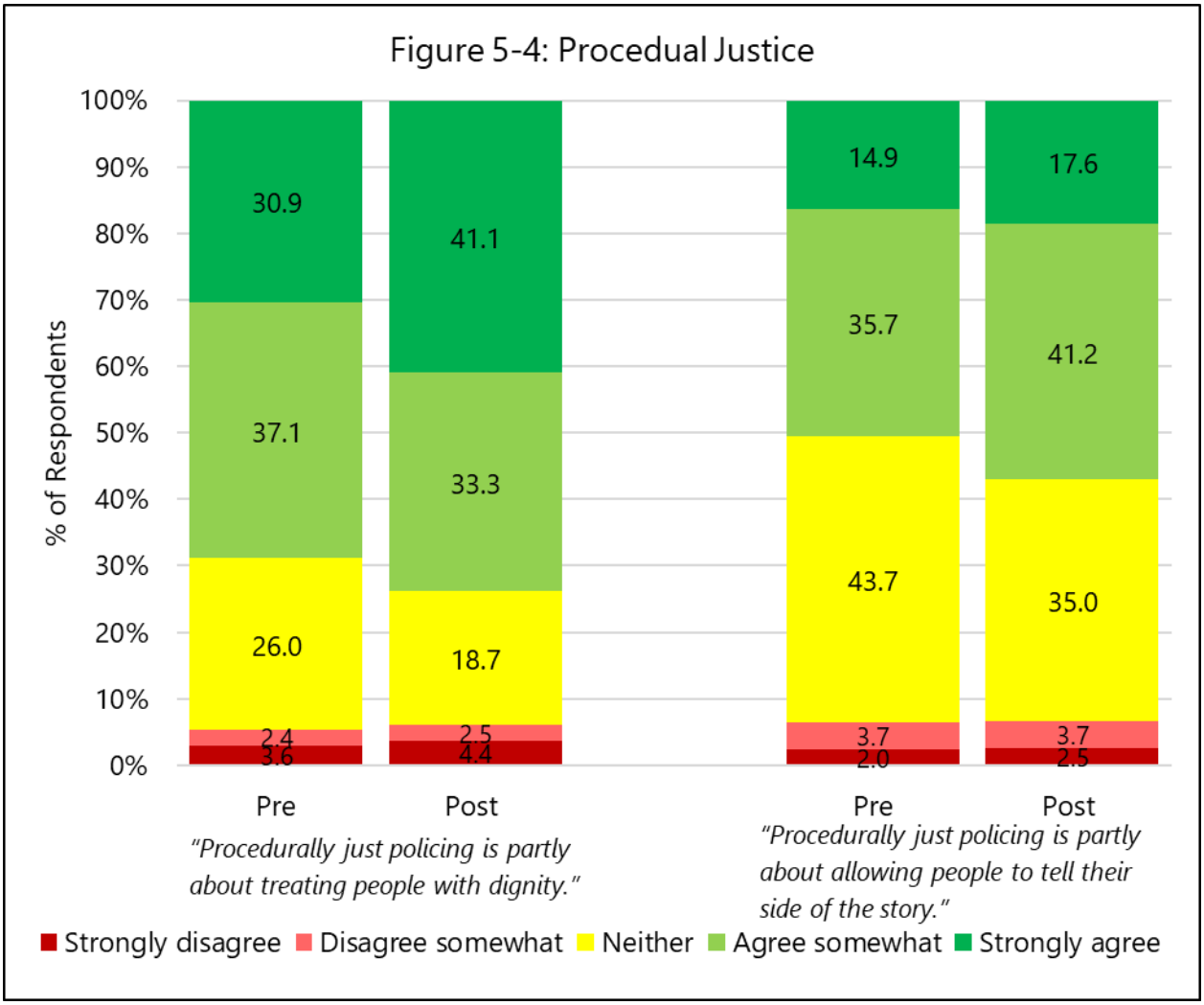
Procedural Justice

Two items capture responses about procedural justice, which is also covered in the FIP training.[195] In general, the procedural justice with which police act turns on the extent to which an officer (1) treats citizens with dignity and respect, (2) allows the citizen an opportunity to explain her situation or express her viewpoint, (3) demonstrates a concern about the citizen's well-being, and (4) establishes that the officer's decisions are based on facts and neutral criteria.[196] One survey item concerns

[195] A factor analysis of these items yields one factor with an Eigenvalue greater than 1.0, explaining 66.9% of the variance. They have an alpha of 0.502.

[196] See Steven L. Blader and Tom R. Tyler, "A Four Component Model of Procedural Justice: Defining the Meaning of a 'Fair' Process," *Personality and Social Psychology Bulletin* 29 (2003): 747-758.

treating people with dignity. The second item has to do with allowing people to explain themselves.



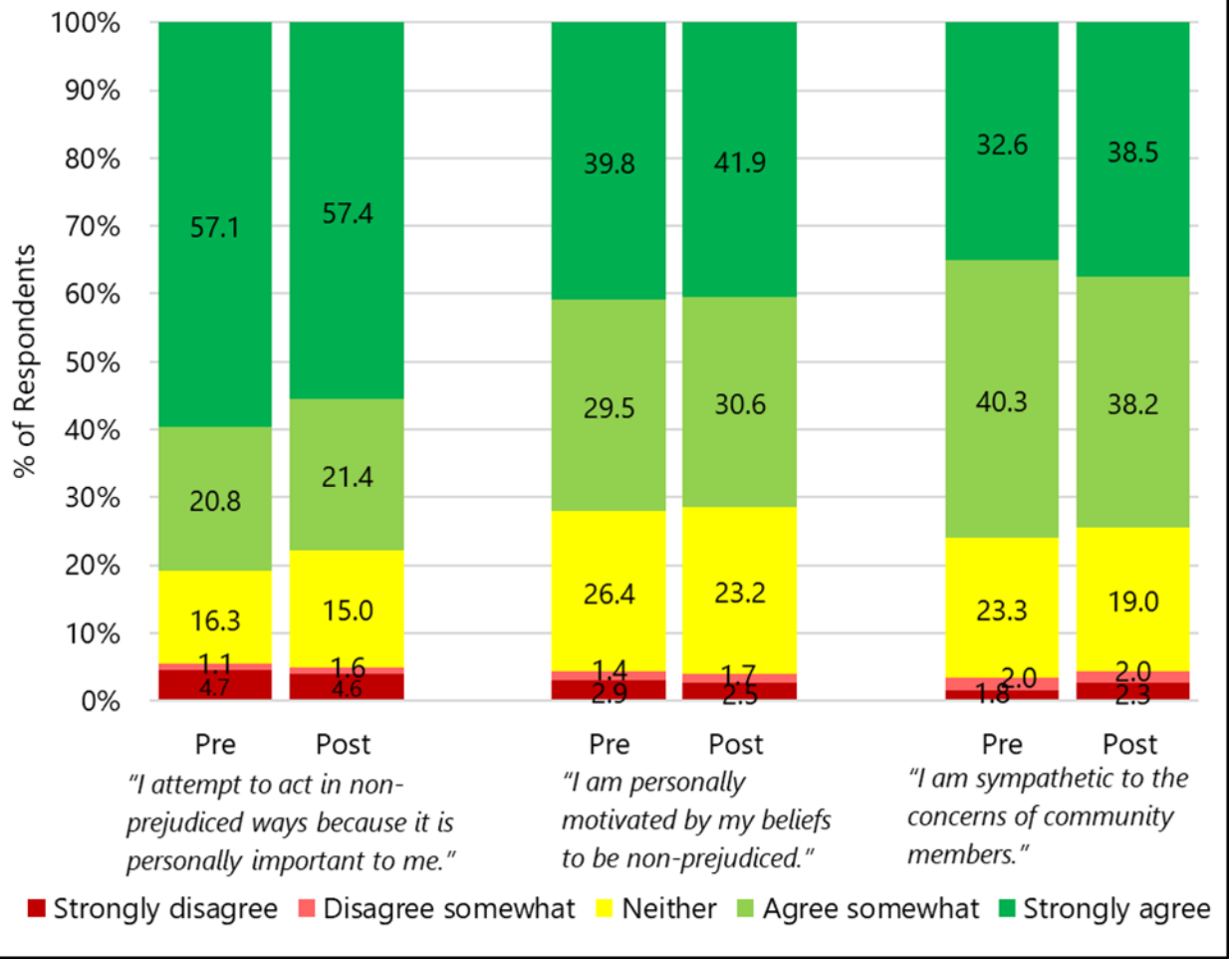
A comprehension of these matters serves to demonstrate a broader understanding of the FIP curriculum's subject and significance. Figure 5-4 summarizes the survey results. Slightly larger proportions of post-training respondents correctly agreed with these statements, over and above a fairly high baseline of comprehension.

Motivation to Respond without Prejudice

Two items were drawn from a previously established scale measuring motivation to respond without prejudice (discussed in Chapter 2), and a third bears a conceptual

and empirical relationship to that construct.[197] As Figure 5-5 shows, most respondents – 70 percent or more – exhibit a motivation to act without prejudice both before and after the training. Three quarters indicated that they are sympathetic to the concerns of the community. Higher scores on the index represent higher levels of motivation to respond without prejudice. We hypothesize that insofar as the FIP training informs trainees about the nature and consequences of implicit bias, post-training scores on this index will be higher than pre-training scores. Moreover, insofar as motivation to act without bias may be a contingency on which training effects on behavior turn, the levels of motivation detected here bode well for training efficacy; the audience appears to be largely receptive to the training intent.

Figure 5-5: Motivation to Respond without Prejudice



[197] The previously established scale is discussed in E. Ashby Plant and Patricia G. Devine, "Internal and External Motivation to Respond Without Prejudice," *Journal of Personality and Social Psychology* 75 (1998): 811-832. A factor analysis of these items yields one factor with an Eigenvalue greater than 1.0, explaining 61.0% of the variance. They have an alpha of 0.676.

Concern about Bias

Three items reflect a concern about discrimination and bias, in general or with respect to policing in particular.[198] Figure 5-6 summarizes officers' responses. About two-thirds of the respondents consider racial and ethnic discrimination to be a serious social problem, while less than ten percent disagree, before and after the training. A broad concern for bias may inspire motivations to respond without bias, with which it is correlated in these survey data. Further, it may be amplified by officers' knowledge about implicit biases, with which it is also correlated in these data. Concern about bias in the specific context of policing is less prevalent, but it is more prevalent among post-training respondents. This is what we would expect to find if the training raises officers' awareness of the nature and implications of implicit bias.

Skepticism about Bias

Several items reflect not concern but skepticism about the significance of discrimination and bias, in general or with respect to policing in particular.[199] We hypothesize that insofar as the FIP training informs trainees about the nature and consequences of implicit bias, post-training scores on this index will be lower than pre-training scores. As shown in Figure 5-7, fewer than half of the respondents expressed such skepticism, and as we would expect if the effects of the training extend to officers' attitudes, smaller proportions of the post-training respondents selected the skeptical responses. For example, one-third of the pre-training respondents agreed that "bias in the police profession is merely a 'fiction' produced by the media," while about 30 percent disagreed. The percentage of post-training respondents who agreed with this statement was somewhat lower, and the percentage who disagreed somewhat larger.

[198] A factor analysis of these items yields one factor with an Eigenvalue greater than 1.0, explaining 64.5% of the variance. They have an alpha of 0.721.

[199] A factor analysis of these items yields one factor with an Eigenvalue greater than 1.0, explaining 40.5% of the variance. They have an alpha of 0.504. This factor was distinct from the factor that represented a concern about bias, though in theory we would expect concern and skepticism to lie at opposite ends of a single dimension.

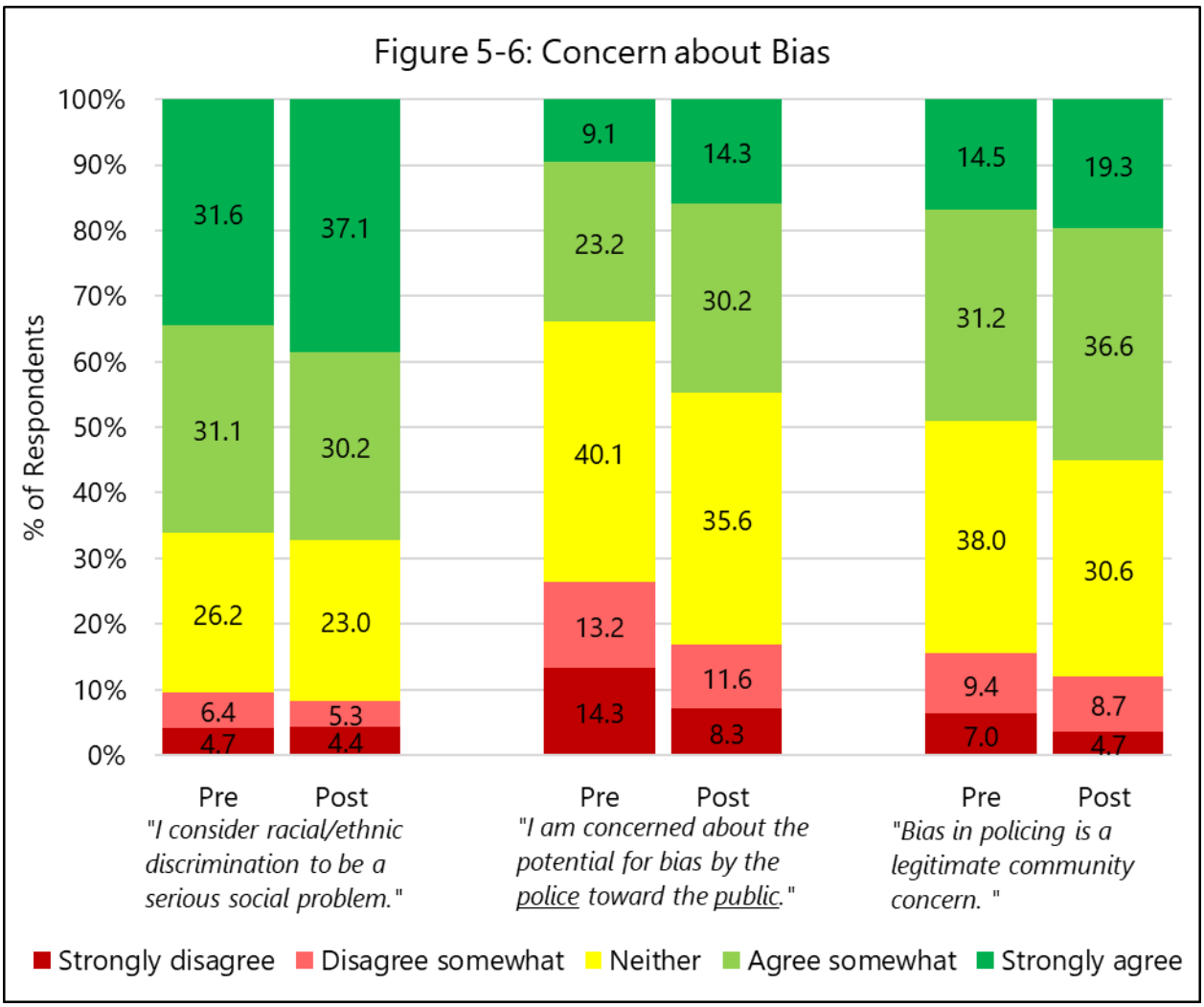
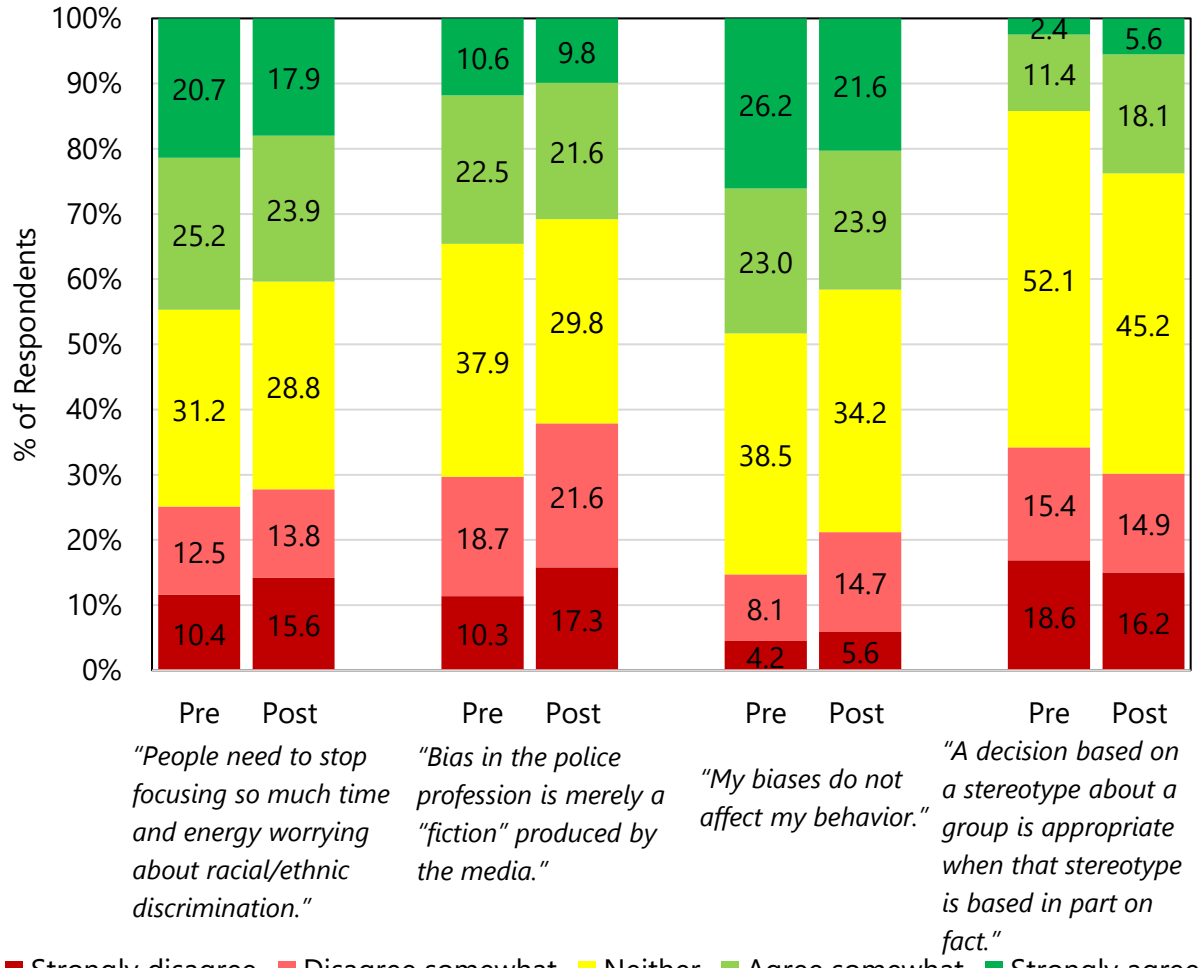


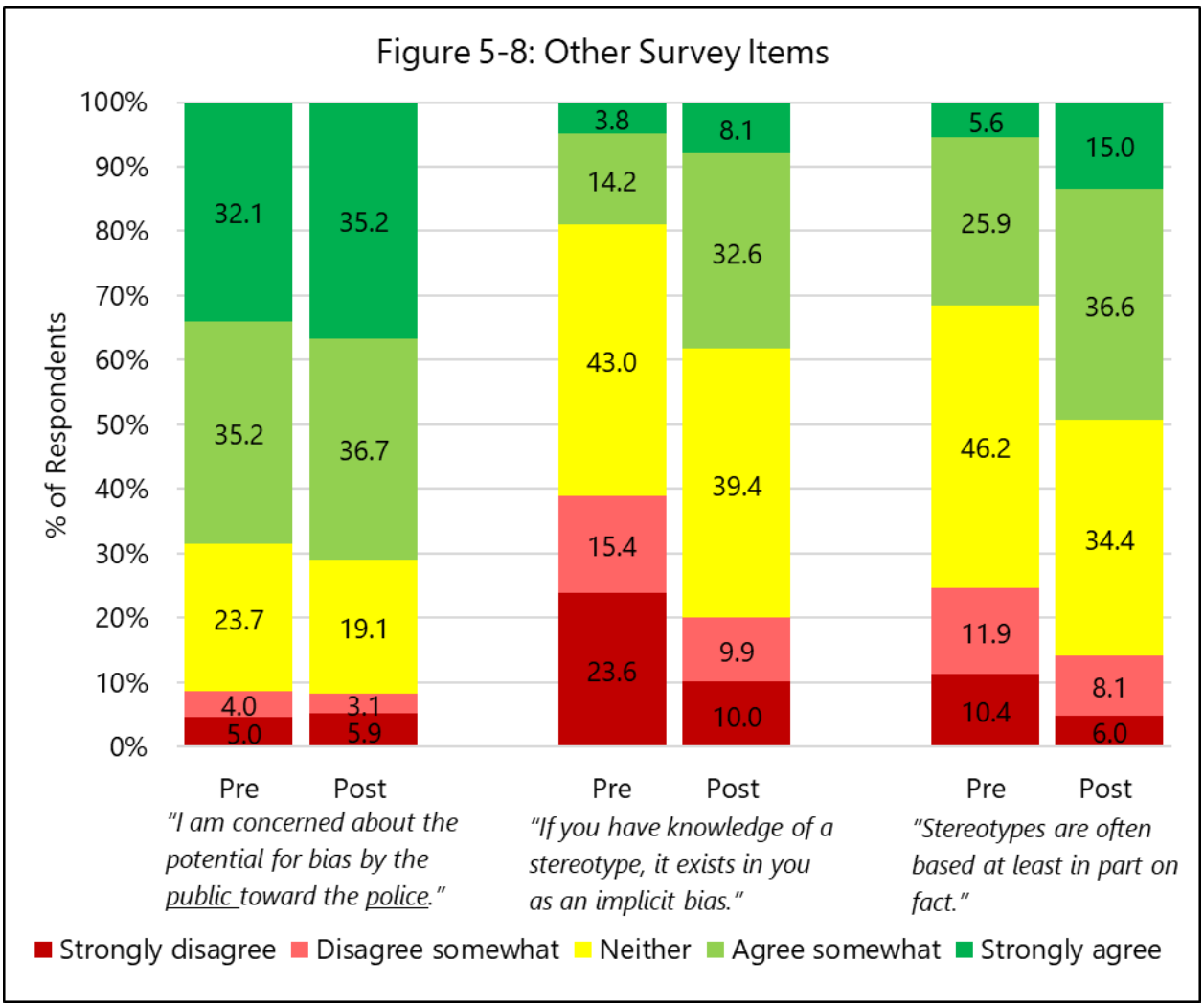
Figure 5-6: Concern about Bias
% of Respondents
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
31.6
31.1
26.2
6.4
4.7
37.1
30.2
23.0
5.3
4.4
9.1
23.2
40.1
13.2
14.3
14.3
30.2
35.6
11.6
8.3
14.5
31.2
38.0
9.4
7.0
19.3
36.6
30.6
8.7
4.7
Pre
Post
"I consider racial/ethnic discrimination to be a serious social problem."
Pre
Post
"I am concerned about the potential for bias by the police toward the public."
Pre
Post
"Bias in policing is a legitimate community concern. "
Strongly disagree
Disagree somewhat
Neither
Agree somewhat
Strongly agree

Figure 5-7: Skepticism about Bias



Other Items

Based on the correlations among the survey items, three items were not a part of any of the sets of items that formed an index; they are shown in Figure 5-8. Two-thirds of the respondents, pre- and post-training, indicated that they are concerned about the potential for bias by the public toward the police; a small fraction of respondents reported that this was not a concern. Recognition that exposure to a stereotype may suffice to form unconscious bias was greater among post-training respondents, as we would expect, along with an understanding that stereotypes are often based, in part, on facts.



Influences on Beliefs and Attitudes

Officers' beliefs about implicit bias and attitudes toward discrimination are patterned to some degree by their backgrounds and characteristics. We regressed each of the scales on respondents' self-reported attributes: years of service with the NYPD; rank (detective or police officer); educational achievement; military experience (yes or no); sex; age; and race/ethnicity. We analyzed pre-training responses and post-training responses separately, allowing for the training to affect the patterns of relationships.

Holding the remaining attributes constant statistically, officers' years of service bears a small but consistent and (with few exceptions) statistically significant relationship to their beliefs and attitudes (see Tables 5-1 and 5-2). In general, prior to the training, the *less* experienced officers had a slightly better knowledge of implicit bias and procedural justice, slightly greater concern about discrimination (and correspondingly less skepticism), and a slightly greater motivation to respond without

prejudice. The same patterns held among post-training respondents, though the effects on neither skepticism nor procedural justice remained statistically significant.

Table 5-1. Regression Analysis of Pre-Training Respondents' Beliefs and Attitudes.

	IB1	**IB2**	**Motiva-tion**	**Concern**	**Skepti-cism**	**PJ**
Years exp	-0.013*	0.007**	-0.018*	-0.017*	0.007**	-0.018*
Detective	0.072	-0.143**	0.003	0.038	-0.100	0.095
Education	0.024*	-0.049*	0.007	0.003	-0.029*	0.011
Military exp	-0.004	-0.015	0.024	-0.047	-0.040	0.021
Female	0.042	-0.028	0.087*	0.125*	-0.049	0.043
Age	0.055*	-0.038*	0.056*	0.085*	-0.061*	0.049*
White	-0.278*	0.170*	0.098**	-0.694*	0.510*	0.016
Hispanic	-0.153*	0.111*	0.042	-0.340*	0.318*	-0.011
Asian	-0.097**	0.336*	-0.105	-0.300*	0.477*	-0.015
Other race	-0.189*	0.183*	0.151*	-0.549*	0.408*	0.060

Table 5-2. Regression Analysis of Post-Training Respondents' Beliefs and Attitudes.

	IB1	**IB2**	**Motiva-tion**	**Concern**	**Skepti-cism**	**PJ**
Years exp	-0.006**	0.009*	-0.010*	-0.015*	0.005	-0.006
Detective	-0.014	-0.280*	0.137	0.052	-0.299*	0.053
Education	0.049*	-0.038*	0.042*	0.047*	-0.025*	0.047*
Military exp	0.051	-0.117*	0.051	0.037	-0.103*	0.040
Female	0.026	-0.072*	-0.037	0.069**	-0.045	-0.014
Age	0.016	-0.035*	0.023	0.060*	-0.027	0.025
White	-0.212*	0.121*	-0.124*	-0.709*	0.534*	-0.027
Hispanic	-0.125**	0.142*	-0.092*	-0.318*	0.380*	-0.031
Asian	-0.180*	0.305*	-0.333*	-0.426*	0.473*	-0.166*
Other race	-0.272*	0.168*	-0.208*	-0.502*	0.424*	-0.111**

Detectives were for the most part indistinguishable from police officers, with two exceptions: (1) detectives were slightly less likely to agree with the incorrect statements for implicit bias 2, and (2) detectives expressed slightly less skepticism about discrimination as a social problem (especially post-training). The same pattern of relationships held for officers with military experience.

Educational achievement was associated with greater comprehension, both pre- and post-training. Post-training, the effects of education also manifested themselves in

somewhat greater concern about discrimination and motivation to respond without prejudice. It may be that the training had a larger effect on more educated officers.

Female officers, pre-training, differed from men only with respect to attitudes – concern about discrimination and motivation to respond without prejudice – and not beliefs, or knowledge. Post-training, the difference between men and women with respect to motivation to respond without prejudice vanished.

Though length of service and age are positively correlated, they are not perfectly correlated, and controlling for length of service, age exhibited somewhat different relationships to officers' beliefs and attitudes. Older officers were, pre-training, both more knowledgeable about implicit bias, more concerned about discrimination, and more motivated to respond without prejudice. Post-training, most of those differences were of smaller magnitude and statistically insignificant.

The largest differences in beliefs and attitudes emerged among officers of different races and ethnicities. Given the nature of the attitude and belief objects, this is scarcely surprising. The reference category for the regression analysis included Black officers, against whom officers of each of the other race/ethnicity categories were compared. Pre-training, White officers were less knowledgeable about implicit bias, and less concerned and more skeptical about discrimination as a social problem, but more motivated to respond without prejudice. All of the same effects held among the post-training respondents, except that, post-training, Whites were less motivated to respond without prejudice. For the most part, similar differences hold between Black officers and those of other races or ethnicities: Hispanics, Asians, and others (e.g., mixed-race).

Immediate Training Impacts on Beliefs and Attitudes

One approach to estimating training impacts on officers' awareness of implicit bias, given the randomized nature of survey administration, compares the means of the indices for the post-training respondents to those of the pre-training respondents (see Table 5-3). All of the index means vary from 2.3 to 4.2. The table also includes the percentage of index items that were answered "correctly" in terms of training content, the sample size (or N), and the statistical likelihood that a difference of means as large as that observed would occur by chance – i.e., the statistical significance of the difference of means.[200] The next-to-bottom row shows the estimated difference on the 5-point index scale, and the bottom row includes a common measure of effect size,

[200] We treat three constructs – motivation to respond without prejudice, concern about discrimination, and skepticism about discrimination – as predominantly matters of personal judgment with no clearly correct or incorrect answers.

Cohen's d.[201] Conventional interpretations of values for Cohen's d hold that a value of 0.8 is large, 0.5 moderate, and 0.2 small. All of the differences achieve statistical significance at the conventional .05 level (all but one are statistically significant at the .01 level), so we can conclude with a high degree of confidence that the post-training respondents answered differently, on average, than the pre-training respondents.[202]

Table 5-3. Comparison of Pre- and Post-Training Means

	IB1	IB2	Motivation	Concern	Skepticism	PJ
Pre- mean	3.42	2.57	4.09	3.38	3.15	3.74
% "correct"	47.4	40.3	NA	NA	NA	59.5
Pre- N	2721	2722	2774	2754	2724	2730
Post- mean	3.77	2.36	4.14	3.59	3.05	3.87
% "correct"	64.4	53.3	NA	NA	NA	66.8
Post- N	3883	3870	3920	3901	3861	3887
H_0 no difference: p <	0.00	0.00	0.02	0.00	0.00	0.00
Δ, 5-pt scale	0.35	-0.21	0.05	0.21	-0.10	0.13
Cohen's d	0.58	-0.30	0.06	0.24	-0.14	0.17

Though we can infer that changes occurred in the immediate aftermath of the training, the changes are not dramatic. The largest average difference, on the first implicit bias index, is less than 0.4 on the 4-point scale, from 3.42 to 3.77.[203] The effect size of 0.58 is considered of moderate magnitude. The next largest difference, on the second implicit bias index, is of small to moderate magnitude, and the remaining effect sizes are small.

While we have good reason to believe that the randomized survey administration controls for differences between groups, we performed additional analyses to confirm that the differences that we detected are not confounded by other factors. Table 5-4 reports the results of regression analyses of each of the indices, in which we statistically control for differences in officers' backgrounds. These findings show that while background factors have a bearing on the index scores, the estimated effects of the

[201] See Jacob Cohen, *Statistical Power Analysis for the Behavioral Sciences* (New York, Routledge, 1988). Also see Daniel Lakens, "Calculating and Reporting Effect Sizes to Facilitate Cumulative Science: A Practical Primer for t-tests and ANOVAs," *Frontiers in Psychology* 4 (2013): 863.

[202] Analyses of individual survey items showed that the pre-/post-training difference in only one item was statistically insignificant: "I attempt to act in non-prejudiced ways because it is personally important to me."

[203] Four of the survey items were identical to those in the instrument that the Urban Institute applied in the six National Initiative sites. Across 7,428 respondents, they too found small differences, i.e., no greater than 0.3 on the 5-point scale. See Jannetta et al., *Learning to Build Police-Community Trust,* op cit., Figure 2.3, p. 26.

training parallel those in the difference-of-means analysis; in Table 5-4 these effects are shown in grey shading. For example, on the Implicit Bias 1 index, the pre-/post-training difference of means is 0.356, while the estimated difference is 0.357 in the regression analysis that controls statistically for background differences. Some divergence could stem from the fact that the regression excludes those for whom background data are missing. However, all of the differences estimated with controls for background characteristics are quite congruent with the simple differences of means tests reported previously.

Table 5-4. Regression Analyses of Officers' Beliefs and Attitudes

	IB1	**IB2**	**Motivation**	**Concern**	**Skepticism**	**PJ**
Education	0.038*	-0.045*	0.027*	0.028*	-0.028*	0.031*
Military experience	0.023	-0.074*	0.035	0.003	-0.078*	-0.026
Length of service	-0.010*	0.007*	-0.015*	-0.016*	0.005	-0.012
Female	0.026	-0.053*	0.007	0.088*	-0.050*	-0.001
Age	0.039*	-0.033*	0.044*	0.073*	-0.036*	0.041
Detective	-0.036	-0.182*	-0.015	-0.022	-0.189*	-0.010
White	-0.240*	0.142*	-0.031	-0.700*	0.525*	-0.009
Hispanic	-0.136*	0.129*	-0.039	-0.327*	0.355*	-0.022
Asian	-0.151*	0.314*	-0.248*	-0.377*	0.470*	-0.112
Other race	-0.235*	0.175*	-0.063	-0.522*	0.417*	-0.046
Post-training	0.357*	-0.219*	0.045	0.189*	-0.087**	0.133*
N	5736	5729	5793	5773	5723	5745
Pre-/post-training difference of means	0.356*	-0.212*	0.046	0.211*	-0.095*	0.126*

Note: Black is the excluded category of race/ethnicity

In the context of the regression analysis, we allowed the estimated effect of the training to vary across racial/ethnic groups. With two exceptions, the effects of the training do not appear to differ across racial/ethnic groups. First, the effect of the training on the Implicit Bias 1 construct among White trainees is somewhat greater than others. The average effect among other racial/ethnic groups is 0.321, and for Whites it is 0.089 higher, a difference that is statistically significant (but substantively quite small). Second, the effects of the training on trainees' motivation to respond without prejudice is greater among Blacks than others. Indeed, the effect holds only among Black trainees; the average training effect among trainees of other races/ethnicities is

statistically insignificant. Pre-/post-training differences in means for the various racial/ethnic groups are shown in Table 5-5.

Table 5-5. Pre-/Post-Training Differences in Means by Racial/Ethnic Group

	IB1	**IB2**	**Motivation**	**Concern**	**Skepticism**	**PJ**
White (pre- N=1152-1166; post- N = 1385-1401)						
Pre- mean	3.33	2.59	4.13	3.16	3.28	3.75
Post- mean	3.74	2.34	4.15	3.35	3.18	3.90
H_0 no difference: p <	0.00	0.00	0.47	0.00	0.00	0.00
Δ, 5-pt scale	0.41	-0.25	0.02	0.19	-0.10	0.15
Black (pre- N=279-285; post- N = 444-449)						
Pre- mean	3.64	2.40	4.08	3.91	2.74	3.76
Post- mean	3.95	2.21	4.26	4.07	2.64	3.91
H_0 no difference: p <	0.00	0.00	0.02	0.01	0.13	0.02
Δ, 5-pt scale	0.31	-0.19	0.18	0.16	-0.10	0.15
Hispanic (pre- N = 684-705; post- N = 954-963)						
Pre- mean	3.47	2.53	4.09	3.54	3.07	3.73
Post- mean	3.81	2.36	4.17	3.74	3.02	3.88
H_0 no difference: p <	0.00	0.00	0.04	0.00	0.19	0.00
Δ, 5-pt scale	0.34	-0.17	0.08	0.20	-0.05	0.15
Asian (pre- N = 222-225; post- N = 282-288)						
Pre- mean	3.54	2.76	3.95	3.58	3.24	3.74
Post- mean	3.77	2.51	3.95	3.63	3.11	3.75
H_0 no difference: p <	0.00	0.00	0.96	0.44	0.04	0.85
Δ, 5-pt scale	0.23	-0.25	0.00	0.05	-0.13	0.01
Multi-racial (pre- N = 144; post- N = 155-160)						
Pre- mean	3.49	2.54	4.30	3.42	3.11	3.85
Post- mean	3.88	2.22	4.16	3.77	2.95	3.96
H_0 no difference: p <	0.00	0.00	0.10	0.00	0.07	0.21
Δ, 5-pt scale	0.39	-0.32	-0.14	0.35	-0.17	0.11
Other race (pre- N = 103-106; post- N = 182-187)						
Pre- mean	3.25	2.66	3.93	3.13	3.22	3.59
Post- mean	3.46	2.56	3.95	3.29	3.24	3.60
H_0 no difference: p <	0.01	0.25	0.85	0.17	0.84	0.97
Δ, 5-pt scale	0.21	-0.10	0.02	0.16	0.02	0.01
Not reported (pre- N = 132-144; pos-t N = 454-477)						
Pre- mean	3.30	2.54	3.98	3.21	3.10	3.69
Post- mean	3.71	2.41	4.12	3.61	3.07	3.83
H_0 no difference: p <	0.00	0.06	0.09	0.00	0.64	0.08
Δ, 5-pt scale	0.41	-0.13	0.14	0.40	-0.03	0.14

Further examination of these data yields additional evidence in terms of which the nature and magnitude of the immediate training effects can be assessed. Tabulating the correct and incorrect responses – that is, responses that are consistent and inconsistent with training content, respectively – item by item (see Table 5-6), it is clear that:

- Pre-training, baseline levels of understanding varied across the statements, from 25.3 to 65.2 percent correct;
- The differences in the percentages of correct responses between pre- and post-training respondents vary in magnitude, from 10.2 to 24.8, reflecting varying impacts of the training;[204]
- On most of the items, small fractions of pre-training respondents chose a response that is inconsistent with the training (incorrect);
- The differences in the percentages of incorrect responses between pre- and post-training residents vary in magnitude, up to 13.2, though not all are in the expected direction.

Across all of the items, the percentage correct went from 45.0 to 60.7. On only five of twelve items did two-thirds of the post-training respondents select a "correct" response. The largest pre-/post-training difference was on item 22, with which 25.3 percent of pre-training respondents and 50.1 percent of post-training respondents correctly agreed. This was a key point of the training, so this is remarkable. The smallest difference was on item 19, on which the baseline level of understanding was fairly high (58.6 percent). In general, and intuitively, we see larger pre-/post-training differences on the items with lower baseline levels ($r = -0.5$). We also see smaller differences with respect to items for which the correct response was disagreement (4, 8, 24, and 26). Differences on the other side of the coin, in percentages *incorrect*, were much smaller overall: 12.7 percent incorrect among pre-training respondents and 9.8 percent among post-training respondents. Officers exhibited lower baseline levels of misunderstanding, such that much of the difference of the training was in imparting knowledge to some of those who simply did not (or professed not to) know (i.e., the "neither agree nor disagree" response option), and not in correcting misunderstandings. The principal exceptions involved the consequences of implicit bias (items 10 and 23).

[204] There is some evidence of response set, as 229 respondents each selected the "somewhat disagree" response to all 27 items in section 2 of the survey, and 15 others consistently disagreed strongly; only 1 respondent consistently agreed strongly (and none who agreed somewhat with all 27 statements). This pattern tends to inflate correct answers at the margin, and is prevalent at equivalent levels in pre- and post-training; it may also account for some of the apparent inconsistencies in substantive responses.

Table 5-6. Training-day and Follow-up Survey Responses, Implicit Bias Items

	CORRECT RESPONSES				INCORRECT RESPONSES				
Item #	Pre-TD%	Post-TD%	TD diff	F-U%	Pre-TD%	Post-TD%	TD diff	F-U%	
13	65.2	79.9	14.7	69.1 - 72.7	5.5	3.7	-1.8	5.7 - 7.3	Even well-meaning people may have biases
24	38.3	49.3	11	39.5 - 45.5	9.8	13.3	3.5	10 - 11	Implicit bias does not affect our actions because it is unconscious
9	55.1	69.4	14.3	60 - 64.6	9.3	7.7	-1.6	8.7 - 11.1	Many well-intentioned people have biases that can negatively impact other people
5	44.3	63.1	18.8	50.5 - 55.5	12.2	8.3	-3.9	13 - 15.4	Implicit biases can affect even individuals who consciously reject prejudices and stereotypes
26	26.9	38.1	11.2	33.4 - 38.5	10.7	15.3	4.6	10.8 - 11.7	Only those with conscious bias also have implicit biases
10	32.8	54.8	22	36.9 - 41	30.2	16.9	-13.3	29.4 - 33	Stereotypes about particular groups could influence my interactions with them, without my awareness
4	48	67.1	19.1	54.6 - 57.9	16.1	8.5	-7.6	14.3 - 15.3	Biased policing is produced only by racist officers
19	58.7	69	10.3	59.3 - 64.8	7.5	7.2	-0.3	8.5 - 11	Policing based on stereotypes or biases can make police unsafe
23	35.7	54.7	19	36.4 - 42.3	20.7	10.7	-10	20.9 - 25.5	Implicit biases may lead officers to be over-vigilant – that is, act aggressively when someone is not a threat
12	63	74.5	11.5	65.2 - 70.7	3.8	4.1	0.3	4.8 - 6.7	The first step toward overriding implicit bias is recognizing how it can affect perception and behavior
22	25.3	50.1	24.8	31.8 - 34.8	15.8	10.2	-5.6	15.3 - 17.2	It is easier to manage implicit biases than to change them
8	48.6	59.6	11	49.3 - 53.3	10.4	11.2	0.8	12.5 - 14.4	If implicit biases are unconscious then there's nothing I can do about them
Mean	45.2	60.8	15.6		12.7	9.8	-2.9		

Table 5-7. Training-day and Follow-up Survey Responses, Attitude Items

	AGREEMENT				DISAGREEMENT				
Item #	Pre-TD%	Post-TD%	TD diff	F-U%	Pre-TD%	Post-TD%	TD diff	F-U%	
Motivation									
3	78.6	79.1	0.5	77.3	5.7	6.1	0.4	8.4	I attempt to act in non-prejudiced ways because it is personally important to me
11	69.8	72.8	3	72.4	4.1	4.1	0	6.1	I am personally motivated by my beliefs to be non-prejudiced
14	73	76.8	3.8	73	3.7	4.2	0.5	6.6	I am sympathetic to the concerns of community members
Concern									
15	62.9	67.3	4.4	58.5	11.1	9.7	-1.4	19.5	I consider racial/ethnic discrimination to be a serious social problem
20	32.7	44.2	11.5	34.4	27.5	20	-7.5	34.5	I am concerned about the potential for bias by the police toward the public
21	46	55.8	9.8	44.1	16.3	13.6	-2.7	26	Bias in policing is a legitimate community concern
Skepticism									
7	45.9	41.7	-4.2	47.2	23.3	29.6	6.3	25.8	People need to stop focusing so much time and energy worrying about racial/ethnic discrimination
17	33.1	31.2	-1.9	33.8	29.3	39.1	9.8	34.2	Bias in the police profession is merely a "fiction" produced by the media
18	49.4	45.4	-4	49	12.2	20.4	8.2	15.4	My biases do not affect my behavior
27	13.6	23.3	9.7	19	34.4	31.4	-3	42.1	A decision based on a stereotype about a group is appropriate when that stereotype is based in part on fact

If we treated these twelve implicit bias items as a final exam, and set the passing threshold at 65 percent, then 32.8 percent of pre-training respondents would have passed even without the training and 56.3 percent of the post-training respondents would have passed. (With a passing threshold set at 75 percent, 21.4 percent of pre-training respondents and 43.5 percent of post-training respondents "passed.") From one perspective, the post-training passing rate represents notable improvement from the pre-training baseline. From another perspective it is disappointing, inasmuch as nearly half of the officers did not demonstrate a comprehension of the basic implicit bias concepts sufficient to achieve a minimally passing score.

Smaller differences in pre- and post-training responses appear with respect to items that formed scales of officers' attitudes – items that are not only factual and involve personal judgment (see Table 5-7). The differences were in the expected directions – post-training respondents were more concerned about discrimination and more motivated to respond without prejudice – but of smaller magnitude, with few exceptions under 10 percentage points. We would expect that attitudes would be less susceptible to change, and moreover, baseline levels of concern and especially motivation were fairly high.

Anticipated Application: FIP Strategies

The FIP training introduces officers to several strategies for or approaches to managing their unconscious biases, and the survey includes items that prompt respondents to indicate their likelihood of using those strategies, namely:

- managing biases – that is, recognizing one's biases and engaging in bias-free behavior;
- avoiding profiling by proxy – that is, being aware of community members' biases and using one's own judgement;
- reducing biases – that is, seeking ways to have positive contact with individuals who are different from oneself;
- slowing it down – that is, checking one's initial impressions and collecting more information to better understand the situation; and
- engaging with community members, so that one has more positive contact with them.

Respondents rated their anticipated likelihood of using each strategy on a scale from 1 (not at all likely) to 7 (very likely). Figure 5-9, below, summarizes their responses.

One-third of the officers said that they were maximally likely (7 on the scale) to use all five strategies, and nearly half rated the likelihood of using all five either a 6 or 7 on the scale. Officers' projected likelihood of using each strategy is, unsurprisingly, correlated with their awareness of and knowledge about implicit bias – the implicit bias

and procedural justice indices – and also with their concern about discrimination and their motivation to respond without prejudice. Any assessment of officers' actual application of the strategies must await the analysis of the follow-up survey, which inquired about officers' use of these approaches; we report those findings below.

Officers' Perceptions of the Training

The post-training survey directly captured officers' perceptions of the FIP course and the instructors, as well as the utility of the training generally. The FIP course and instructors received very favorable ratings (see Table 5-8, below). More than half of the respondents rated the course as excellent, overall, and nearly 80 percent rated it as at least four on a five-point scale. More than 80 percent of the respondents rated the instructors as excellent.

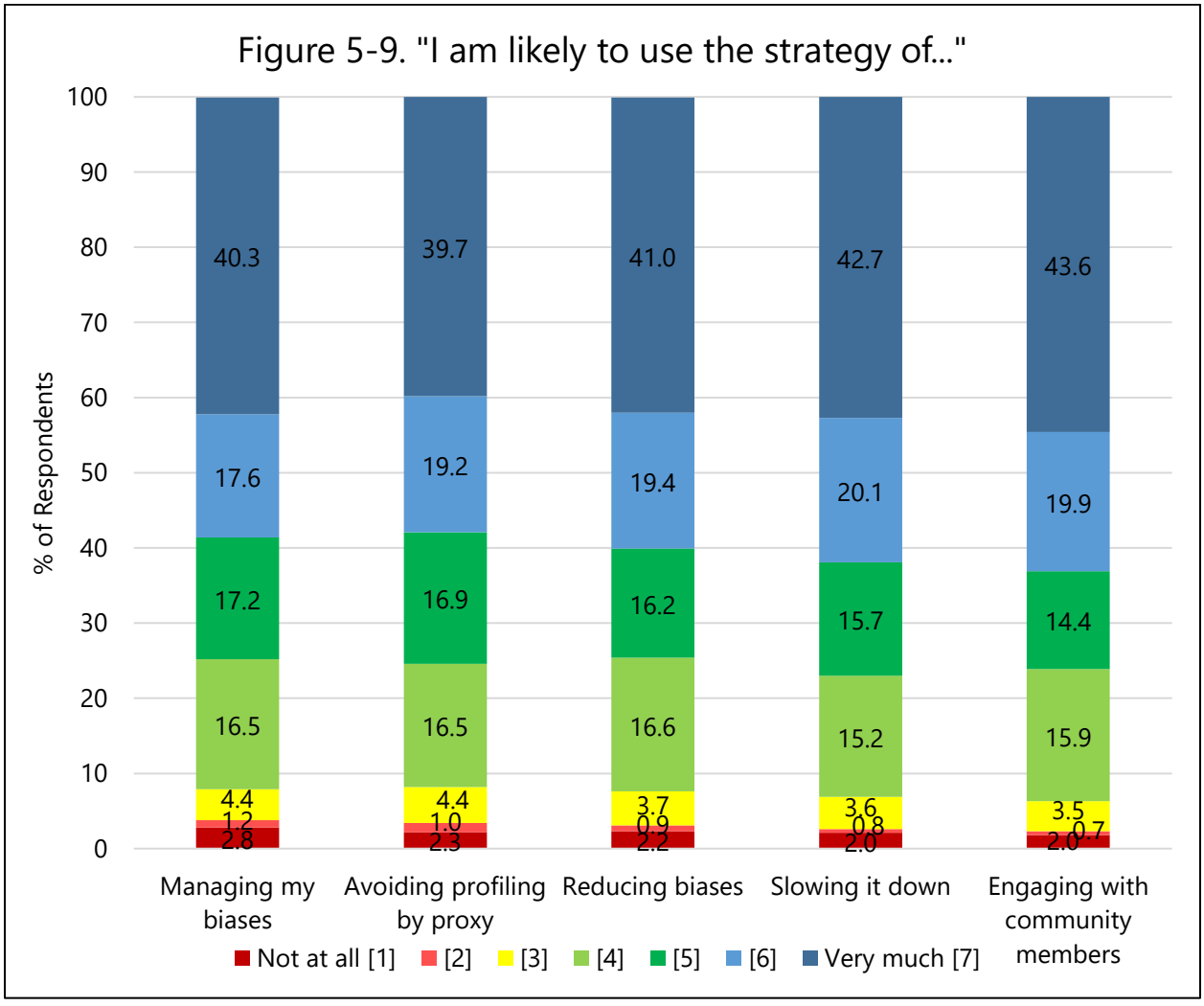
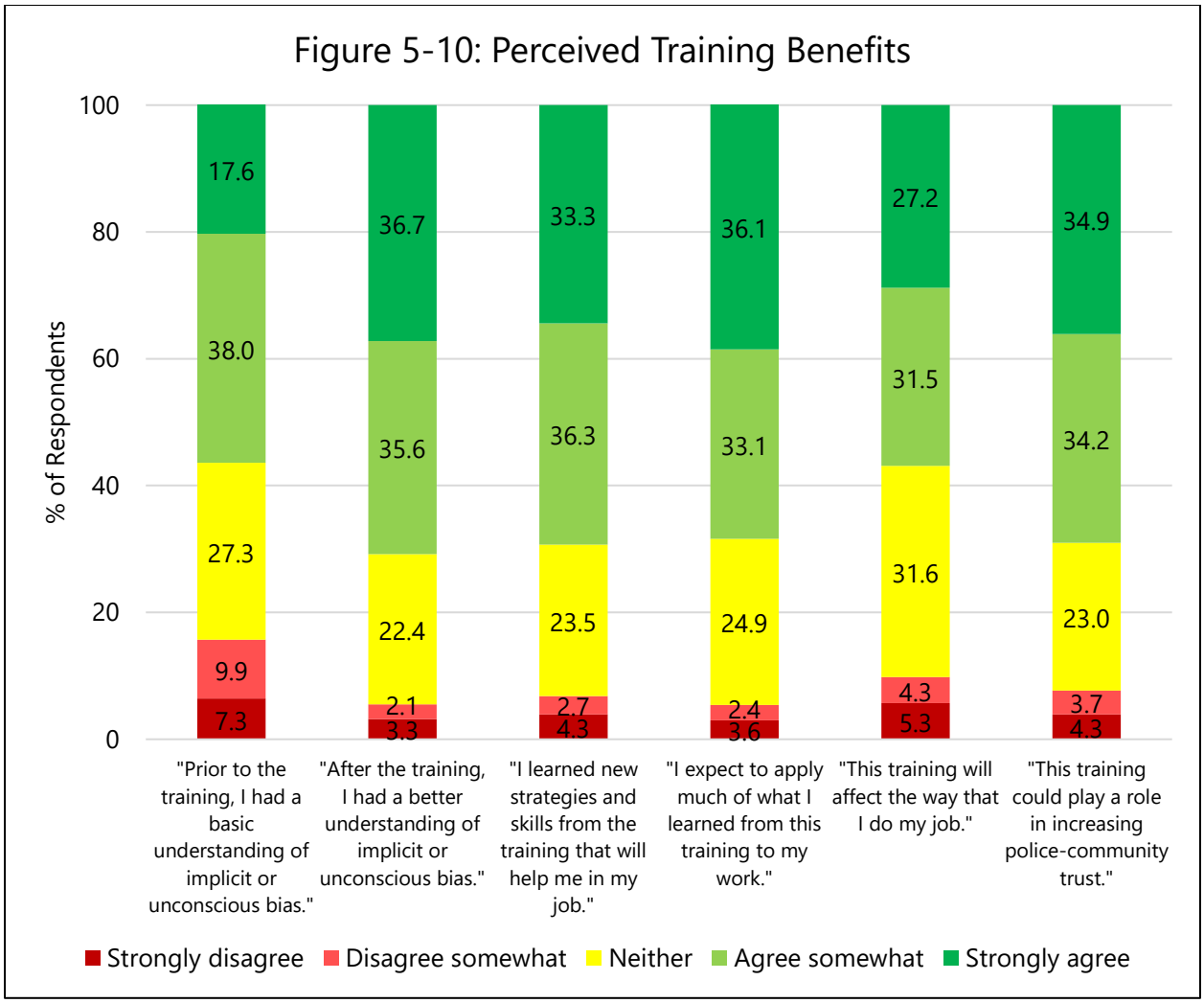


Table 5-8. Ratings of the FIP Training and Instructors

	[1] Unsatisfactory	[2]	[3]	[4]	[5] Excellent
How would you rate the Fair and Impartial Policing course overall?	2.9	2.7	15.2	25.2	54.1
How would you rate the instructors (professionalism, knowledge and preparedness)?	0.8	0.7	6.8	9.3	82.3

The post-training respondents were asked to assess the utility of the training. As Figure 5-10 (below) shows, more than half indicated that they already had a basic understanding of implicit bias prior to the training. Even so, 70 percent reported gaining a better understanding of implicit bias following the training. Moreover, more than two-thirds of the respondents reported learning new strategies and skills, which they expected to apply to their work. Nearly 60 percent responded that the training would affect how they do their jobs.[205]

[205] Some logical inconsistencies appeared in the responses. For example, twenty-nine respondents (3 percent) agreed that they expected to apply much of what they learned in the training and disagreed that the training will affect the way that they do their job. Likewise, an additional 98 respondents (10 percent) agreed that they expected to apply much of what they learned in the training and neither agreed nor disagreed that the training will affect the way that they do their job.



Longer-Term Impacts on Beliefs

We originally hypothesized that training effects on officers' beliefs (or knowledge) about implicit bias would decay over time, a hypothesis to be tested by treating the post-training survey (on the day of training) as a baseline against which follow-up survey responses could be compared. Recognizing the potential for non-response bias, we have weighted the follow-up survey to replicate the distribution of attitudes on the training-day surveys, with one set of weights for each of the ten attitude items; see Table 5-6, which shows a range of follow-up survey percentages defined by the weighted results. Attitudes, as we noted above, were largely unaffected by the training; weighting the follow-up survey respondents in such a way as to reproduce the attitudinal distributions on the training-day survey serves to adjust for – or at least mitigate – the non-response bias. Thusly weighted, the ranges are not wide and generally exhibit a pattern of differences from the training-day survey that are consistent with the hypothesis of decay. In fact, the drop from training-day to follow-up

in the percentages of correct responses on the implicit bias statements rises with the difference between pre- and post-training responses: the greater the training difference on the day of training, the greater the difference between post-training and follow-up. The follow-up "test" scores (percentage correct) do not differ significantly by training block, suggesting that the decay is fairly rapid. Nor is the decay complete, however; we see evidence of some longer-term retention of training content.

Semi-structured interviews with officers months following their training, during the spring and summer of 2019, provide some additional depth to the survey-based findings. All but one of the 42 officers with whom we spoke reported receiving the FIP training offered by NYPD (one could not recall). Very few officers described the training in negative terms and a number, despite not being prompted to describe the trainers or the course more generally, indicated that they found the trainers to be engaging and effective.

> *"The presenters did a good job getting their point across. It was a good training, but it was no major shocker."*

> *"It was an interesting training and the instructor was very entertaining. Enjoyable."*

> *"It was a good training. Retired officers from other agencies spoke. It was nice to hear from officers from other agencies speak about stories from their experiences."*

We asked officers, "Did the FIP training raise your awareness and add to your knowledge about implicit bias?" The majority (75 percent) responded in the affirmative, while the remainder indicated that it did not provide them with information over and above what they knew going into the training. When pressed to describe more specifically what they had learned, officers generally indicated having either further developed prior knowledge or learning completely new information. For some, the training enabled them to place information or vague knowledge they previously had into a larger context, for example, learning terms and research they had not previously known or how to be more aware and mindful of their actions and what could be driving them. Examples of what we are characterizing as the acquisition of new knowledge included: learning that all people hold implicit biases, learning how to be alert to biases, that biases can be managed or controlled, and gaining a better understanding of the diversity of associations that drive implicit biases.

When asked if they had taken real-world strategies learned in the classroom and applied them on the streets, responses were not consistent with the pattern we found in response to raising knowledge and awareness. The majority of officers indicated the FIP training did raise knowledge and awareness about implicit bias. Despite prompting,

most did not elaborate on how it did (or did not). Among those who did elaborate, we heard:

> *"I went into it [the training] thinking that there are certain biases, but I learned about other biases – like, you can have biases about older people."*
>
> *"I knew about it [implicit bias] going in, but it is always beneficial looking at numbers and research factors."*
>
> *"I had a general idea, but there were terms I didn't know."*
>
> *"One point I learned is that white officers were less likely to shoot a Black person, and Black officers are more likely to shoot a Black person. If you are the same demographic you are more likely to shoot. If there are opposite, ethnicities, it's less likely. I think hesitation and second-guessing is horrible."*
>
> *"Yes. It [the training] taught to never judge a book by its cover. You can't always react by the appearance of someone. You have to always investigate."*

When asked if the training offered practical real-world strategies that could be taken from the classroom and applied on the streets, the majority (71 percent) indicated they did not believe the training offered practical real-world strategies for their application on the streets of New York City. From those who indicated that they had not applied FIP strategies we heard that the examples used to teach the strategies did not translate to the New York City environment:

> "*They gave us an example of a car stop in Colorado where the guy got out of the car with a rifle. I guess guns are common there so the officer can just keep issuing commands. Not here*"
>
> "*The trainers were from other agencies, but they don't really understand how it was for NYPD. … The outside perspective was nice, but it was dampening because we weren't connecting on the same level."*

Some officers did not recall that the training curriculum included a strategy component. Others acknowledged that strategies were taught but did not regard them as personally useful because the officer either did not have any biases, did not allow biases to influence their professional interactions, or was not assigned to a unit that afforded the opportunity to apply the strategies. Among the smaller group of officers who indicated

that they had applied strategies, most elaborated that the application of the strategies in their work pre-dated the training.

Even though we detected among officers with whom we spoke little overt opposition to the training, and that most acknowledged that it had raised their awareness and knowledge about implicit biases, we also detected a disconnect (i.e., less support) for the practical utility of the training. In effect, while most officers with whom we spoke found the training interesting and many went so far as to compliment the trainers, far fewer found it truly useful. We believe this can be explained in large part by their views about the context of New York City policing and NYPD officers.

We asked officers a set of general questions not directly tied to the FIP training, including: "Do you believe that implicit bias is something about which police departments should be concerned?"; "Do you believe that implicit bias is something about which you as an officer should be mindful or concerned?"; and "Do you think that implicit biases could affect officers' behavior on the street?" After each question, officers were prompted to elaborate on their answers. Officers' responses to these general questions serve to place their perspective of the FIP training into context.

One perspective offered by some officers suggests that those who grew up in, live in, or have spent a large amount of time working in diverse areas believe that they are less likely to have biases or to let biases influence their behavior on the street. Officers' own words serve to describe this sentiment:

> *Officers are human. It's like when I say, 'Boys will be boys. Girls will be girls.' But you come to the city and boys are dressed like girls and girls are dressed like boys. It can be a culture shock you are not from the city.*
>
> *When I work with someone newer, I try to let them know that it* [implicit bias] *exists. Some people weren't raised in New York City so they don't know the dynamics.*
>
> *We are culturally desensitized to ethnicity. I don't feel there are any problems related to bias. I work in an area that is predominantly Black and Hispanic.*
>
> *We are a melting pot. You may not be aware of bias, it's where you grew up. … Officers may be from different boroughs, Long Island, West Chester, … They don't understand the culture of others from different places. I had an officer say to me that guys were hanging out on a street corner. I said, 'didn't you hang out at a 7-11 when you were younger?' He said yes, but that it was different. I told him it wasn't any different. They may not be doing anything. This is just where they hang out.*

Growing up in New York … it's so diverse here. All my friends are mixed. In my friend group I have my black friend, my Indian friend … This is not Canada. It's not all white and Swedish. This is normal in New York City.

Others allowed that while they personally did not "*have a biased bone in their body,*" they could imagine some officers who might learn something from the training that would guide their behavior. Others indicated the training did not have practical utility, as NYPD officers already acted in a professional manner, not allowing biases to drive decisions. This helps to explain why they would see the training as informative but not personally useful in application.

Conclusions

Officers who completed the training had favorable views of it. Moreover, it appears that officers are, on average, more knowledgeable about implicit bias following the training. These findings are consistent with the hypothesized impacts of the training on officers' beliefs and attitudes. Furthermore, officers reported that they expect to apply what they learned during the training, including the FIP skills in managing implicit bias. However, the impacts of the training on officers' beliefs and attitudes are of no greater than moderate magnitude, and most are quite small even if they are statistically significant – i.e., very likely larger than zero. These findings are similar to those reported by the Urban Institute on its evaluation of the implicit bias training delivered as part of the National Initiative.[206] Results across the six agencies varied, but in the aggregate, the pre-/-post-training differences were of no greater magnitude than those in the NYPD.

Officers emerged from the training with greater knowledge about implicit bias than they had at the start of the training, but on average, their comprehension of the training content was incomplete. Furthermore, and as we would intuitively expect, some of the gains on the day of training decayed over time, and appear to have decayed fairly quickly. Such findings are reminiscent of the findings of studies of other implicit bias interventions, reviewed in Chapter 2, many of which report small effects of short duration.

[206] Jannetta et al., *Learning to Build Police-Community Trust.*

Chapter 6
Organizational Reinforcement: Supervisors' Role and Practices

As we noted previously, no one should suppose that a one-day training in implicit bias would by itself eliminate or minimize the effects of officers' unconscious biases. The organizational context could reinforce and even amplify the influence of the training; or the context could contradict, undermine, or negate the influence of the training. Among the many features of organizational context that potentially impinge on officers' management of their unconscious biases – including the executive's tone, departmental policies, recruitment and hiring practices, programs designed to engage with communities, and others discussed in Chapter 2 – the practices of field supervisors are the most immediate and perhaps among the most important. Supervisors mediate the application of policies and implementation of programs; follow-through can be facilitated by supervisors' support or stymied by their opposition.

The FIP training for supervisors reflects the recognition that supervisors play a unique role in bias-free policing. They are trained in the sources of data in terms of which they can monitor officers' performance for signs of bias, and in how to respond – with the required delicacy – when potential bias is detected. They are also trained in how to communicate about bias, and reminded that as role models, their words and deeds carry weight with their subordinates. Supervisors can reinforce the training with periodic reminders to be mindful of the potential impacts of implicit bias and the need to manage it. We surveyed sergeants, who at the time of the survey had already received the FIP training, about 22 percent of whom responded to at least one section of the survey, and 12 percent of whom reached the end of the survey.[207]

We begin by assessing the extent to which supervisors embrace the role that the FIP training prescribes for them, followed by an examination of supervisors' role enactment, or how they detect potential bias in patrol behaviors, and address and communicate about implicit bias. We caution that our analysis of the supervisor survey data should not be construed as evidence of the impact of the training on supervisors. We have no data on supervisors' beliefs, attitudes, or practices prior to the training that they received, and we cannot draw inferences about changes over time. We summarize data on supervisors' self-reported application of the training, but any inferences about training effects turn on respondents' own judgments about their individual baselines. Our purpose in administering the supervisor survey and analyzing these data was to illuminate the extent to which the training is reinforced in the field.

[207] As noted in Chapter 3, respondents may have skipped individual survey items, and the number of usable survey responses varies from section to section.

Role Conception

The FIP training stresses that supervisory monitoring is critical to detect and address potential bias in officers' street behaviors. We therefore asked sergeants to respond to survey items that concern whether they believe that monitoring is part of their supervisory role. Sergeants were prompted to indicate the extent to which they agreed or disagreed with four statements:

- It is part of my job to monitor the performance of subordinate MOS [members of service] in terms of potential bias;
- It is appropriate to expect supervisors to discuss bias with their subordinates;
- Biased policing is something that supervisors can impact;
- Supervisors can provide direction to subordinates to ensure that they act in a fair and impartial manner.

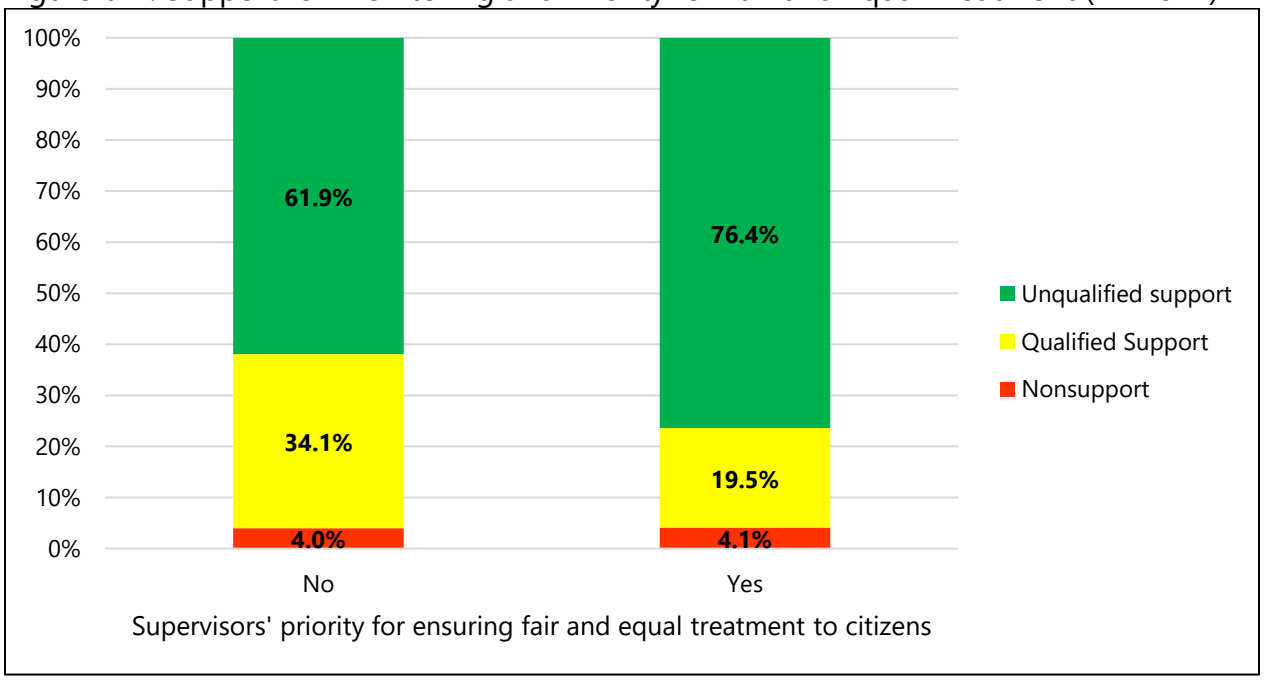
These four items are strongly correlated and form a reliable scale (alpha = .819). The supervisor monitoring index is based on 972 responses and ranges from 4 to 20, with a mean of 15.9. For the purpose of presenting findings in more economical form, we represent the values in three categories: unqualified support for monitoring; qualified support; and non-support.[208] The distribution of responses is skewed—roughly two-thirds of supervisors (64.1 percent) agree that monitoring officers' performance for bias should be a part of their supervisory role. Conversely, less than five percent (4.3 percent) reject the role, and about a third (31.7 percent) give it only qualified support.

Monitoring is merely one of many responsibilities supervisors are required to perform. In recognition of their limited time and competing responsibilities, we inquired into supervisors' priorities in patrol work. We asked supervisors to identify the three most important supervisory functions from among a list, such as enforcing department rules and regulations, and distributing the workload fairly. Of the 879 supervisors who responded, 25 percent identified *ensuring fair and equal treatment of citizens* as one of their three most important supervisory functions. As expected, supervisors who prioritize fair and equal treatment of citizens are also more likely to view monitoring for potential bias as part of their supervisory responsibilities compared to their counterparts. It appears that supervisors' priority for fair and impartial policing may color how they view their role and associated responsibilities.[209]

[208] The unqualified support category includes supervisors who agreed (strongly or somewhat) with all four items. Conversely, the non-support category includes supervisors who disagreed (strongly or somewhat) with all four items.

[209] Performing a crosstabulation of these two factors, we find that this relationship is statistically significant (X^2 = 16.63; p = .000).

Figure 6-1. Support for Monitoring and Priority for Fair and Equal Treatment (N = 847)



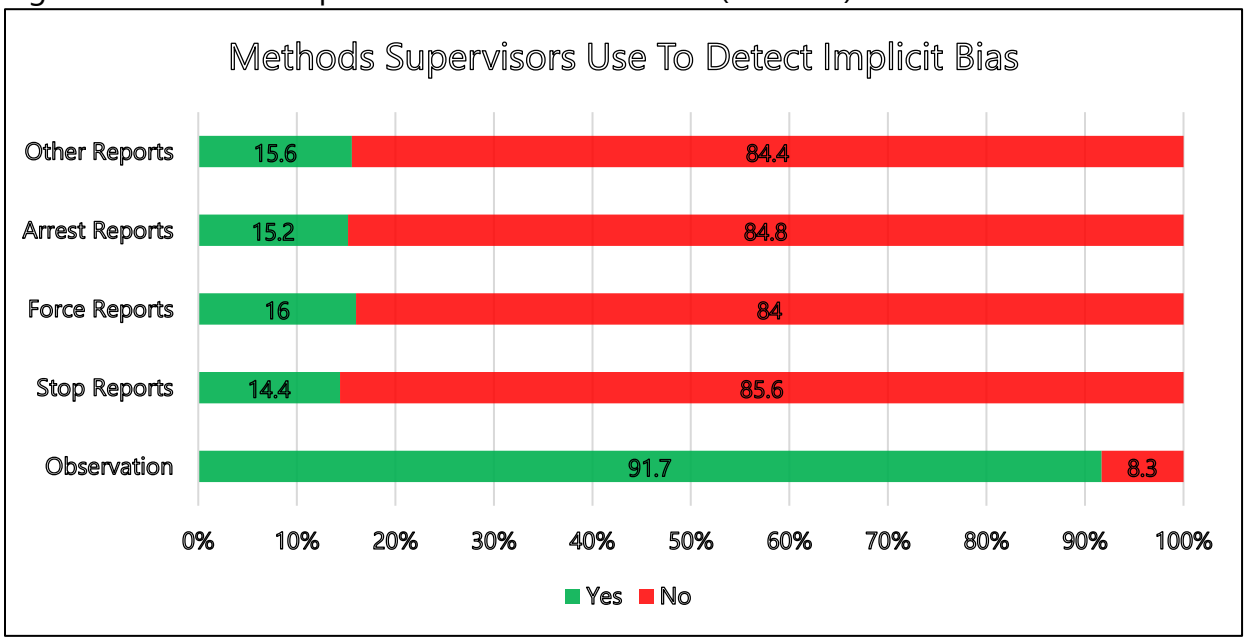
Role Enactment

The enactment of supervisors' FIP role entails activities to check for and detect potential bias and, if detected, to address it. The role also involves communicating more generally with subordinates about implicit bias and fair and impartial policing.

Detecting Bias

The FIP training tasks supervisors with identifying patrol behaviors that may signify biased policing, and points them toward several sources of information in which they may be able to detect bias in officers' street behaviors. Such monitoring may take the form of a post-hoc review of an officer's written reports or real-time observation of officers engaged in routine patrol. To assess the extent to which supervisors utilize these methods, we asked them whether they have looked for indications that their officers exhibited bias. Supervisors were instructed to select all of the sources they use to monitor officers' performance, including: (1) direct observation; (2) traffic and pedestrian stop reports; (3) use of force reports; (4) arrest reports; and (5) other reports. A majority of supervisors (91.7 percent) reportedly observe officers' street activities, while roughly 30 percent use one or more of the four report options to detect potential bias.

Figure 6-2. Sources Supervisors Use to Detect Bias (N = 494)



The use that supervisors make of these different sources of information is not strongly shaped by the extent to which they embrace their FIP role. Scores on the monitoring index are no more than weakly related to supervisors' reported reliance on these methods. Supervisors who support supervisory monitoring as one of their responsibilities, without or with qualifications, are more likely to review officers' written reports for this purpose (31.4 percent and 28.8 percent, respectively), compared to supervisors who do not embrace this role (20.0 percent). However, these differences are small and do not reach statistical significance. Nearly all supervisors reportedly perform direct observations to detect biased policing, regardless of their orientation to monitoring.

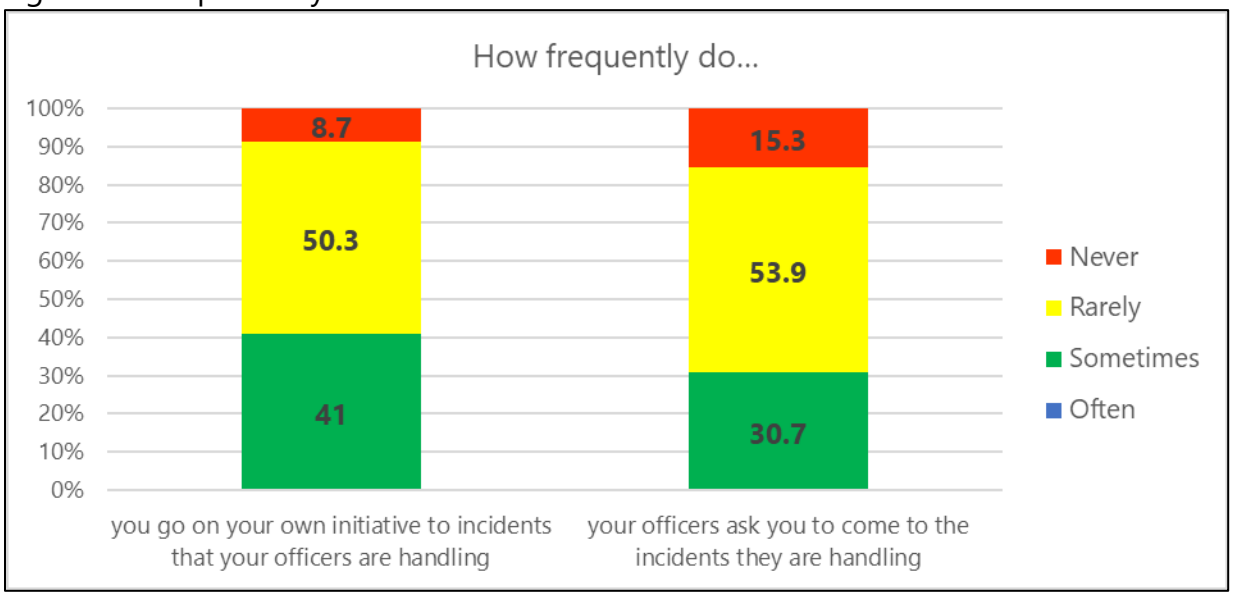
With observation as the most common source of supervisors' information, the survey asked supervisors the frequency with which:

- They go on their own initiative to incidents that subordinate officers are handling.
- Officers ask the supervisor to come to the incidents they are handling.

As demonstrated in Figure 6-3, the majority of supervisors (59.0 percent) report rarely or never going to an incident on their own initiative, and more than two-thirds (69.2 percent) report rarely or never being requested by officers to arrive on-scene. Thirty to forty percent report that they observe their officers directly under these circumstances.[210]

[210] Patrol supervisors are required to "[i]mmediately respond to and direct activities at radio runs involving any weapons (firearms, knives, etc.), serious crimes, burglaries and emergencies." (NYPD Patrol Guide, Procedure No. 202-17, Patrol Supervisor, p. 1.) In addition, NYPD policy also directs supervisors to "periodically review BWC video as appropriate, to provide positive feedback and address any performance

Figure 6-3. Supervisory Presence On-Scene



Responding to Bias

Supervisory responses to potential bias may take several forms. The FIP training instructs supervisors on how best to respond to potential bias, particularly through dialogue. Supervisors should mentor and coach officers about implicit biases by (a) communicating the importance of recognizing potential bias, and (b) guiding officers to perform patrol work that is free from bias, whether conscious or unconscious. We address communications about implicit bias later in this section.

The FIP training offers steps that supervisors may take in response to officers' potentially biased behavior. The FIP training stressed that how supervisors handle the conversation with officers is important to enhance their credibility and reputation as leaders. If and when supervisors become aware of potential bias, in an individual incident or in a broader pattern of conduct, they are advised to have a "crucial conversation" with their subordinate – crucial considering the stakes, the probability of differing perspectives, and the potential for a strong emotional reaction. The conversation begins with a statement of the facts (e.g., observations) and a (non-accusatory) explanation of how those facts could be interpreted as manifestations of bias. The officer's perspective is solicited and repeated or rephrased to ensure and

or tactical deficiencies observed." (NYPD Patrol Guide, Procedure No. 212-123, Use of Body-Worn Cameras, p. 6.) The frequency with which supervisors respond to radio runs as required by policy or utilize body-worn camera video to monitor officers' performance was not assessed in our evaluation.

establish that the supervisor understood. The conversation should end with a plan of action and next steps, if appropriate, with an arrangement to follow-up at a later date.

To assess how supervisors may respond to bias detected in written reports, we asked respondents to read a hypothetical scenario and select the one of four potential actions that best describes how they would most likely respond. The scenario reads:

> *In your review of stop reports, you form the impression that a subordinate MOS [member of the service] frequently conducts consent searches on Latino males but not people of other races or ethnicities. A check of the stop reports confirms your impression.*

The four potential responses include:

- wait to see if it happened again before taking any action
- warn the officer that there will be consequences if this happens again
- engage the officer in a dialogue to learn why this pattern is occurring
- accept that force is necessary sometimes and you can't second-guess the person in the situation

Confronted with the scenario provided, nearly 90 percent of supervisors reported that they would engage the officer in dialogue to learn why the behavior may be occurring (see Figure 6-4, below). Other respondents reported that they would accept that officers have a sixth sense and their decision-making cannot be second-guessed, warn the officer, or wait to see if the behavior occurred in the future.

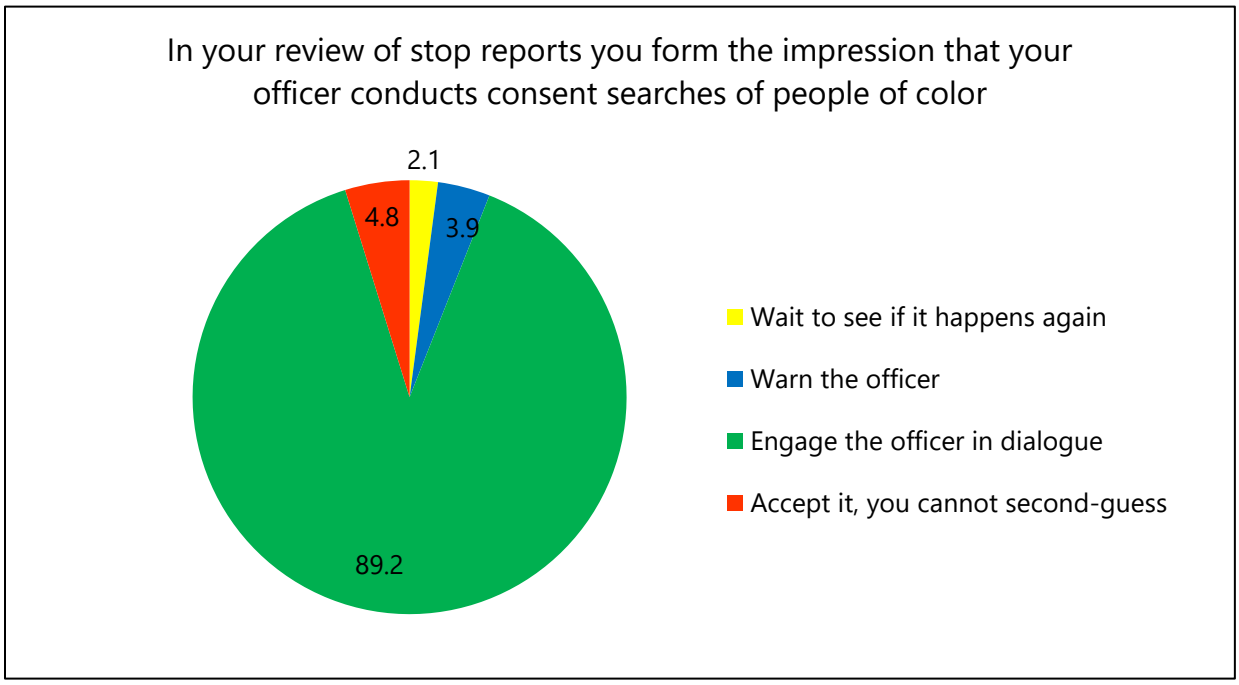
We also asked respondents if they have intervened with a subordinate officer who may have exhibited bias. Roughly one-quarter of supervisors reported that they have intervened upon detecting potential bias. These supervisors reportedly not only attended to signs of bias but also followed up.

Communicating about Implicit Bias

The FIP training discusses supervisors' role in internal communications with officers and external communications with community members. Here we concentrate on internal communications.

The FIP training casts supervisors' role as communicating the elements of fair and impartial policing and encouraging officers to be cognizant of potential biases while on patrol. Insofar as supervisors attend to this issue as an important matter with their subordinates, supervisors can reinforce the messages conveyed in the training. Supervisors must recognize their mid-level role in communications with both

Figure 6-4. Supervisors' Responses to the Hypothetical Scenario (N = 560)

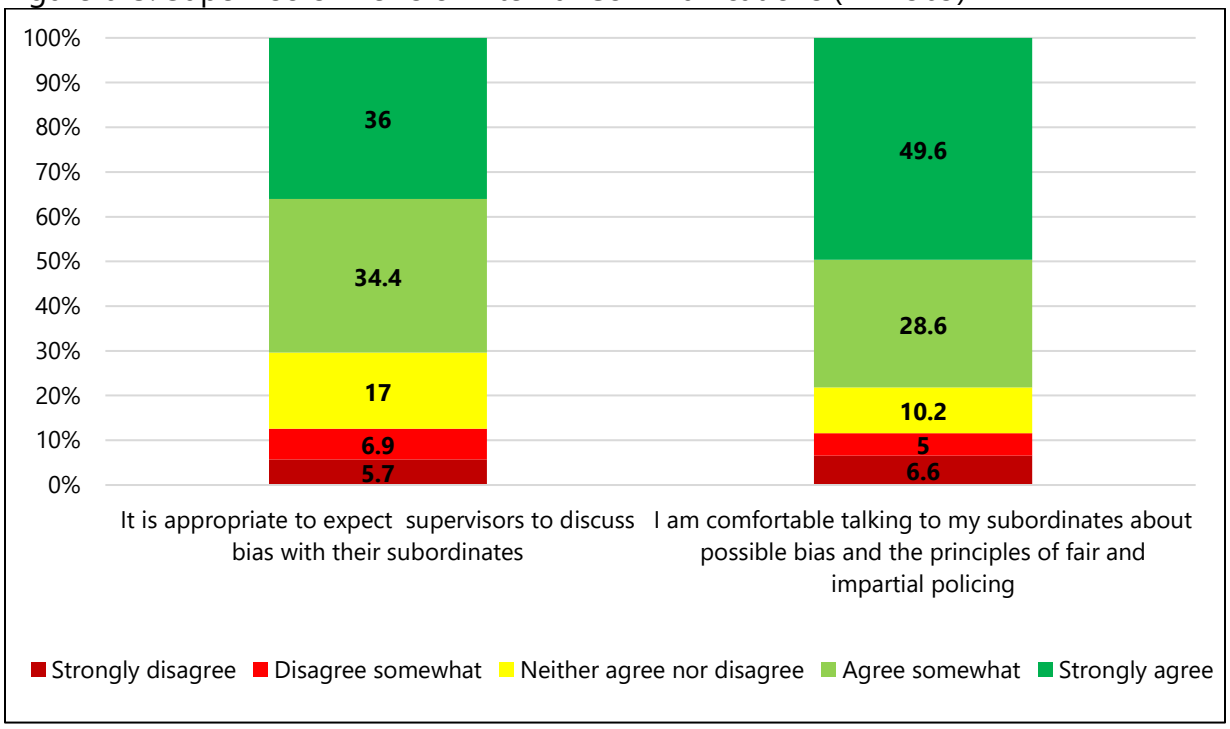


subordinates and superiors, and be comfortable engaging in discussions about sensitive topics. We asked supervisors the extent to which they agree or disagree with the following two statements:

- It is appropriate to expect supervisors to discuss bias with subordinates (part of the monitoring index, discussed above)
- I am comfortable talking to my subordinates about possible bias and the principles of fair and impartial policing

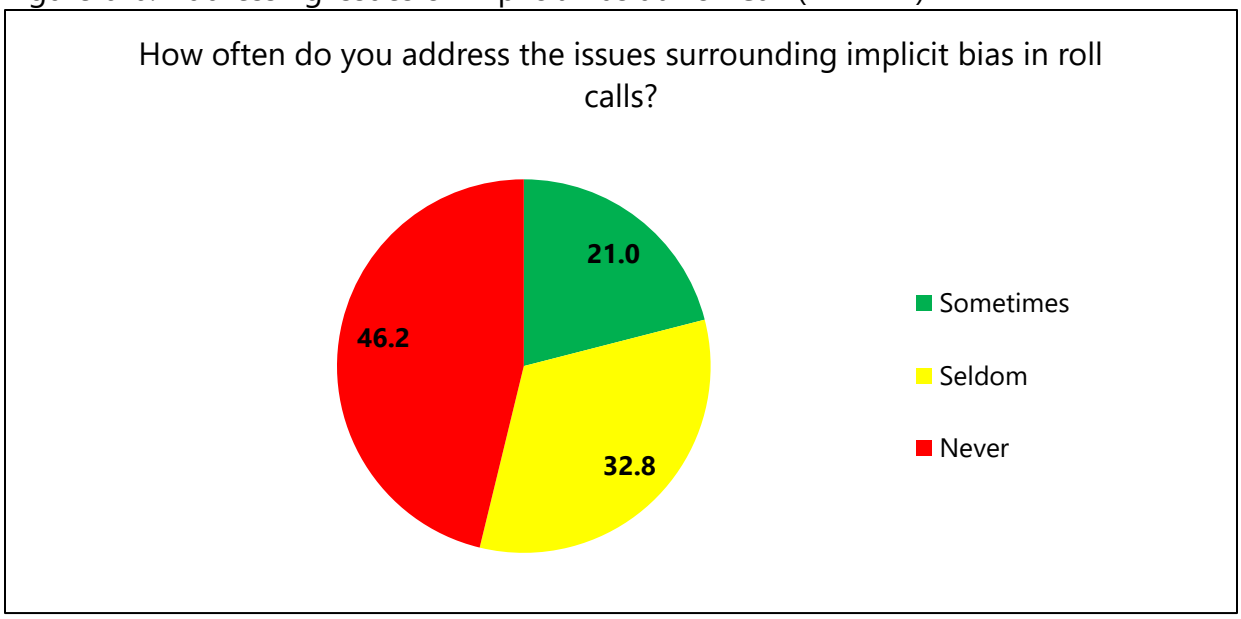
As shown in Figure 6-5, the majority of respondents agree that supervisors should be expected to discuss bias with subordinates (70.4 percent) and report that they are comfortable speaking with officers about potential bias or principles of fair and impartial policing (78.2 percent). Conversely, only one in eight reported that supervisors should not be expected to discuss issues surrounding implicit bias with officers, and a comparably small fraction reportedly feel uncomfortable talking about the subject with their subordinates. These two items are strongly correlated (r = .65), such that those who agree that supervisors should be expected to communicate with their subordinates about fair and impartial policing are more likely to feel comfortable talking about it.

Figure 6-5. Supervisors' Views of Internal Communications (N = 969)



Roll calls are one venue in which supervisors can communicate the importance of fair and impartial policing, so we asked supervisors how often they do so: sometimes, seldom or never. (Given that supervisors would be unlikely to regularly feature implicit bias and its consequences as a topic of roll calls, we did not provide for a "frequently" or "nearly always" response option.) As Figure 6-6 shows, roughly 20 percent of supervisors sometimes address implicit bias at roll call, and one-third do so seldom; nearly half reportedly never raise issues of implicit bias in roll calls.

Figure 6-6. Addressing Issues of Implicit Bias at Roll Call (n = 442)



Application of the FIP Training

The FIP training provided supervisors with knowledge about implicit bias and how they should best respond to and address bias detected in officers' street performance. We examine the extent to which they apply the FIP training to their work. We also discuss the challenges supervisors face in performing the tasks prescribed by the training.

We asked supervisors to indicate the extent to which they agree or disagree that:

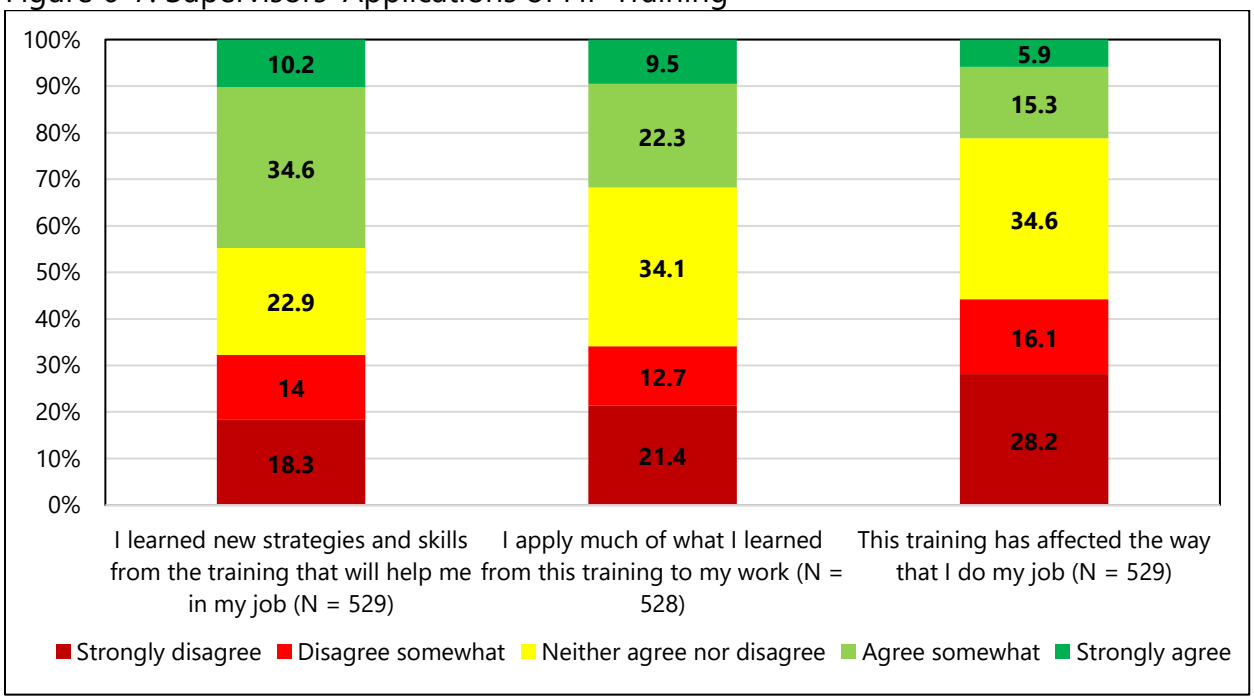
- I learned new strategies and skills from the training that will help me in my job
- I apply much of what I learned from this training to my work
- This training has affected the way I do my job

As reported in Figure 6-7, nearly half of the survey respondents reportedly learned new skills and strategies that they believe will assist them in doing their jobs, while about one-third (32.3 percent) disagreed. Approximately one-third of the supervisors indicated that they apply much of what they learned from this training to their work, and slightly more than 20 percent agreed that the training has affected how they perform their job. These three items are strongly correlated and form a reliable (alpha = .907) index of supervisors' FIP applications.

Supervisors' characterizations of their application of the FIP training is associated with their reported behavior, though not strongly. Supervisors whose application of FIP training is high are more likely to have intervened with an officer who may have exhibited bias compared to other supervisors, and they are more likely to sometimes discuss issues of bias in roll call. Supervisors whose FIP application is high are more likely to review their officers' written reports for potential bias, compared to supervisors

whose application is lower, though the differences are small and not statistically significant.

Figure 6-7. Supervisors' Applications of FIP Training



Applying any new initiative or policing strategy poses challenges. We asked supervisors to rate (on a scale from 1 to 5) the level of difficulty in using the FIP strategies, and to indicate the nature of any difficulties that they had experienced with application. As shown in Figure 6-8, roughly half of supervisors (51.7 percent) report that they have no or very little difficulty using FIP strategies, while approximately 10 percent (11.8 percent) report having a lot or some difficulty applying the principles.

We also asked supervisors to review a list of common obstacles faced during organizational change and to check all those which prevented them from using the FIP strategies. These obstacles include:

- No opportunity
- Lack of time
- Could not remember the strategies
- Did not know how to use the strategies
- Did not want to use the strategies
- Did not think the strategies would work

Figure 6-8. Supervisors' Reported Difficulty Using FIP Strategies (N = 441)

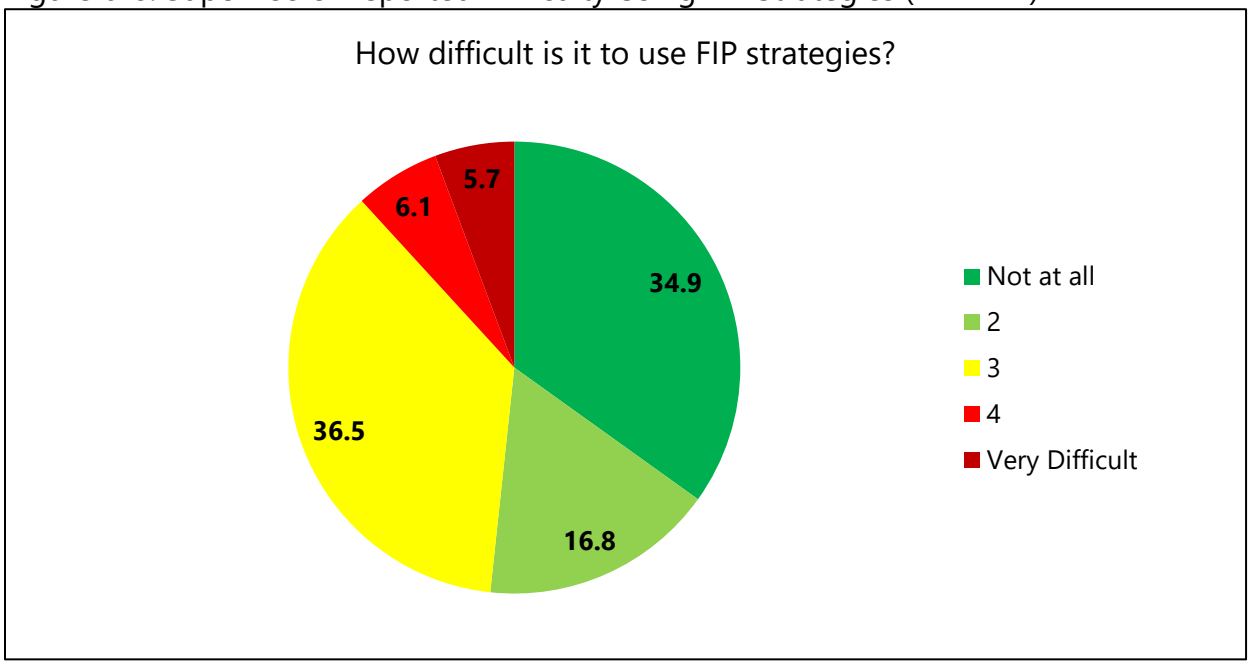
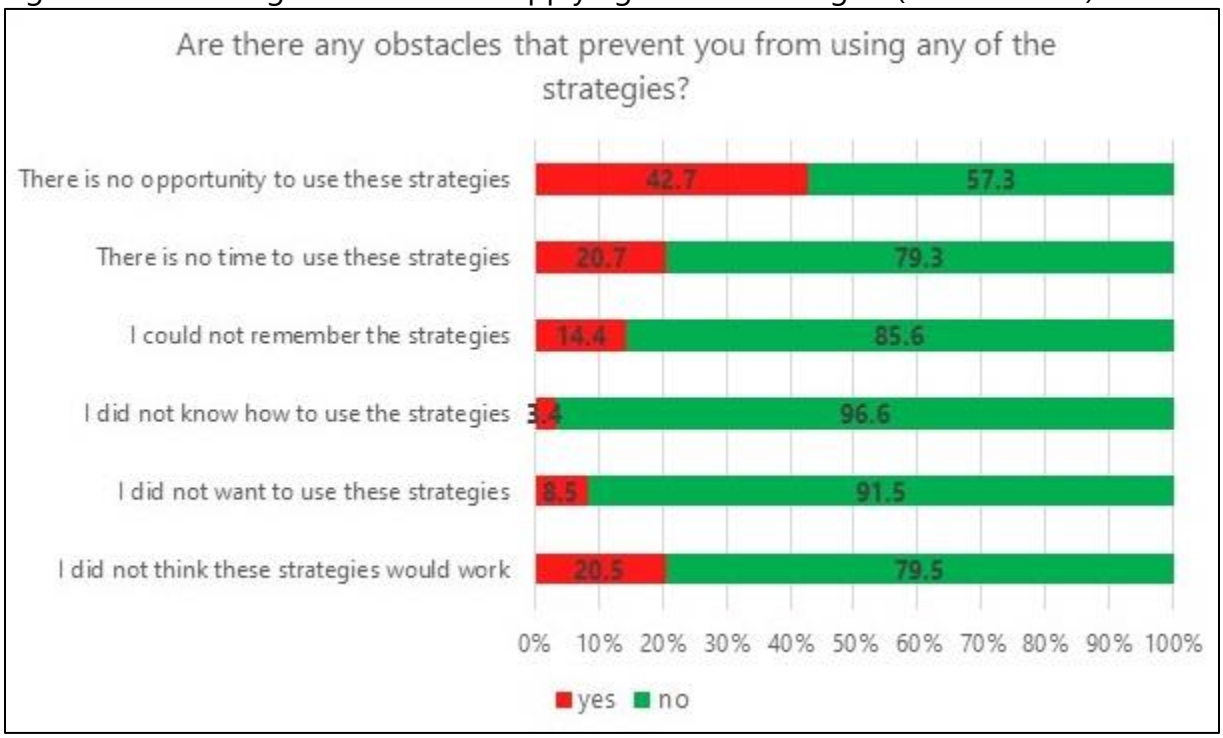


Figure 6-9. Challenges Faced When Applying the FIP Strategies (N = 449-461)



As shown in Figure 6-9, the most common difficulty experienced was having little opportunity to use the strategies. This may be a result of supervisors' assignments (e.g. to a homogeneous community population, or a specialized or administrative

assignment). Twenty percent of supervisors indicated that they did not have time to use the strategies, a characteristic common to many large and urban police departments. Very few supervisors reported that they could not remember (14.4 percent) or did not know how to use (3.4 percent) the strategies. Finally, a substantial minority of supervisors cited as obstacles either their disinclination to use the strategies or their disbelief in their efficacy: 8.5 percent indicated that they did not want to use the strategies, and 20.5 percent indicated that they did not think that the strategies would work.[211] Insofar as these supervisors serve as role models for their subordinates, and set a tone in the police workplace, they would appear unlikely to reinforce the training content.

Semi-structured interviews with sergeants provide some additional depth to the survey-based findings. We asked supervisors whether monitoring subordinates' performance for potential bias should be a routine part of a supervisor's job. Three schools of thought emerged: the majority (62 percent) of supervisors agreed it should be; the second group (21 percent) disagreed; while the remainder (17 percent) found this desirable, yet had reservations as it appeared to be an unrealistic supervisory function. Supervisors who believed that monitoring for potential biases should be a routine part of the job stated that doing so was no different than monitoring for any number of other performance measures such as productivity, time on calls, or the quality of treatment shown by officers to those with whom they interact.

Supervisors who did not agree that monitoring for implicit bias was an appropriate supervisory function provided several reasons for their belief. Some saw it as unnecessary in that biased-policing is not a problem at the NYPD. These supervisors did not reject the idea of implicit biases, rather they trusted that officers' professionalism would not allow biases to manifest in their workplace interactions. Another common reason for rejecting monitoring for biases as an appropriate supervisory function turned on the subjective nature of biases and unequal treatment to which it would lead. The subjectivity of biases opens the door to a range of perspectives on which behaviors might be indicative of biases or the threshold an officer must cross before intervention is necessary, leaving officers subject to disparate standards.

Supervisors who neither fully agreed with nor outright rejected this monitoring role offered several explanations. Most common was that it was simply unrealistic to expect supervisors to add this function to those already competing for their limited time. Several supervisors reported that distilling patterns of biased behavior from multiple interactions, each with its own context, was an unrealistic expectation and not something in which supervisors could be formally trained. Finally, some supervisors suggested that detecting bias is not something for which they could be held accountable. However, their comments suggested optimism for a somewhat more

[211] Slightly more than 5 percent of respondents (5.9 percent, N = 87) reported that they experience a problem other than one of the options mentioned.

limited and discretionary expectation. Good supervisors can and do detect biases in individual incidents. We infer that these middle ground supervisors acknowledge biases exist and can manifest themselves in the workplace, and they believe that supervisors should intervene when they come to their attention, but it is unrealistic to formalize the process.

After talking with supervisors about their general views on monitoring subordinates' performance, we then turned to their application of the FIP training. Nearly all (25 of the 29) supervisors with whom we spoke had reportedly attended the training. The discussion that follows is drawn from the responses of those twenty-five supervisors who attended the FIP training. The majority (75 percent) of supervisors indicated they had not applied what they had learned in the FIP training. Following prompts to elaborate on why this was the case, two explanations emerged: (1) there was no need to do so, or (2) the opportunity had not presented itself (e.g. supervisor is assigned to a unit where subordinates do not have street interactions). The majority of supervisors who reported no need to apply FIP training believed that biased policing is not a problem at the NYPD and, therefore, the training did not address a relevant/real issue. For example, supervisors indicated that officers already treat people "equally," "with dignity," "the same," or "as a blank canvas." Other supervisors described the training in a way that suggested they understood it as solely focused on knowledge and awareness building: "we learned we all have biases" and the most we can do is be "aware" of them. They did not seem to recall the training included a skills-based/practical application component.

A smaller proportion (25 percent) of supervisors who attended FIP training reported that they had applied some of what they learned in the training. The most common tactic we heard was supervisors taking advantage of opportunities to remind subordinates to leave their biases at home, to interact "professionally and not personally," to treat people as they would like to be treated or to look for common ground with the people with whom they interact.

We asked officers if supervisors had addressed the issue of biased policing in any way in the last six months, and their responses were consistent with self-reports from supervisors. The majority (90 percent) of officers indicated they had not had a supervisor address biased policing in any way. We would expect that even among officers whose supervisors do take steps to address biased policing, occasional reminders of the nature described to us by supervisors might not register as efforts to address biased based policing. The handful who recalled supervisors addressing biased policing provided descriptions consistent with those offered by sergeants themselves (e.g. *"We are always reminded to treat people with respect"* or "*… at roll call … Guys keep behaviors and ideas to yourself."*) Through follow up discussion with officers, we gathered that the majority of officers did not view it as particularly consequential if supervisors did not take steps to address biased policing. First, many officers did not see biased policing as an issue at

the NYPD so conclude, therefore, supervisors need not direct attention to the issue. For others, even if supervisors attempted to raise the issue of implicit bias with subordinates through the regular channels of information sharing (e.g., roll call), it would simply be lost among all the other streams of information officers take in. The few officers who indicated to us that it could be valuable for the department to take steps to address biased policing over and above the FIP training offered reservations about front-line supervisors doing so. For example, we heard concern that supervisors did not have the time or the skills to detect true biases in an officer's interactions with the public, and it might be a task better performed by an individual who is less closely tied to the officer and who has the proper skills to do so.

Conclusions

Given the central role that sergeants play in interpreting and implementing reforms at the street level, the buy-in and support of supervisors is critical to the success of any initiative. The FIP training portrays supervisors as role models. Supervisors who express their concern about bias in policing, monitor their officers' performance for signs of bias, and engage officers in dialogue upon detecting behavior that signifies potential bias, demonstrate to their subordinates the importance of fair and impartial policing and their willingness to hold officers accountable to such a standard of performance.

These survey findings indicate that most sergeants view monitoring for bias as one of their many responsibilities. Furthermore, they are willing to intervene as needed: all but a small fraction report that, in our hypothetical scenario, they would engage the officers in discussion upon detecting a pattern of potentially biased behavior; and one-quarter report that they have actually – not hypothetically – intervened with an officer whose performance warranted intervention. That most supervisors appear to rely on only observation as a source of information about potential bias, to the exclusion of stop, arrest, or use of force reports, suggests that their efforts in detecting bias could be improved, though we acknowledge that the information to which they routinely have access is quite ambiguous for this purpose. There are likely some additional opportunities to affirm a commitment to fair and impartial policing on which some supervisors are not capitalizing. Importantly, 20 percent of supervisors report that they do not believe that using the FIP strategies will be effective.

In the context of multiple reforms that implicate immediate supervisors, and the accompanying expansion of their responsibilities, we find it remarkable that the glass of supervisory reinforcement for fair and impartial policing is as full as it appears in these survey findings. Room for improvement remains, to be sure, but supervisors' receptivity to the role prescribed for them forms a fairly wide base from which those improvements can be realized.

Chapter 7
Impacts on Enforcement Disparities

The FIP training introduces several strategies by which officers may be able to mitigate the effects of their unconscious biases:

- managing biases – that is, recognizing one's biases and engaging in bias-free behavior
- avoiding profiling by proxy – that is, being aware of community members' biases and using one's own judgement;
- reducing biases – that is, seeking ways to have positive contact with individuals who are different from oneself;
- slowing it down – that is, checking one's initial impressions and collecting more information to better understand the situation; and
- engaging with community members, so that one has more positive contact with them.

Insofar as disparities in enforcement – stops, frisks, searches, arrests, summonses, or the use of force – arise at least partly from officers' implicit biases, the effective application of these strategies could be expected to diminish the magnitude of such disparities. We evaluate that hypothesis here.

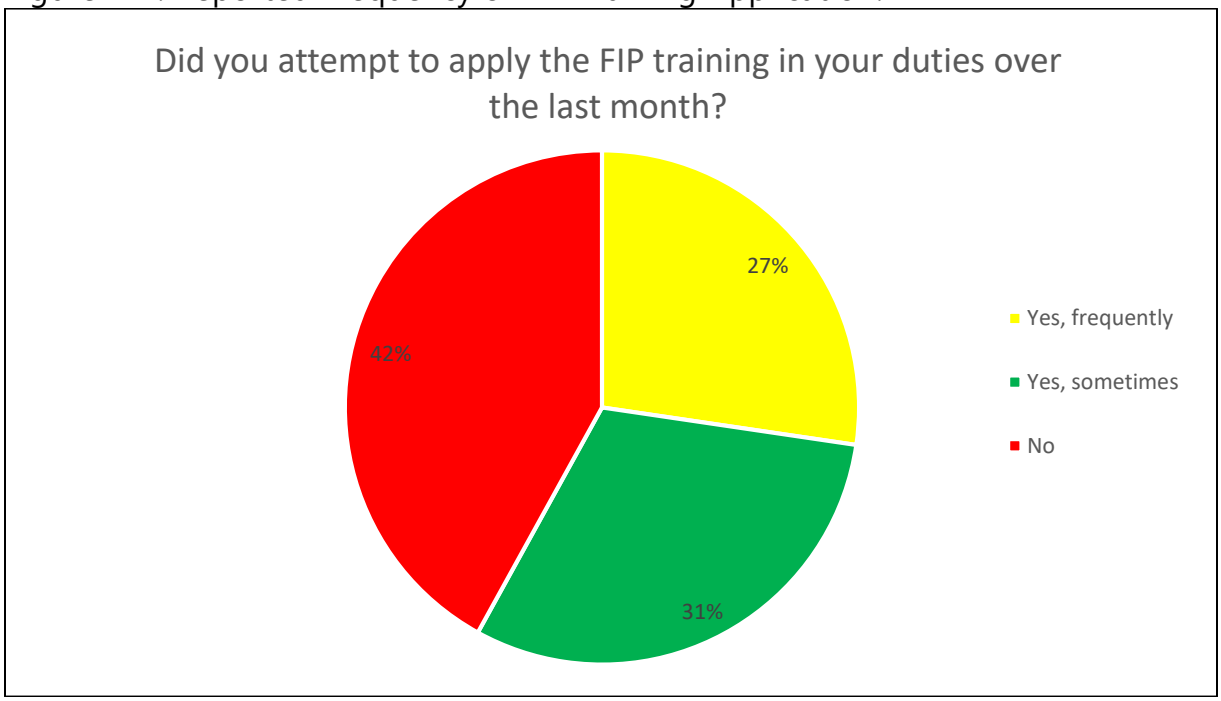
Our examination of the behavioral impacts of the training include, first, an analysis of follow-up survey data on officers' self-reported use of the FIP strategies for bias reduction, management and control, and second, analyses of enforcement behaviors pre- and post-training.

Officers' Use of FIP Strategies

Follow-up survey responses regarding officers' use of FIP skills and tactics show some variation with regard to general application of the training, and perceptions of its feasibility in the field. Of the valid answers to the question "did you attempt to apply the FIP training in your duties over the last month," 58.0 percent (N = 828 of 1,427) answered in the affirmative.[212] Of those who indicated that they used FIP strategies in the previous month, 27.3 percent said they attempted them *frequently*, while 30.7 percent stated they attempted them *sometimes*. See Figure 7-1, below.

[212] While only half of all respondents (50.2%, or N = 1568) participated in the survey beyond opening the link, of those who did, the majority (87.7%) completed the survey in full. For the purpose of describing officers' self-reported FIP strategy use in the field, only those who reached 30% or greater survey completion would have at minimum valid answers (N = 1486, or 47.6%). However, references to valid answers in this analysis only indicate answers valid to the individual question.

Figure 7-1. Reported Frequency of FIP Training Application.



In order to determine which FIP strategies were most utilized, we asked respondents to check any (or all) of the five main FIP tactics (see Table 7-1). Roughly 15 percent of officers stated that they had used strategies that revolve around community contact ("seek ways to have positive contact with individuals who are different from me" and "engage with community members so that I have more positive contact with them"). Similarly, 15.5 percent indicated that they attempted to "slow it down." Responses to "control your responses" and "avoid profiling by proxy" were lower, but still within a similar range (12 percent and 10.1 percent, respectively). A substantial proportion of survey respondents (22.3 percent) reported that they attempted to use all five strategies, but the majority (67.7 percent) stated that they attempted to use fewer than three.

Table 7-1. Reported Use of the Individual FIP Strategies.

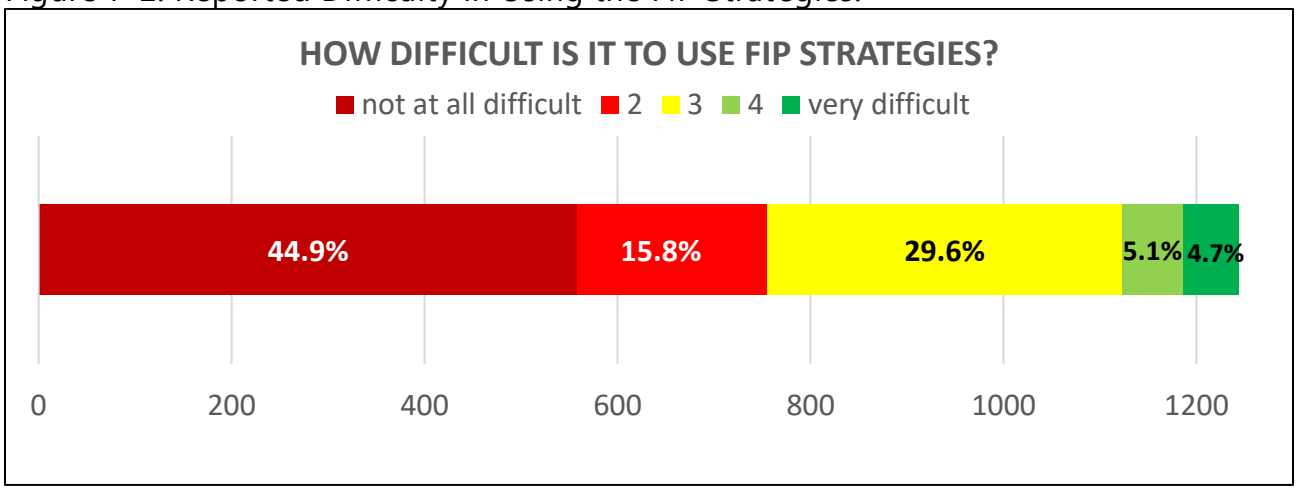
Did you…	N	Percent
Control your responses	375	12.0%
Avoid profiling by proxy	314	10.1%
Seek ways to have positive contact with individuals who are different from you	474	15.2%
Slow it down	484	15.5%
Engage with community members to have more positive contact with them	478	15.3%

Correlations between reported strategy use and attitudes and beliefs are weak but significant at the .01 level. This is represented in two ways: first, a correlation based

on responses to the question, "Did you attempt to apply the FIP training in your duties over the last month?" against respondents' scores on attitude and belief scales; and second, a correlation based on a cumulative count of strategies individual officers reportedly used against their attitudes and belief scores. These relationships follow an intuitive pattern. Officers who said that they had not attempted to use the FIP strategies tended to have lower (or less egalitarian) scores on attitude and belief scales, and those who answered "Yes, sometimes" or "Yes, frequently" were more likely to have higher (or more egalitarian) scores on the attitude and belief scales. The same pattern was observed for the measure of strategy use as a cumulative number of used strategies: officers who reported using more strategies tended to have more egalitarian scores in attitudes and beliefs.

To better understand the impediments to using the FIP strategies that officers might perceive, several instructive items were included in the follow-up survey. Of the valid responses (N = 1244) to the question "how difficult is it to use the FIP strategies," most officers (44.9 percent) stated that the FIP strategies were "not at all difficult," 29.6 percent rated difficulty as a 3 on the 5-point scale, and only 4.7 percent said the FIP strategies were "very difficult" to use. See Figure 7-2.

Figure 7-2. Reported Difficulty in Using the FIP Strategies.



In order to better illustrate the reasons that officers may not have used the FIP strategies in the field, an additional check-all item listing likely or possible causes was included (see Table 7-2). A small minority of officers stated that they did not use the strategies because they "did not want to" or "did not know how" (1.7 percent and 2.3 percent, respectively). Slightly larger proportions of officers cited "could not remember" (6.1 percent) and "did not think they would work" (4.2 percent) as reasons for not employing FIP strategies. Still larger percentages of officers (both 12.7 percent) stated that they had "no opportunity" and "no time." Most officers indicated that there was

only one reason for not employing the FIP strategies (83.1 percent), and very few listed more than three (1.5 percent).

Table 7-2. Reported Reasons for Not Using the FIP Strategies.

	N	Percent
No opportunity	361	29.0%
No time	384	30.9%
Could not remember	174	14.0%
Did not know how	65	5.2%
Did not want to	47	3.8%
Did not think they would work	123	9.9%
Other	192	15.4%

Pre- and Post-FIP Disparities in Enforcement

Disparities in enforcement by the NYPD are well-documented; whether and to what extent the disparities stemmed from police bias has been contested. For the purpose of estimating the impact of implicit bias awareness training in NYPD, we need not address the question of bias, as such, for we can test the hypothesis that the training led to diminished disparities, which is a more tractable question.

We first consider an analysis of disparities in stops between 2011 and 2015, which was conducted by two of the NYPD monitor's experts. Then we discuss trends in enforcement more generally – stops, arrests, and criminal summonses – in the 2017-2019 period for which we obtained enforcement data. We thereupon present findings about racial and ethnic disparities in the same enforcement outcomes in the 2018-2019 time frame of the stepped-wedge RCT.

Stops, 2011-2015

Stop, question, and frisk (SQF) contacts by NYPD officers have attracted particular attention for racial and ethnic disparities. John MacDonald and Anthony Braga report that in 2011, "the racial composition of individuals subjected to reported SQF encounters were 54% Black, 33% Hispanic, 9% White, and 4% Asian or other races," while the representation of the same racial and ethnic groups in New York City was 24.5 percent, 27.0 percent, 33.3 percent, and 15.2 percent, respectively.[213] The number of stops dropped dramatically during and following litigation against the NYPD, falling 96 percent between 2011 and 2015. Even so, racial and ethnic disparities remained fairly stable year by year (see Table 7-3).

[213] John MacDonald and Anthony A. Braga, "Did Post-*Floyd et al.* Reforms Reduce Racial Disparities in NYPD Stop, Question, and Frisk Practices? An Exploratory Analysis Using External and Internal Benchmarks," *Justice Quarterly* (2018).

Table 7-3. Racial/Ethnic Composition of People Stopped, 2012-2015.

Year	Stops (N=)	Black (%)	Hispanic (%)	White (%)	Asian/Other (%)
2012	532,911	53.3	30.9	9.5	6.3
2013	191,851	54.4	28.6	10.8	6.2
2014	45,787	53.1	27.2	11.9	7.6
2015	22,563	52.9	28.8	11.1	7.2

Source: MacDonald and Braga, table 1.

MacDonald and Braga analyze stops in 2012-2015, replicating (in many respects) the analysis performed on behalf of the *Floyd* plaintiffs by Jeffrey Fagan. They analyzed monthly counts of stops in New York City's precincts or Census tracts in terms of crime levels, socioeconomic factors, and the racial and ethnic composition of the precincts. Fagan asserted that this analytic approach treats crime and the racial composition of the population as the appropriate benchmarks against which to assess disparities. The estimated "place-based disparities" led to the conclusion that the NYPD's stops were discriminatory and in violation of the 14th Amendment. As MacDonald and Braga explain:

> A series of regression analyses used the resident racial distributions in NYC census tracts and NYPD precincts and found evidence that stops were significantly higher in areas with a higher fraction of Black and Hispanic residents, after controlling for arrest and crime patterns in previous periods as well as other socioeconomic factors (Fagan, 2010; Gelman, Fagan, & Kiss, 2007). These analyses suggest there were disparities in SQF activities based on places, and were highly influential in the ruling made by a federal judge in the Floyd et al. litigation that the use of SQF as practiced by the NYPD during the 2000s was unconstitutional.

Indeed, the court drew an inference of bias – discrimination – from the results of Fagan's analysis:

> The crux of plaintiffs' Fourteenth Amendment claim is that blacks and Hispanics are stopped more frequently than they would be if police officers did not discriminate based on race when deciding whom to stop. Assessing this claim required comparing statistics about rates of stops of blacks and Hispanics to [a benchmark]. … a central dispute between the experts regarding the Fourteenth Amendment claim was the appropriate benchmark for measuring racial bias in stops. … Each expert submitted voluminous reports and testified at trial in support of his choice of benchmark. … I conclude that Dr. Fagan's benchmark is the better choice.[214]

This claim should not be accepted uncritically. Crime in this model is a rough indicator of the size – but not the composition – of the population whose behavior

[214] *Floyd et al. v. City of New York* 959 F. Supp. 2d 540 (S.D.N.Y. 2013), pp. 49 & 51.

would legitimately arouse reasonable, articulable suspicion. Furthermore, analysis conducted by the Independent Monitor showed that in many Census tracts, the predicted numbers of stops were lower than the actual numbers of stops, though the actual numbers are generally under 100 per tract.[215] It also showed that in a number of tracts with 50 to 100 percent Black residents, the regression predicts very large numbers of stops – 200 to 400 or more – even though the actual numbers of stops are very small (or zero). These predictive errors should give everyone pause in drawing inferences about bias from the estimated parameters of this regression model.

Be all that as it may, MacDonald and Braga do not make claims about bias. Their point is that the estimated effects of the racial and ethnic composition of the surrounding Census tract and its socioeconomic characteristics declined over time, other things (especially crime levels) being equal, and by 2015 the effects were statistically insignificant. They conclude that the "place-based racial disparities" on which the Court's *Floyd* ruling was based "were no longer statistically different from zero in 2014 and 2015." We can infer that although racial and ethnic disparities in stops remain, they have become more closely associated with crime levels. As police have made fewer stops, stop activity has become more concentrated in places with higher crime. The decline in stops can probably be attributed to several factors: (1) regression to a more natural mean, not inflated due to administrative pressures; (2) the more challenging climate of public opinion;[216] and (3) *anticipated* restrictions imposed in court-ordered reforms, which were first applied in 2015.

Enforcement, 2017-2019

Stops continued to decline after 2015, numbering fewer than 9,000 in 2017, with small upticks in 2018 and 2019 (though these increases could be partly attributable to better reporting of stops). Furthermore, across the period for which we obtained data, split by year and by FIP training, the initiation of stops changed somewhat. Table 7-4 divides the period in two ways: the top panel splits 2018 at the point at which FIP training for patrol officers was begun (on May 14); the bottom panel splits 2018-2019 based on the dates on which the training was attended by the individual officers who conducted the stops.[217] Either way, post-FIP stops tended to be based more on suspicion of more serious offenses and less on officers' own discretion. Pre-FIP, robbery and other violent crimes were the suspected offenses in 30-31 percent of stops, inching up to 33 percent of stops post-FIP. Pre-FIP, 26 percent of the stops were self-initiated by officers, declining to 22 percent post-FIP. Roughly comparable shifts are detectable

[215] Peter L. Zimroth, *Fifth Report of the Independent Monitor: Analysis of NYPD Stops Reported, 2013-2015* (New York: Arnold & Porter Kaye Scholer LLP, 2018), pp. 29-30.
[216] Zimroth, *First Report of the Independent Monitor*, p. 8.
[217] The counts include only actions by officers in the three bureaus on whom the evaluation focuses.

Table 7-4. Stops, 2017-2019.

	2017	2018		2019*
		Pre-FIP Start	Post-FIP Start	
Number of stops	8,584	3,468	5,897	5,924
Suspected offense				
Percentage for robbery	14.55	15.43	14.62	15.06
Percentage for other violent	14.92	15.86	18.21	17.56
Percentage for weapons	26.63	25.61	26.23	26.52
Percentage for drugs	3.81	4.84	2.73	2.46
Percentage for larceny	11.40	13.06	13.38	14.96
Percentage for other property	23.61	20.56	19.06	18.43
Percentage for other crime	5.07	4.64	5.77	5.01
Initiation				
Percentage Self-initiated	25.7	26.5	22.3	22.0
Percentage Radio run	60.0	58.1	62.7	62.4
Percentage Complainant/witness at scene	14.3	15.4	15.0	15.6

	2017	2018-2019	
		Pre-training	Post-training
Number of stops	8,584	7,683	7,606
Suspected offense			
Percentage for robbery	14.55	15.33	14.61
Percentage for other violent	14.92	16.82	18.04
Percentage for weapons	26.63	24.64	27.78
Percentage for drugs	3.81	3.59	2.62
Percentage for larceny	11.40	13.99	13.84
Percentage for other property	23.61	20.29	18.01
Percentage for other crime	5.07	5.34	5.10
Initiation			
Percentage Self-initiated	25.7	22.7	23.3
Percentage Radio run	60.0	61.6	61.2
Percentage Complainant/witness at scene	14.3	15.7	15.4

when we treat the dates of individual officers' training to mark pre- and post-FIP periods. These shifts are subtle, to be sure, and these comparisons do not take account of many other factors that could influence patterns of stops. But they are consistent

with an expectation that the training is received by officers as a message that officers should engage in lower levels of enforcement.

Arrests, and particularly misdemeanor arrests, in New York City have exhibited some marked trends over time. Felony arrests peaked in 1989 and, but for brief periods of proportionally modest increases, declined steadily through 2017.[218] Misdemeanor arrests increased from 1980 to 2010, rising 282 percent. We note that the steeply increasing numbers of misdemeanor arrests is a pattern that the NYPD shared with a small number of police departments in the U.S. Cynthia Lum and Heather Vovak show that only a small fraction of the nation's larger municipal and county police agencies exhibited a similar trend in misdemeanor arrests.[219]

Between 2010 and 2017, the NYPD's misdemeanor arrests dropped 37 percent. In the time period covered by our evaluation, felony arrests have increased while arrests for misdemeanors and other lesser offenses have decreased (see Table 7-5). The increase in the proportion of arrests that were for felony offenses largely reflects a change in the enforcement of laws governing the possession of marijuana. Arrests for unlawful possession of marijuana (a violation) and criminal possession of marijuana in the 5th degree (a B misdemeanor) dropped from 800-1,200 per month in 2017 to 200-300 per month after the policy change was announced in June, and to fewer than 100 per month after the policy change became effective. Even before the announcement of the policy change, in May, these arrests dropped under 700.

Like stops and arrests, the issuance of criminal summonses was on a downward trajectory in New York City prior to the FIP training. Summonses issued by the NYPD decreased from 269,361 in 2015 and 243,714 in 2016, to 148,538 in 2017 and 77,086 in 2018. See Figure 7-3. The issuance of criminal summonses declined sharply in 2017 due to the Criminal Justice Reform Act (CJRA), which became effective on June 13, 2017.

> The CJRA changes the enforcement of some lower-level offenses in New York City by creating the presumption, absent certain exclusionary factors …, that some behaviors will result in a civil rather than a criminal summons. The behaviors that became eligible for a civil summons under CJRA include public consumption of alcohol, public urination, littering, unreasonable noise, and all NYC Parks Rules offenses.[220]

---

[218] Meredith Patten, Quinn O. Hood, Cecelia Low-Weiner, Olive Lu, Erica Bond, David Hatten, and Preeti Chauhan, *Trends in Misdemeanor Arrests in New York, 1980 to 2017* (New York: John Jay College of Criminal Justice, 2018), p. 20.

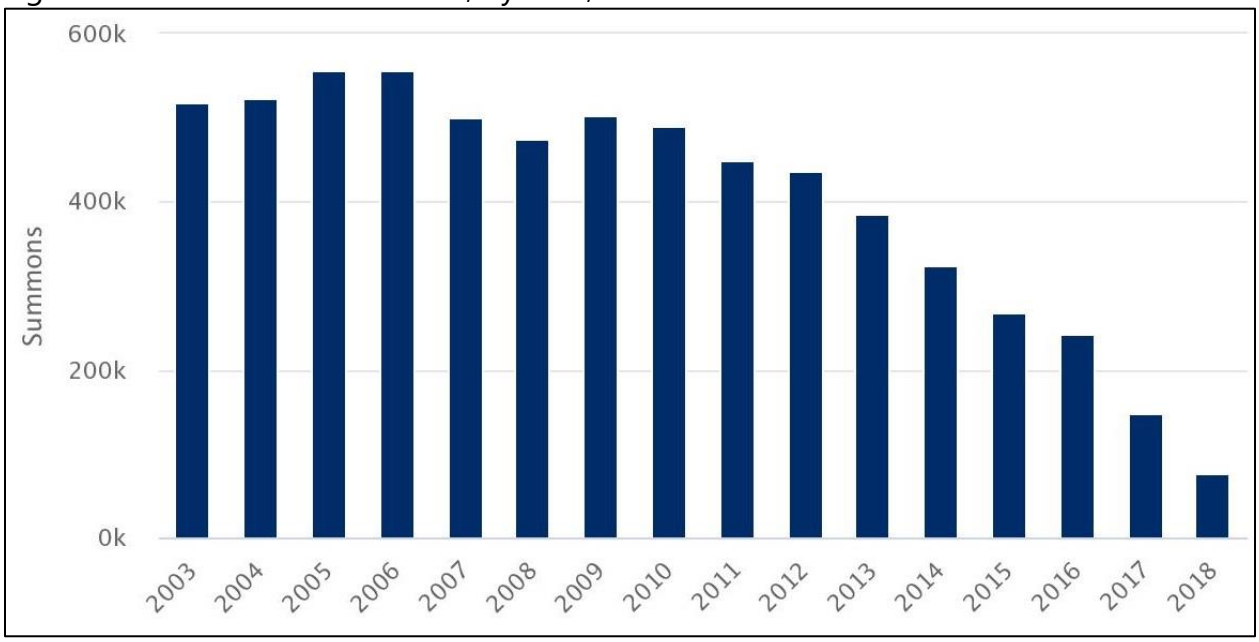
[219] Cynthia Lum and Heather Vovak, "Variability in the Use of Misdemeanor Arrests by Police Agencies from 1990 to 2013: An Application of Group-Based Trajectory Modeling," *Criminal Justice Policy Review* 29 (2018): 536-560.

[220] Kerry Mulligan, Celina Cuevas, Edwin Grimsley, and Preeti Chauhan, *The Criminal Justice Reform Act Evaluation: Post-Implementation Changes in Summons Issuance and Outcomes* (New York: John Jay College of Criminal Justice, 2018), p. 11.

Table 7-5. Arrests, 2017-2019.

	2017	2018		2019*
		Pre-FIP Start	Post-FIP Start	
Number of arrests	156,429	58,998	92,157	70,504
Percentage for felony offenses	24.66	25.83	30.40	32.34
Mean charge seriousness	3.209	3.234	3.282	3.305
	2017	2018-2019		
		Pre-training	Post-training	
Number of arrests	156,375	125,446	96,057	
Percentage for felony offenses	24.66	27.77	32.45	
Mean charge seriousness	3.209	3.255	3.304	

Figure 7-3. Criminal Summonses, by Year, 2003-2018.



Source: https://datacollaborativeforjustice.org/data-home/data-dashboard/

We find similar declines among patrol officers in the three bureaus whose training we evaluate, which also reflect the impacts of the CJRA (see Table 7-6). Prior to the implementation of the CJRA, summonses for public consumption of alcohol represented a significant fraction of the total, dropping thereafter to 10 percent or less.

With the announcement of the change in the NYPD's marijuana enforcement policy, summonses for marijuana remained stable, with seasonal and month-to-month fluctuation between 1,000 and 1,600 per month, as arrests for marijuana decreased.

Table 7-6. Criminal Summonses, 2017-2019.

	2017^	2018		2019#
		Pre-FIP Start	Post-FIP Start	
Number of summonses	32,452	28,084	39,087	27,548
Alleged offense				
Percentage for alcohol*	8.95	6.29	10.84	6.57
Percentage for marijuana**	22.98	21.87	24.67	29.88
Percentage for disorderly conduct	9.01	3.20	3.05	2.90
Percentage for urination***	1.13	1.11	1.32	1.15
Percentage for engine****	4.66	6.80	3.84	3.29
Percentage for trespass	4.29	4.03	3.82	4.64
Percentage for obstruction*****	4.46	3.69	4.25	3.51
	2017^	2018-2019		
		Pre-training	Post-training	
Number of summonses	32,452	56,165	38,554	
Alleged offense				
Percentage for alcohol*	8.95	8.12	8.45	
Percentage for marijuana**	22.98	23.00	28.79	
Percentage for disorderly conduct	9.01	3.15	2.92	
Percentage for urination***	1.13	1.14	1.31	
Percentage for engine****	4.66	5.30	3.47	
Percentage for trespass	4.29	4.19	4.03	
Percentage for obstruction*****	4.46	4.03	3.64	

^ July 1 – December 31
# January 1 – June 30
* Public consumption of alcohol
** Possession of marijuana
*** Public urination
**** Engine on/key in ignition
***** Obstruction of vehicular or pedestrian traffic

Disparities, 2017-2019

Aggregate disparities in stops remained fairly stable (see Table 7-7), before and after the FIP training for patrol officers was initiated on May 14, 2018. Somewhat more than half – 55 to 60 percent – of stops were of Blacks, and about 30 percent were of Hispanics.

Table 7-7. Racial/Ethnic Disparities in Enforcement, 2017-2019.

	2017	2018		2019*
		Pre-FIP	Post-FIP	
Number of stops	8,529	3,452	5,707	5,884
Stops - % Black	56.2	56.3	57.1	59.3
Stops - % Hispanic	31.8	31.4	31.3	29.3
Frisks in stops - % Black	57.2	58.4	60.0	60.6
Frisks in stops - % Hispanic	32.9	31.4	31.0	30.7
Searches in stops - % Black	53.7	52.9	56.7	57.8
Searches in stops - % Hispanic	33.8	35.6	31.8	30.0
Force in stops - % Black	57.7	52.2	58.9	62.6
Force in stops - % Hispanic	30.8	34.6	29.2	26.2
Number of summonses	88,827	26,681	37,956	26,948
Summonses - % Black	43.8	51.9	50.0	51.0
Summonses - % Hispanic	35.0	30.8	34.2	33.1
Number of arrests	156,429	58,998	92,157	70,504
Arrests - % Black	47.5	46.9	47.9	47.6
Arrests - % Hispanic	34.5	35.3	34.5	34.3
Force in arrests - % Black	61.3	62.4	60.5	61.2
Force in arrests - % Hispanic	29.2	28.6	30.4	27.6
Number of non-felony arrests	117,810	43,742	64,087	47,666
Non-felony arrests - % Black	46.0	45.1	45.9	45.4
Non-felony arrests - % Hispanic	35.3	36.4	35.5	35.5

* January 1 – June 30

Across these years, 57 to 60 percent of stops involved a frisk, 33 to 36 percent involved a search, and 22 to 23 percent involved police use of force. Blacks and Hispanics were represented among those who were frisked, searched, or subjected to force nearly in proportion with their representation in the stopped population – that is, given the stop, post-stop outcomes exhibit little or no disparity. This is congruent with the findings reported by the Independent Monitor, whose experts analyzed post-stop outcomes in 2013-2015.[221] By 2015, they found few statistically significant differences associated with race and ethnicity, controlling statistically for the relevant factors that

[221] Zimroth, *Fifth Report of the Independent Monitor*, pp. 33-37.

could be measured, and the differences that achieved statistical significance were quite small in magnitude.

Disparities in misdemeanor arrests have changed much less dramatically than their annual numbers. In 2010, the misdemeanor arrest rate for Blacks was 5.8 times that for Whites, and the arrest rate for Hispanics was 3.2 times that for Whites. By 2017, these "rate ratios" had dropped to 5.2 and 3.0, respectively; both Blacks and Hispanics remained overrepresented among misdemeanor arrestees.[222] Though felony arrests have increased and arrests for misdemeanors and other lesser offenses have decreased, the representation of Blacks and Hispanics in the arrestee populations has remained steady.

A small fraction (1.3 percent) of arrests involve police use of physical force (see Table 7-7, above). Blacks are overrepresented among those to whom police force is applied, while Hispanics and Whites are underrepresented.

Training Impacts

As we explained above, in describing the evaluation design, we have applied two analytic approaches to estimating the magnitude and direction of the impacts of FIP training on enforcement disparities.[223] First, we analyze aggregate enforcement behavior by officers in the training clusters of commands, pre- and post-training – that is, as the training was delivered in the steps of the stepped-wedge design. Though the FIP training is not intended to affect enforcement levels as such, if the training reduced enforcement disparities, then we would expect to see either (a) reductions in enforcement overall, driven by reductions in enforcement actions against people of color as officers apply bias-management strategies, or (b) changes of differing magnitude among different categories of race/ethnicity – greater decreases among people of color than among Whites, or smaller increases among people of color than among Whites. We present findings for each of the several forms of behavior: stops; summonses; arrests; frisks and searches in stops, and physical force in stops and arrests. We also analyze counts of citizen complaints.

Second, we analyze behavior in individual events, pre- and post-training. Analyzing stops, summonses, and arrests, respectively, we treat the race/ethnicity of the individual as the outcome, estimating the effect of the training on the likelihood that the individual is Black or Hispanic, other things being equal. Post-training decreases in either or both of these likelihoods would confirm the hypothesized reduction in enforcement disparities. Further, among stops, we analyze frisks, searches, and the use

[222] Patten et al., *Trends in Misdemeanor Arrests in New York, 1980 to 2017*, tables 29 and 30 on p. 90.

[223] Across all of these analyses, we generate numerous estimates of regression coefficients associated with the training intervention. We discuss only those that reach a conventional .05 level of significance. We also report standard errors, so that readers who prefer to apply a different standard may do so.

of force as outcomes, estimating the effects of the training on the likelihood that these actions were taken against Black and Hispanic suspects, respectively, relative to Whites, other things being equal. We apply the same analytic approach to the analysis of the use of force in arrests.

Aggregate Analyses

The primary advantage of the stepped-wedge design is that it allows for randomization in the application of an intervention when all personnel will be exposed to the intervention. Randomization turns on the timing of treatment rather than the presence or absence of treatment. In the more common cluster randomized trials, a designated number of subjects receive treatment and a matched number of subjects does not.

One potential concern with the standard stepped-wedge model on which we relied is that assessment happens over an extended period of time in which the proportion of clusters exposed to the intervention gradually increases.[224] Thus, control clusters will, on average, contribute more to observations from an earlier calendar time than later intervention clusters. It is therefore critical to adjust for time in order to reduce the bias associated with secular trends in the outcomes caused by external forces.[225]

The primary dependent variables used in this aggregate analysis are most consistent with count regression modeling. Generalized Linear Modeling (GLM) estimation via Maximum Likelihood was used as a supplemental analysis to examine the impact of the implicit bias training across the intervention clusters.[226] We therefore used Poisson regression estimation.[227] Although the mean and variance are substantially different for many of the outcomes examined, including formal tests of overdispersion, research indicates that relying on Negative Binomial estimation as a default for overdispersed outcomes has the potential to lead to biased estimation.[228] Perhaps most importantly, robustness checks using a conditional fixed-effects negative binomial model return results that are empirically similar (particularly for point estimates and

[224] Hussey and Hughes, "Design and Analysis of Stepped Wedge Cluster Randomized Trials."
[225] Karla Hemming, Monica Taljaard, and Andrew Forbes, "Analysis of Cluster Randomised Stepped Wedge Trials with Repeated Cross-Sectional Samples," *Trials* 18 (2017).
[226] See J. Scott Long, *Regression Models for Categorical and Limited Dependent Variables* (Thousand Oaks, CA: Sage Publications, 1997); and J. Scott Long and Jeremy Freese, *Regression Models for Categorical Dependent Variables using Stata* (College Station, TX: Stata Press, 2001).
[227] Dwayne W. Osgood. "Poisson-Based Regression Analysis of Aggregate Crime Rates," *Journal of Quantitative Criminology* 16 (2000).
[228] Richard Berk and John M. MacDonald, "Overdispersion and Poisson Regression," *Journal of Quantitative Criminology* 24 (2008).

significance thresholds) to those obtained via Poisson regression, which suggests that overdispersion is not a significant issue in our analyses.[229]

The time frame we used to examine changes in counts of enforcement actions was from April 14, 2018 to May 16, 2019. This equated to one month of observational data prior to the onset of training, the eleven months of the entire training regimen, and one month post-training for the final cluster (or training block) in the design. The first step of training onset occurred May 14, 2018 – June 12, 2018, with Treatment block A; the last training step ended April 16, 2019 for Control block E. The primary reason for the relatively short pre- and post-intervention periods for the first and last training blocks was the large variations in arrests and summonses from January 1, 2017 through May 1, 2018 (i.e., in the 16 months prior to the start of training). This variation in behavioral outcomes was likely due to a series of reforms and changes in policies and practices implemented by the NYPD during that time period.[230] Thus, the analyses of event counts focus on the immediate and short-term association with implicit bias training.

We estimated for each behavioral outcome three models represented by the following three regression equations, respectively:

1) $Y^{J}_{it} = \beta 0 + \beta 1 T_{it} + \theta_i + \rho_t + \varepsilon_{it}$
2) $Y^{J}_{it} = \beta 0 + \beta 1 T_{it} + \beta 2 P_{it} + \theta_i + \rho_t + \varepsilon_{it}$
3) $Y^{J}_{it} = \beta 0 + \beta 1 T_{it} + \beta 2 P_{it} + \beta 3 T_{it} x P_{it} + \theta_i + \rho_t + \varepsilon_{it}$

In each equation, Y^{J} it represents the number of behavioral outcomes of type J (e.g., stops, arrests, uses of force, etc.) generated by police in commands in each cluster i in time period t. For equations 1-3, T_{it} represents the contemporaneous timing of the permanent movement into the treatment group (i.e., implicit bias training) for commands assigned to cluster i in time period t, and where θi and ρt represent individual and time period fixed effects, respectively, that account for time- and individual-invariant unobserved heterogeneity, and ε is based on Huber-White Robust sandwich estimators to ensure the coefficient variances were robust to violations of homoscedastic error distributions. In Equation 2, P represents the contemporaneous timing of the NYPD marijuana enforcement policy change in September 2018.[231] It is

[229] Jerry Hausman, Bronwyn H. Hall, and Zvi Griliches, "Econometric Models for Count Data with an Application to the Patients-R&D Relationship," *Econometrica* 52 (1984).

[230] M Sisitzky, "Police Reform is Coming to New York City, but will the NYPD Follow the Law?" (2018). American Civil Liberties Union: https://www.aclu.org/blog/criminal-law-reform/reforming-police-practices/police-reform-coming-new-york-city-will-nypd.

[231] As noted by McGowan, the NYPD consistently reduced the number of marijuana arrests over the past decade from 53,000 marijuana arrests in 2010 to 19,000 in 2017. Beginning in June 2018 and fully enacted by September 2018, the NYPD relaxed marijuana enforcement, authorized to issue summonses rather than make custodial arrests for marijuana violations. Furthermore, McGowan reported that 9 out of 10

also important to note that alternative policy indicator variables were included to estimate the effect of the implementation of Neighborhood Policing in specific precincts over the period of inquiry here, and none yielded any substantive impacts on event counts examined herein.

Finally, in Equation 3, $T_{it}xP_{it}$ represents a fixed-effect interaction term between the contemporaneous timing of the treatment crossover during the period following the NYPD marijuana enforcement policy change.[232] Since marijuana arrests disproportionately involved Blacks and Hispanics, the altered policy would at the margin restrict the occasions on which the application of FIP strategies would reduce disparity in arrests, and – if marijuana enforcement overall remained stable over time – correspondingly expand the occasions on which the application of FIP strategies would reduce disparity in summonses. In Equation 3, we have conditional (rather than unconditional) parameter estimates of treatment cross over, before (β1) and after (β1 + β3) the marijuana policy change, respectively.

For each outcome, we estimate the impact of implicit bias training on the counts of all such events and those for each race/ethnicity separately. To reiterate: the hypothesized reduction in disparity would be confirmed by either decreases among Blacks and/or Hispanics that are greater in magnitude – to a statistically significant degree – than any decrease among Whites, or increases among Blacks and/or Hispanics that are smaller than an increase among Whites. Table 7-8 summarizes the results; Tables B-1 through B-11 in Appendix B display the details (coefficients and standard errors).[233]

---

marijuana arrests typically involved a non-White (i.e., Black and/or Hispanic) suspect. See https://www.ny1.com/nyc/all-boroughs/news/2018/09/01/nypd-marijuana-enforcement-relaxation-policy-takes-effect-most-marijuana-smokers-to-get-summonses

[232] We specified and estimated this model at the suggestion of the NYPD Monitor's experts. They explained: "To control for the marijuana policy change they use a dichotomous variable to account for the change that occurred in September 2018. This variable will be perfectly correlated with the strata-time fixed effect control variables that they use to control for time and cluster specific effects. This will then mean that the estimates of the FIP post training are only for the months before September 2018, greatly reducing the effective sample size and negating nearly half of the treatment commands. To assess the impact of FIP in the post-marijuana policy change, Worden and colleagues could simply add an interaction term to the time-strata fixed effect they use. This would in effect see if the main effect of FIP exists once one accounts for specific clusters (strata – high and low arrest locations) being treated in the post-marijuana policy change." "Memo Regarding the Evaluation of the Fair and Impartial Policing Training Program," May 13, 2020. However, the rationale for this specification is not only or primarily – if at all – statistical, as we explain in the text.

[233] Readers with an eye for detail may note that in many instances, the estimates of β1 and its standard error are identical in Equations 1 and 2. We assure you that these are the results we obtained and not an error.

Stops

The count regression results provide no evidence of a change in the total number of stops that corresponded with the timing of the implicit bias training. There is no statistically significant parameter in any of the equations (1-3) for any of the equations (i.e., unconditional impact, impact controlling for the marijuana policy change, and impact conditional on the marijuana policy change) overall or by suspect race/ethnicity. Nor did we find significant differences between coefficients for the different race/ethnicity categories, indicating no change in the aggregate racial disparities in stops.

Summonses

The number of summonses did not exhibit any significant change in the post-training period, overall or for summonses issued to only Black suspects, only Hispanic suspects, Black and Hispanic suspects, or only White suspects.

Results from Equation 2 show that the marijuana policy change corresponded with a 32.4 percent *decline* in the monthly number of summonses. The reduction was statistically significant ($p < .01$). Given the nature of the policy change, one might expect an *increase* in summonses, more or less corresponding to a decrease in arrests for marijuana offenses. Controlling for this effect, however, did not alter the coefficient for the intervention cross over. Finally, Equation 3 shows a statistically significant post-training increase in the period that preceded the marijuana policy change, and a significantly smaller (near zero) post-training change in the period following the policy change. The contingent effect prior to the policy change is not readily interpretable in terms of the training; that is, we know of no theoretical reason to posit that the training would increase the issuance of summonses under conditions that prevailed prior to the training, and have no effect on the issuance of summonses under the conditions of the new marijuana enforcement policy. A very similar pattern of estimated changes is seen across each of the groups defined by suspect race/ethnicity, and we find no differences across categories of suspect race/ethnicity that would signal changes in disparities.

Arrests

The number of arrests did not exhibit any significant change in the post-training period. Similar to the pattern in summonses, Equation 2 shows that the marijuana policy change corresponded with a decline in the monthly number of arrests by roughly 10-13 percent, depending on the race/ethnicity of arrestees. These reductions were statistically significant ($p < .01$) for all racial and ethnic groups. Again, however, controlling for the policy change did not alter the estimated change with the intervention cross over. The results for Equation 3 indicate that post-training arrest counts were higher only prior to the marijuana policy change, paralleling the pattern for summonses. As with summonses, this pattern is not interpretable as a training effect. Disaggregating arrests

by race/ethnicity also reveals a pattern like that found among summonses: a statistically significant post-training change (increase) only in the period that preceded the marijuana policy change (excepting arrests of Whites, counts of which neither increased nor decreased post-training before or after the marijuana policy change). In summary, these results suggest that there was no statistically significant change in the number of or disparities in arrests that is attributable to the implicit bias training.

Table 7-8. Summary of Statistically Significant Coefficients, Cluster-Level Counts of Enforcement Actions, by Suspect Race/Ethnicity

	Suspect Race/ Ethnicity	Eq 1	Eq 2	Eq 3 - Pre	Eq 3 - Post
Stops	All	-	-	-	-
	White	-	-	-	-
	Black	-	-	-	-
	Hispanic	-	-	-	-
	Persons of color	-	-	-	-
Summonses	All	-	-	1.17	-
	White	-	-	-	-
	Black	-	-	1.14	-
	Hispanic	-	-	1.23	-
	Persons of color	-	-	1.14	-
Arrests	All	-	-	1.06	-
	White	-	-	-	-
	Black	-	-	1.07	-
	Hispanic	-	-	1.09	-
	Persons of color	-	-	1.08	-
Frisks	All	-	-	-	-
	White	-	-	-	-
	Black	-	-	-	-
	Hispanic	-	-	-	-
	Persons of color	-	-	-	-

	Suspect Race/ Ethnicity	Eq 1	Eq 2	Eq 3 - Pre	Eq 3 - Post
Searches	All	-	-	-	-
	White	-	-	-	-
	Black	-	-	-	-
	Hispanic	-	-	-	-
	Persons of color	-	-	-	-
Use of Force in Stops	All	-	-	1.33	.
	White	-	-	-	-
	Black	1.39	1.39	1.57	-
	Hispanic	-	-	-	-
	Persons of color	1.32	1.32	1.52	1.13
Use of Force in Arrests	All	-	-	-	-
	White	-	-	-	-
	Black	-	-	-	-
	Hispanic	-	-	1.13	-
	Persons of color	-	-	-	-
Complaints		-	-	-	-

Note: Entries for statistically significant coefficients expressed as odds ratios

Frisks and Searches in Stops

Analyses of cluster-level counts of frisks and searches produced results that largely mirrored those of counts of stops, in that there were no statistically significant training parameters that emerged in any of the models.

Use of Force in Stops

Analyses of the use of force in stops, like other post-stop outcomes, revealed no estimate of a training effect that achieved statistical significance, but with one exception. Post-training, counts of use of force against Black suspects increased, increasing disparity in use of force, contrary to the hypothesized training effect.

Use of Force in Arrests

The findings concerning the use of force in arrests mirror the results for arrests, which is to be expected, since the enumerated uses of force represent a subset of arrests. The total number of arrests in which force was used did not exhibit a statistically significant change in the post-training period, excepting use of force against Hispanics prior to the marijuana policy change (and only prior to the policy change).

Citizen Complaints

The analysis of counts of citizen complaints indicates that there were no statistically significant differences associated with training-as-delivered across the command clusters of the experimental design.

Sensitivity Analyses

The primary source of complexity in the analysis of the impact of implicit bias training on police officer behavior is the policy change in the enforcement of marijuana laws that was intended to (and apparently did) reduce the number of low-level arrests (including drug and other misdemeanor arrests) for non-White citizens. Once the timing of this marijuana policy change is included in the behavioral count analyses, there is hardly any evidence of statistically significant changes attributable to implicit bias training. The results consistently indicate that the implicit bias training had no beneficial impact on racial/ethnic disparities in enforcement.

It was also important to examine the types of arrests that changed in total and across racial/ethnic groups over time. A series of supplemental analyses (available upon request) were conducted to provide additional context for the results presented above. For Blacks and Hispanics, the driving force behind the arrest declines was a significant reduction in misdemeanor arrests. Non-White felony arrests remained stable net of implicit bias training and the marijuana policy change. However, no reduction in misdemeanor arrests was observed for White suspects, accounting for both implicit bias training and the marijuana policy changes. Felony arrests were significantly lower for White suspects over the duration of the study period here, which likely correlated with the estimated declines observed in White arrests over time (and post-training). The change in uses of force across racial groups also calibrated identically with the total arrest patterns.

In sum, our tests of the hypothesized impacts of implicit bias training on aggregate enforcement behaviors across the various steps in the RCT wedge design do not yield evidence that supports the rejection of the null hypothesis. Our findings indicate that the change in marijuana enforcement policy affected enforcement patterns, though not all of those effects were in the anticipated directions. When the policy change is treated as a watershed that defines distinct enforcement environments, the evidence suggests post-training increases confined to the earlier period, which are not interpretable as training effects, and with no statistically significant effect on racial/ethnic disparities.

Event-Level Analyses

The event-level analyses all examine outcomes through logistic regression. The models are analogous to that estimated for aggregate analyses:

4) $Y^J_{it} = \beta 0 + \beta 1 T_{it} + \beta 2 P_{it} + \theta_i + \rho_t + \varepsilon_{it}$

5) $Y^J_{it} = \beta 0 + \beta 1 T_{it} + \beta 2 P_{it} + \beta 3 T_{it} x R_{it} + \beta 4 T_{it} x E_{it} + \theta_i + \rho_t + \varepsilon_{it}$

Y^J it represents the behavioral outcomes of type J by officers in commands in each cluster i in time period t. T_{it} represents the contemporaneous timing of the permanent movement into the treatment group (i.e., implicit bias training) for officers assigned to commands in cluster i in time period t, and where θi and ρt represent individual cluster and time period fixed effects, respectively.

For the "veil of training" analyses of stops, arrests, and summonses, we estimate the parameters of Equation 4 in a multinomial logistic regression of the trichotomous race/ethnicity outcome to estimate the likelihood that the subject of the event is Black or Hispanic, respectively, relative to Whites as the reference category. The hypothesized training effect would be supported by negative and statistically significant estimates of β1.[234] This is arguably a more intuitive test of the anticipated reduction in disparities than that offered by aggregate analyses. In this model, P represents a set of factors for which we statistically control: place (precinct location); time (day of week and time of day); the officer's bureau (patrol, transit, housing); the change in the NYPD's marijuana enforcement policy; Raise the Age legislation; the implementation of Neighborhood Policing in selected precincts; and the reason for the stop (for stops), the nature of the alleged violation (for summonses), or the level of the alleged offense (for arrests).

For the analyses of frisks and searches in all stops, and the use of physical force in all arrests or stops, we estimate the parameters of Equation 5, in which R and E represent the suspect's race/ethnicity: Black and Hispanic, respectively. In this model, β1 is the likelihood that the enforcement action was taken against Whites, post-training relative to pre-training. β3 and β4 are the estimated differences in the effects of the training on the likelihood of enforcement actions against Blacks and Hispanics, respectively, relative to the effect among White suspects. In this model, P also includes forms of resistance in analyses of the use of force in arrests.[235]

Stops

Our aggregate analysis detected no statistically significant changes in the counts of stops following the training, overall, with or without a statistical control for the marijuana enforcement policy change. Our event-level analysis of individual stops reveals no post-training changes that are statistically significant at the .05 level (see Table 7-9, below). Thus, neither approach to the analysis of training impacts –

[234] Alternatively, we represented the training intervention as a binary pre-/post-training based on the officer's date of training relative to the date of the event. The results of these analyses do not differ from those reported below.

[235] Stop data do not include information on suspect resistance.

aggregate, training cluster-level and event-level – allows us to reject the null hypothesis of no training effect, applying a conventional .05 level of confidence, though we revisit this conclusion below.

Table 7-9. Veil of Training Analyses of Stop, Summonses, and Arrests.

	Coefficient	St. Error	Exp(B)	% Change
Stops				
Black subject stopped				
Intervention Crossover	-0.136	0.133	0.87	-12.72
Intercept	1.147*	0.581		
Hispanic subject stopped				
Intervention Crossover	-0.279	0.148	0.76	-24.35
Intercept	-0.784	0.675		
Arrests				
Black subject arrested				
Intervention Crossover	0.021	0.032	1.02	2.12
Intercept	0.997*	0.129		
Hispanic subject arrested				
Intervention Crossover	-0.034	0.034	0.97	-3.34
Intercept	-0.015	0.142		
Summonses				
Black subject cited				
Intervention Crossover	-0.085	0.055	0.92	-8.15
Intercept	0.450	0.236		
Hispanic subject cited				
Intervention Crossover	-0.121	0.062	0.89	-11.4
Intercept	-0.988*	0.255		

*p < .05; **p<.01

Summonses

The cluster-level analysis reflected the shifting trends in the issuance of summonses that accompanied the announcement and effective date of the marijuana enforcement policy change, with no changes that could be attributed to the training. The event-level analysis of summonses, like the analysis of stops, indicates no post-training changes that are statistically significant (see Table 7-9). Once again, then, neither the aggregate nor the event-level analyses allows us to reject the null hypothesis of no training effect, applying a conventional .05 level of confidence, though we revisit this below.

Arrests

At the event level, the training had no detectable effect on the likelihoods that the arrestee was Black or Hispanic, respectively, rather than White. We note that these findings are not sensitive to the exclusion of subsets of arrests:

- felonies, to better allow for the detection of training effects on lower-level offenses, about which officers have greater discretion;
- arrests of 16-year-olds, to eliminate any potentially confounding effects of legislative changes to the age of criminal responsibility in New York State; and
- arrests on marijuana charges, to eliminate any potentially confounding effects of policy changes with respect to marijuana.

These event-level findings are consistent with those at the aggregate level.

Frisks and Searches in Stops

We analyzed frisks and searches as three binary outcomes: frisks (whether or not searches were also conducted in the same stops); frisks only (i.e., frisks conducted in stops in which searches were not conducted); and searches. In none of these analyses do we find evidence of a training effect on disparities. The incidence of frisks and searches of White suspects (the reference category) were unaffected by the training, and the coefficients for the interaction terms indicate that the training had no greater effects on frisks or searches of Black or Hispanic suspects (see the top three panels in Table 7-10).

Physical Force in Stops

The use of physical force in stops increased post-training. The increase was not significantly different for Black and Hispanic suspects than for White suspects, however, such that racial/ethnic disparities in the use of force in stops was unchanged by the training.

Physical Force in Arrests

At the event level, the likelihood that physical force was used was unaffected by the training for Whites, Blacks, or Hispanics. Thus, disparity was unchanged following the training.

Officer Race/Ethnicity

The estimated impacts of the implicit bias training on officers' beliefs and attitudes were not, with few exceptions, different for officers of different races or ethnicities. However, Black officers had higher levels of knowledge about implicit bias than officers of other races and ethnicities before and after the training. Furthermore, Black officers exhibited higher levels of concern and lower levels of skepticism about

discrimination as a social problem, and higher levels of motivation to respond without prejudice. We might, then, expect to find more pronounced behavioral effects of the training on non-Black officers. Thus, we conducted separate analyses of enforcement behavior. At the aggregate level, we conducted separate analyses for White officers, Black officers, Hispanic officers, and officers of other races/ethnicities. At the event level, we conducted separate analyses, first excluding Black officers, and then excluding Black and Hispanic officers.

Table 7-10. Training Effects on Racial Disparities in Frisks, Searches, Use of Force

	Coefficient	St. Error	Exp(B)	% Change
Frisks				
Intervention Crossover	-0.144	0.160	0.87	-13.41
X Subject Black	0.179	0.156	1.04	3.56
X Subject Hispanic	0.240	0.176	1.10	10.08
Intercept	1.994*	0.169		
Only Frisks				
Intervention Crossover	-0.142	0.170	0.87	-13.24
X Subject Black	0.195	0.165	1.05	5.44
X Subject Hispanic	0.224	0.189	1.09	8.55
Intercept	0.353	0.171*		
Searches				
Intervention Crossover	-0.115	0.153	0.89	-10.86
X Subject Black	-0.030	0.148	0.87	-13.50
X Subject Hispanic	-0.109	0.171	0.80	-20.07
Intercept	-0.813*	0.154		
Use of Force – Stops				
Intervention Crossover	0.460	0.263	1.58	58.41
X Subject Black	-0.246	0.253	1.24	23.86
X Subject Hispanic	-0.333	0.298	1.14	13.54
Intercept	-2.282*	0.262		
Use of Force – Arrests				
Intervention Crossover	-0.062	0.168	0.94	-6.01
X Subject Black	-0.067	0.163	0.88	-12.10
X Subject Hispanic	-0.098	0.183	0.85	-14.79
Intercept	-5.61*	0.294		

*p < .05; **p<.01

We first present findings from aggregate analyses of the training clusters. A summary table of estimated effects that achieved statistical significance is included below, in Table 7-11; tables with the details of the results are included in Appendix B (Tables B-12 to B-15). The results can be succinctly summarized. First, we find no evidence of a training impact on aggregate stops by officers of any race/ethnicity.

Second, the baseline intervention models indicated that summons counts were stable pre-/post-training across officer races/ethnicities. When the marijuana policy change was taken into account, one difference emerged: post-training, Hispanic officers issued more summonses (see Table B-13).

Table 7-11. Summary of Statistically Significant Coefficients, Cluster-Level Counts of Enforcement Actions, by Officer Race/Ethnicity.

	Officer Race/ Ethnicity	Eq 1	Eq 2	Eq 3 - Pre	Eq 3 - Post
Stops	All	-	-	-	-
	White	-	-	-	-
	Black	-	-	-	-
	Hispanic	-	-	-	-
	Other	-	-	-	-
Summonses	All	-	-	1.17	-
	White	-	-	1.14	-
	Black	-	-	-	-
	Hispanic	1.14	1.14	-	-
	Other	-	-	-	-
Arrests	All	-	-	1.06	-
	White	-	-	-	-
	Black	-	-	-	-
	Hispanic	-	-	1.09	-
	Other	-	-	-	-
Frisks	All	-	-	-	-
Searches	All	-	-	-	0.82
Use of Force in Stops	All	-	-	1.33	-

	Officer Race/ Ethnicity	Eq 1	Eq 2	Eq 3 - Pre	Eq 3 - Post
Use of Force in Arrests	All	-	-	-	-
	White	-	-	-	-
	Black	-	-	-	-
	Hispanic	-	-	-	-
	Other	-	-	-	-

Note: Entries for statistically significant coefficients expressed as odds ratios

Third, arrests were also stable pre-/post-training across racial/ethnic groups of officers, with or without controls for the marijuana policy change (see Table B-14). Fourth, use of force patterns mirror the arrest patterns (see Table B-15), which is expected, since use of force is derived from the arrest data.

The combination of officer race/ethnicity and suspect race/ethnicity is a logical supplemental analysis for the analyses of arrest counts. And a clear and consistent pattern emerged: post-training, arrests of Black suspects by White and Hispanic officers increased, while arrests of White suspects by White and Hispanic officers were unchanged; arrests of neither Black nor White suspects by Black officers were affected by the training (see Table B-16 in Appendix B). This is not a pattern that is interpretable as a training effect.

We also conducted event-level analysis of enforcement actions to estimate training impacts on enforcement by officers of different race/ethnicity. The models included all of the variables in the event-level models summarized above. Excluding Black officers, we find a post-training decrease in the likelihood that a person who is stopped is Hispanic (27.31 percent less likely) rather than White, a change that is statistically significant at the .05 level. We also find a statistically significant decrease in the likelihood that a person to whom a summons is issued is Hispanic, rather than White – a 12.54 percent decrease. Arrests were unchanged (see Table 7-12). Excluding Black and Hispanic officers, the same analysis shows only a statistically significant decrease in stops of Hispanics, however.

The more granular analysis at the level of individual events reveals two pre-/post-training differences that are consistent with the hypothesized training effect, but no larger pattern of effects on disparities. These analyses include not only the experimental control of the stepped-wedge randomization of the training intervention but also statistical controls for many of the other changes over time in the NYPD.

Table 7-12. Veil-of-Training Results, Excluding Black Officers

	Coefficient	St. Error	Exp(B)	% Change
Stops				
Black subject stopped				
Intervention Crossover	-0.155	0.141	0.86	-14.36
Intercept	1.090	0.612		
Hispanic subject stopped				
Intervention Crossover	-0.319*	0.158	0.73	-27.31
Intercept	-0.989	0.733		
Arrests				
Black subject arrested				
Intervention Crossover	0.040	0.034	1.04	4.08
Intercept	0.951	0.140		
Hispanic subject arrested				
Intervention Crossover	-0.015	0.037	0.99	-1.49
Intercept	0.017	0.155		
Summonses				
Black subject cited				
Intervention Crossover	-0.106	0.057	0.90	-10.06
Intercept	0.441	0.242		
Hispanic subject cited				
Intervention Crossover	-0.134*	0.065	0.87	-12.54
Intercept	-0.893	0.266		

*p < .05; **p<.01

We do not find evidence of training effects in analyses of other enforcement actions – frisks, searches, use of force – that exclude Black and/or Hispanic officers; see Table 7-13, below, for results excluding Black officers.

Officers' Attitudes

If the training has effects on enforcement behavior, we might expect the effects to be more pronounced among – or confined to – officers who are motivated to respond without prejudice and/or those who are concerned about discrimination as a social problem. In other words, the effect might be contingent on officers' attitudes and detectable primarily – or at all – among those who are most receptive to the training content and likely to apply the bias-management strategies. We might, in addition, hypothesize that the effects of the training would be more pronounced among officers whose understanding of the science of implicit bias is greater.

Table 7-13. Other Enforcement Actions, Excluding Black officers

	Coefficient	St. Error	Exp(B)	% Change
Frisks				
Intervention Crossover	-0.207	0.169	0.81	-18.7
X Subject Black	0.320	0.166	1.12	11.96
X Subject Hispanic	0.168	0.189	0.96	-3.82
Intercept	2.075*	0.180		
Only Frisks				
Intervention Crossover	-0.198	0.180	0.82	-17.96
X Subject Black	0.315	0.176	1.13	12.52
X Subject Hispanic	0.268	0.201	1.07	7.25
Intercept	0.365	0.183		
Searches				
Intervention Crossover	-0.132	0.161	0.88	-12.37
X Subject Black	0.001	0.156	0.88	-12.28
X Subject Hispanic	-0.094	0.178	0.80	-20.23
Intercept	-0.784*	0.161		
Use of Force – Stops				
Intervention Crossover	0.256	0.277	1.29	29.18
X Subject Black	0.015	0.270	1.31	31.13
X Subject Hispanic	-0.178	0.318	1.08	8.11
Intercept	-2.301*	0.276		
Use of Force – Arrests				
Intervention Crossover	0.045	0.179	1.05	4.60
X Subject Black	-0.197	0.175	0.86	-14.10
X Subject Hispanic	-0.243	0.195	0.82	-17.96
Intercept	-5.572*	0.317		

*p < .05; **p<.01

The training-day survey was administered anonymously, but it included items on officers' sex, race/ethnicity, age, and length of service. Among the 7,413 respondents who answered any of these items, 5,557 respondents answered all of these items; 488 answered three of the four. In addition, the survey platform captured the date and time at which respondents began the survey, which quite likely was the same day on which the officers were trained. Since the personnel information provided by NYPD also includes information on officers' sex, race, age, and length of service, as well as the date on which officers were trained, it is possible in many instances to reliably link training-day survey data to personnel records and, on that basis, to records of enforcement activity using the anonymized tax IDs of individual officers.[236]

[236] We are grateful to Greg Drake for suggesting this kind of data linkage.

From the personnel records, we formed the various permutations of the values for officers' sex, race/ethnicity, age (in the same categories specified on the survey instrument), and years of service. Aggregating those permutations by date, we identified the permutations that were unique on each day, and then merged the personnel data with the survey data. For a subset of 1,973 officers for whom a match was possible, then, we can analyze the relationships between officers' attitudes and their enforcement activity.

The officers whose survey responses could be linked in this way to administrative records are not much different, as a group, than the larger population of patrol officers who attended the training on days on which the survey was administered (see Table 7-14, below). Women are overrepresented among the former, inasmuch as they were individually more likely to form a unique permutation of characteristics on any given day. Transit officers were slightly underrepresented and housing officers slightly overrepresented among the "matched" respondents, who were also somewhat older and more experienced. White officers were somewhat underrepresented.

The "matched" officers were also quite similar to the entire population of survey respondents in terms of their beliefs and attitudes (see Table 7-15, below). They were slightly more knowledgeable about implicit bias and procedural justice, both before and after the training, but the increments on the scales are of small magnitude. They were also somewhat more motivated to respond without prejudice, somewhat more concerned about discrimination and somewhat less skeptical about discrimination as a social problem.

We performed event-level analyses of the matched officers' stops, arrests, and summonses, predicting the race/ethnicity of the subjects against whom officers took enforcement action. We formed interaction terms that allowed the estimated effect of the training to vary across groups of officers: officers: (1) those with high or lower motivation to respond without prejudice (with high motivation operationalized as a scale score of 4 or greater); and officers with high- or lower concern about discrimination (with high concern operationalized as a scale score of 4 or greater); (2) those with a stringer grasp of implicit bias concepts, based on their scores on the two implicit bias scales. The results are shown in Tables 7-16 and 7-17, below.

These analyses provide no support for the hypothesized effects of the training. The only statistically significant coefficient in the analyses of attitudes indicates that the likelihood that a person stopped was Hispanic decreased for officers with *lower* levels of motivation and concern. The only statistically significant coefficients in the analyses of implicit bias beliefs indicate that the likelihoods that a person stopped was Hispanic or Black decreased for officers with *lower* levels of implicit bias knowledge. Insofar as we can test for the effect of the training conditional on officers' attitudes or knowledge, we can find no evidence of a training effect.

Table 7-14. Training-Day Survey Respondents Matched to Administrative Records

	All trained	Surveyed population	Matched
Bureau			
Patrol	79.75%	79.13%	79.22%
Transit	9.86%	9.92%	11.2%
Housing	10.38%	10.94%	9.58%
Female	19%	19%	26%
Age	33.64	33.62	34.85
Years of service (May 2018)	5.98	6.01	6.70
Race			
White	44.1%	44.1%	38.2%
Asian	9.5%	9.6%	10.9%
Black	15.2%	15.2%	16.9%
Hispanic	31.2%	31.0%	34.0%
Indian	0.1%	0.1%	0%
Training block			
Treatment A	9.4%	0%	0%
Treatment B	9.4%	6.4%	4.8%
Treatment C	10.2%	16.0%	20.3%
Treatment A&B make-up	1.6%	2.5%	3.1%
Treatment D	10.4%	16.3%	17.1%
Treatment E	10.0%	15.7%	14.2%
Control A	9.9%	15.5%	16.5%
Control B	9.1%	10.2%	7.9%
Control C	9.8%	0%	0%
Control D	9.5%	0.7%	0.5%
Control E	10.7%	16.7%	15.5%
N	14,741	9,415	1,973

Table 7-15. Beliefs and Attitudes of Matched Respondents

	All surveyed		Matched	
	Pre	Post	Pre	Post
IB1	3.42	3.77	3.48	3.84
IB2	2.57	2.36	2.55	2.30
Motivation	4.09	4.14	4.13	4.15
Concern	3.38	3.59	3.47	3.69
Skepticism	3.15	3.05	3.11	3.00
PJ	3.74	3.87	3.76	3.93
N (range)	2721-2774	3861-3920	868-886	1055-1065

Table 7-16. Veil-of-Training Results by Officers' Attitudes

	Coefficient	St. Error	Exp(B)	% Change
Stops				
Black subject stopped				
Intervention Crossover	-0.413	0.517	0.66	-33.83
X Motivation	0.125	0.459	0.75	-25.02
X Concern	0.076	0.416	0.71	-28.61
Intercept	-0.415	1.370		
Hispanic subject stopped				
Intervention Crossover	-1.00	0.606	0.37	-63.21
X Motivation	0.301	0.519	0.50	-50.29
X Concern	-0.017	0.448	0.36	-63.83
Intercept	-2.239	1.560		
Arrests				
Black subject arrested				
Intervention Crossover	0.107	0.133	1.11	11.29
X Motivation	-0.004	0.141	1.11	10.85
X Concern	-0.017	0.128	1.09	9.42
Intercept	0.157	0.321		
Hispanic subject arrested				
Intervention Crossover	0.099	0.132	1.10	10.41
X Motivation	0.110	0.141	1.23	23.24
X Concern	-0.213	0.132	0.89	-10.77
Intercept	0.353	0.345		
Summonses				
Black subject cited				
Intervention Crossover	-0.202	0.193	0.82	-18.29
X Motivation	0.097	0.232	0.90	-9.97
X Concern	0.182	0.210	0.98	-1.98
Intercept	-0.851	0.739		
Hispanic subject cited				
Intervention Crossover	-0.129	0.214	0.88	-12.10
X Motivation	-0.048	0.262	0.84	-16.22
X Concern	-0.115	0.250	0.78	-21.65
Intercept	-1.520	0.673		

*p < .05; **p<.01

Table 7-17. Veil-of-Training Results by Officers' Implicit Bias Beliefs

	Coefficient	St. Error	Exp(B)	% Change
Stops				
Black subject stopped				
Intervention Crossover	-0.232	0.464	0.79	-20.71
X IB1	0.439	0.395	1.23	23.00
X IB2	-0.681	0.433	0.40	-59.87
Intercept	0.191	1.264		
Hispanic subject stopped				
Intervention Crossover	-0.465	0.522	0.63	-37.19
X IB1	-0.293	0.440	0.47	-53.14
X IB2	-0.432	0.477	0.41	-59.22
Intercept	-2.474	1.400		
Arrests				
Black subject arrested				
Intervention Crossover	0.005	0.116	1.01	0.50
X IB1	-0.000	0.134	1.01	0.50
X IB2	0.200	0.139	1.23	22.75
Intercept	-0.092	0.411		
Hispanic subject arrested				
Intervention Crossover	0.069	0.121	1.07	7.14
X IB1	-0.120	0.142	0.95	-4.97
X IB2	0.192	0.142	1.30	29.82
Intercept	-0.328	0.467		
Summonses				
Black subject cited				
Intervention Crossover	-0.285	0.167	0.75	-24.80
X IB1	0.301	0.214	1.02	1.61
X IB2	0.242	0.196	0.96	-4.21
Intercept	0.016	0.692		
Hispanic subject cited				
Intervention Crossover	-0.368	0.201	0.69	-30.79
X IB1	0.151	0.264	0.80	-19.51
X IB2	0.272	0.245	0.91	-9.15
Intercept	0.006	0.747		

*p < .05; **p<.01

Conclusions

Our analysis of officers' responses to the follow-up survey indicate that officers have to a significant extent applied the FIP strategies for managing the effects of implicit bias. Our estimates of strategy utilization may be understated, inasmuch as we have reason to believe that the survey respondents disproportionately represent the officers whose concern about discrimination is lower than that of other officers. Thus, we see evidence in the survey data of efforts by officers to minimize the effects of implicit bias.

Our analyses of enforcement behavior spanned a number of forms of enforcement, including stops, frisks, searches, arrests, summonses, and use of force. Analysis was conducted at two levels – at the aggregate, training-block level and at the level of individual enforcement events. We also conducted aggregate analysis of citizen complaints. We sought to detect training effects among subsets of officers defined by their race/ethnicity and by their knowledge and attitudes. In short, we endeavored to detect training effects wherever we might expect that they would be detectable.

Empirical evidence that is supportive of the hypothesized training effect of reduced disparity in enforcement is, on the whole, spotty and weak at best. Of the many coefficients estimated, very few achieved statistical significance at the conventional .05 level of significance, and fewer still (two) were interpretable as an intended training effect. (Nor do the results indicate that the training "backfired," with unintended effects on enforcement patterns.) We find then, little evidence of the reductions in racial and ethnic disparities that we hypothesized would follow if officers practiced the FIP strategies for managing their biases – assuming that pre-training enforcement disparities stemmed, at least in part, from officers' unconscious biases. The analyses do not lead us to reject the null hypothesis of no training effects on enforcement behavior. That is not to say that the null hypothesis should be accepted on the basis of these findings, and we cannot draw the inference that the training had no effects on officers' behavior. Some readers may find such a conclusion consternating, but that is the logic of null hypothesis statistical testing.[237] We also must acknowledge that the analysis was complicated by other trends over time, even within the restricted time frame of the stepped-wedge design, so our inferences are tempered by the challenge of isolating training effects.

[237] See David Weisburd, Cynthia M. Lum, and Sue-Ming Yang, "When Can We Conclude that Treatments or Programs 'Don't Work'?" *The Annals of the American Academy of Political and Social Science* 587 (2003): 31-48.

Chapter 8
Summary and Implications

We designed and executed this evaluation of the NYPD's implicit bias training with the expectation that the empirical evidence that it would produce would add to the very thin base of knowledge about the efficacy of providing implicit bias training for police. A large body of social psychological evidence shows that many people hold unconscious biases and are susceptible to the influence of those biases. Police officers are as likely as anyone to form such biases. However, if those biases affect police officers' behavior, the consequences are of profound significance for the community, the police department, and the officers themselves. Under those circumstances, effective training would be of enormous benefit.

Our findings can be summarized succinctly. First, we found that the training was associated with detectable pre- and post-training differences in officers' awareness of and knowledge about implicit bias, which were all in the hypothesized direction, and which we interpret as a training impact. Post-training, officers were more likely to understand that unconscious biases could affect their interactions with particular groups of people even if they consciously reject the stereotypes on which those biases are based. They were also more likely to grasp that the effects of such biases could lead them to be either over-vigilant and act with undue aggressiveness, or under-vigilant and act in ways that make them less safe. And they were more likely to believe that though it is difficult to eliminate unconscious biases, it is possible to manage them. Correspondingly, we found that officers were less likely, post-training, to believe that only racist officers engage in biased policing, or that nothing could be done about their unconscious biases and their behavioral consequences.

Enhancements of officers' knowledge about and awareness of implicit biases, and their significance for officers' and community members' well-being, represent the most proximate impacts of the training. Such effects would also seem to be necessary, if not sufficient, for achieving effects on officers' performance on the street: on the quality of their interactions with the citizenry and on their judgments about enforcement, including decisions to stop, frisk, search, arrest, and issue a summons. The findings about pre- and post-training differences in officers' beliefs are quite important, then.

Pre- and post-training differences extended to officers' attitudes as well. Even though all racial and ethnic groups of officers espoused a high mean level of motivation to act without prejudice prior to training, we found a still higher level in most groups post-training – a motivation based on their personal values. In addition, officers tended to indicate, post-training, a greater concern about discrimination as a social problem, and a recognition that bias in policing is a legitimate concern of the public. Officers were less likely, post-training, to characterize bias in policing as a "fiction" generated by the media.

None of the differences that we detected in officers' beliefs and attitudes were of greater than moderate magnitude, however, and some were quite small. On our scales, values of which potentially ranged from 1 to 5, the differences were no greater than 0.35. To illustrate the differences more concretely: 32.9 percent of the pre-training respondents agreed (somewhat or strongly) that "stereotypes about particular groups could influence my interactions with them, without my awareness," and 54.7 percent of the post-training respondents agreed with the statement. Fifty-eight percent of the pre-training respondents, and 69 percent of post-training respondents, agreed that "policing based on stereotypes or biases can make police unsafe." Overall, little more than half of the officers achieved "passing" scores on the post-training survey on the day of training, if we treated the survey items as questions on a final exam. Though the size of the sample for the training-day survey was sufficient to establish that the differences were greater than zero – i.e., "statistically significant" – the data do not suggest that the training had substantively large effects on officers' knowledge or attitudes. In this respect, our findings parallel those of the Urban Institute's evaluation of a different implicit bias curriculum.

Second, we found that NYPD supervisors are, to a significant degree, playing the role prescribed for them in monitoring their subordinates' performance for signs of bias and intervening as needed. They did so despite the challenges that inhere in this task: the ambiguity of the information on which a suspicion of bias can rest; the sensitivity of the issue; the difficulty of broaching the question without seeming to make an accusation. Furthermore, sergeants undertook this function in the context of a number of other burdens that court-ordered reforms placed on their shoulders. Thus, one important element of the organizational environment offered some support and reinforcement for the training.

Third, we found that officers' self-reported use of the bias-management strategies in which they were trained was moderately high. Officers reported that the strategies were not difficult to use. About one-quarter used them frequently, and nearly one-third used them sometimes (and these estimates may be understated, given the likely pattern of non-response).

Even so, we could detect little evidence that the effects of training extended to the reduction of racial and ethnic disparities in enforcement, the likes of which would represent behavioral manifestations of training impacts. We should repeat, in this connection, that neither the NYPD nor FIP, LLC, identified changes in officers' behavior as an objective of the training. We should also add that it is very difficult to isolate the effects of the training from other forces that produce disparate enforcement outcomes. Training impacts might be a signal that is easily lost in the noise of everyday police work. Research on police arrest decisions has pointed to a host of potential influences, including those that are attributes of the immediate situation (such as the seriousness of the offense, preferences of complainants, and the sobriety and civility of the suspected

offender), characteristics of the officer(s), features of the neighborhood context, and structure and practices of the police organization.[238] Estimating the effect of a single training curriculum on officers' decisions to invoke the law or otherwise exercise police authority may well be akin to finding the proverbial needle in a haystack.

Furthermore, it has been presumed but not demonstrated that enforcement disparities stem, at least in part, from officers' implicit biases. As we discussed in Chapter 1, even though research has shown that police officers, like the general public, hold unconscious biases, no scientific evidence directly links officers' implicit bias with enforcement disparities. To the contrary, the evidence – which is thin, to be sure – suggests that officers practice controlled responses even without implicit bias training. (Indeed, some of officers' self-reported use of the FIP strategies may reflect long-standing practice – e.g., to "slow it down" – rather than the application of new concepts.) If disparities stem from forces other than implicit bias, then even a well-designed training that is flawlessly delivered cannot be expected to alter patterns of police enforcement behavior.

There are still other possible explanations for the largely null findings with respect to behavioral impacts. First, it might be that the strategies for managing bias are not effective, though they are consistent with findings of social psychological research, as we reviewed it in Chapter 2. Second, it might be that too few officers chose to use the strategies, or they underutilized them, though their self-reported use of the strategies may refute this conclusion. Third, and more plausibly, it might be that the training did not suffice to make officers proficient in the strategies. The effective application of the FIP skills probably requires some practice, for which eight hours of training could not provide.

One factor that is sometimes at the root of evaluation findings that fall short of expectations is a failure of implementation, that is, that the intervention whose effects are estimated was not applied, or not applied fully or properly. One threat in the evaluation of a training curriculum is that the instructors would deviate from the prescribed curricular content, or even undermine it (intentionally or unintentionally). We can most likely rule out implementation failure as an explanation for weak or null findings, since the training was delivered by FIP instructors; we presume that the curriculum as delivered enjoyed a high degree of fidelity to the curricular model.

The training might have salutary effects that our evaluation was not designed to assess, despite its breadth. Any agency's investment in implicit bias training sends a

[238] See National Research Council, *Fairness and Effectiveness in Policing: The Evidence*, Committee to Review Research on Police Policy and Practices, Wesley Skogan and Kathleen Frydl, eds. Committee on Law and Justice, Division of Behavioral and Social and Behavioral Sciences and Education (Washington, DC: National Academies Press, 2004), chaps. 4 & 5. Also see Robin S. Engel, Robert E. Worden, Nicholas Corsaro, Hannah D. McManus, Danielle L. Reynolds, Hannah Cochran, Gabrielle T. Isaza, and Jennifer Calnon Cherkauskas, *The Power to Arrest: Lessons from Research* (New York: Springer, 2019).

signal to both the rank-and-file and the public that fair and unbiased policing is a priority for the agency. It may contribute at the margin to police legitimacy, though we did not (and could not) incorporate public trust as an outcome into the evaluation design. Over the longer term, implicit bias training may also contribute at the margin to an agency environment that prizes equitable service and will not tolerate discrimination, sowing seeds of generational change among officers.

To put these findings in a larger context, we should note that research has seldom examined the effects of police training, and in the absence of empirical evidence, the public and its representatives may have an exaggerated faith in the benefits of police training. In 2000, a committee of experts formed by the National Research Council (NRC) undertook an assessment of the state of research on police policies and practices. Their report, published in 2004, included a discussion of research on police training that was as brief as the studies of training were scarce:

> Knowledge of the effects of police training is limited primarily to whether more training produces the desired change in police practice. This is typically acquired by conducting a controlled experiment (comparing police who have received training with those who have not) or by a study that measures the correlation between the amount of training officers have received and some police practice, while statistically controlling for the effects of other influences, such as years of experience. There are too few of either type of study available to shed light on the effects of training.[239]

They added that "prior research has not taken into account the substantive content of the training, modes of instruction, the abilities of the instructors, the timing of the training, or the organizational support for reinforcing the objectives of the training program."[240]

The state of the research on police training has not improved much since the NRC Committee completed its review. Writing in 2016, Lum et al. affirmed that the NRC Committee's characterization of the evidence on training effects remained accurate: "there is little or no evaluation evidence for most of the categories of training recommended by the Task Force."[241] They succinctly summarized the evidence on crisis intervention team (CIT) training, which has been shown to have positive impacts on officers' beliefs and attitudes relating to interactions with persons with mental illness; they also note that a systematic review found "null overall effects" on arrests of and use of force on persons with mental illness.[242] Procedural justice training has been subjected to some evaluation, which has detected some effects on beliefs and attitudes; behavioral impacts have not been

[239] National Research Council, *Fairness and Effectiveness in Policing: The Evidence*, p. 142.
[240] Ibid., p. 141.
[241] Lum et al., *An Evidence-Assessment of the Recommendations of the President's Task Force on 21st Century Policing — Implementation and Research Priorities*, p. 34.
[242] Ibid., p. 36. Also see Amy C. Watson, Victor C. Ottati, Melissa Morabito, Jeffrey Draine, Amy N. Kerr, and Beth Angell, "Outcomes of Police Contacts with Persons with Mental Illness: The Impact of CIT. *Administration and Policy in Mental Health and Mental Services Research* 37 (2010): 302-317.

assessed.[243] The United Kingdom's College of Policing conducted an experimental evaluation of a pilot stop and search training program, delivered in six police forces. The impact evaluation found that the training: (1) marginally improved officers' stop and search knowledge, which was already strong; (2) had a modest impact on officers' attitudes; (3) affected officers' anticipated, or hypothetical search decisions (based on responses to vignettes); and (4) had a small (but statistically insignificant) effect on officers' recorded search rates, and no effect on racial/ethnic disparities in searches.[244] Engel, McManus, and Herold recently assessed 64 studies of de-escalation training, mainly in nursing and psychiatry; "Only one study evaluating a training explicitly designed to reduce officer use of force in their interactions with citizens was identified."[245] Even on the training domains that have been subjects of inquiry, the evidence base is not strong. Consequently, prior research does not afford us a baseline of estimated training impacts against which to compare our findings about the effects of implicit bias training in the NYPD.

A priori, a single day of classroom training should perhaps not be expected either to fully inform officers about the science of bias and the utility of bias-management strategies, or to translate directly into practice in the hurly-burly of police work. Indeed, Lum and her colleagues point to the challenge to any form of training in the "transfer" of learning into performance. Gaps between learning and performance "can be explained by a combination of learner (e.g., cognitive ability, motivation level), intervention (e.g., reinforcement, error-based examples, modeling), and work environment (e.g., peer and supervisor support, organizational culture) characteristics."[246]

To put the findings in a still larger context, we would note that the implicit bias training could not by itself eliminate disparities in policing, for the disparities stem from many sources in a society marked by yawning economic inequality, the lingering effects of historical discrimination, and the effects of contemporary discrimination in many domains of life other than policing. Police work is conducted in a social and economic context, and disparities in employment, education, housing, health care, etc., inevitably manifest themselves in the process and outcomes of police work. If and when implicit

---

[243] Dennis P. Rosenbaum and Daniel S. Lawrence, *Teaching Respectful Police-Citizen Encounters and Good Decision Making: Results of a Randomized Control Trial with Police Recruits* (Chicago: National Police Research Platform, no date); Wesley G. Skogan, Maarten Van Craen, and Cari Hennessy, "Training Police for Procedural Justice," *Journal of Experimental Criminology* 11 (2015): 319-334.

[244] Joel Miller and Banos Alexandrou, *College of Policing Stop and Search Training Experiment: Impact Evaluation* (London: College of Policing, 2016). Also see Also see Chris Giacomantonio, Tal Jonathan-Zamir, Yael Litmanovitz, Ben Bradford, Matthew Davies, Lucy Strang, and Alex Sutherland, *College of Policing Stop and Search Training Experiment: Process Evaluation* (London: College of Policing, 2016).

[245] Robin S. Engel, Hannah D. McManus, and Tamara D. Herold, *The Deafening Demand for De-escalation Training: A Systematic Review and Call for Evidence in Police Use of Force Reform* (Cincinnati: IACP/UC Center for Police Research and Policy, 2019), p. 30.

[246] Lum et al., *An Evidence-Assessment of the Recommendations of the President's Task Force on 21st Century Policing — Implementation and Research Priorities*, p. 34.

bias training has the intended effects, disparities will be reduced at the margin, and only at the margin.

We intended this evaluation not only – or even primarily – as a report card for the NYPD but rather as an addition to the body of knowledge about police practice and management. The unusual and in some respects unique attributes of the NYPD demand caution in generalizing, and in any case, the effects of training are likely to hinge on characteristics of the organization in which it is provided, and perhaps also the broader historical climate in which it is received. The training and our evaluation were undertaken following years of declining enforcement levels, reducing the incidence of discretionary decision-making that is potentially subject to the influence of implicit bias, and correspondingly reducing the likelihood that effective efforts to control implicit bias would have detectable effects. For all of these reasons, we would refrain from offering recommendations that police agencies should or should not implement implicit bias training; such recommendations should, we believe, await further research in other agency and community settings. That research would, ideally, attend not only to estimating the impacts of the training but also to the accumulation of evidence concerning variations in training content and delivery modalities. We trust, however, that the cooperation of the NYPD with the demands of the evaluation, and the support of the Arnold Ventures, advances our understanding of policing and police reform.

Appendix A

Table A-1. Differences between Pre- and Post-Training Respondents

	Pre-training	**Post-training**	**H_0 no difference**
Length of service			
Less than 1 year	3.6	3.5	p = 0.891*
1-3 years	34.9	33.6	
4-6 years	19.6	20.0	
7-13 years	24.0	23.1	
14 to 20 years	12.5	12.8	
More than 20 years	2.2	1.8	
Mean	6.82	6.84	
Unreported	3.2	5.2	p < .001*
Rank			
Patrol officer	95.1	93.0	p = .243*
Detective/other	2.2	1.8	
Unreported	2.6	5.2	p < .001*
Sex			
Male	82.2	79.1	p = .001*
Female	16.3	19.6	
Unreported	1.4	1.3	p = .570*
Age			
18 to 24	9.1	9.0	p = .813*
25 to 30	35.8	36.9	
31 to 35	24.1	22.8	
36 to 40	15.6	15.8	
41 to 45	8.5	8.6	
Over 45	5.9	6.0	
Unreported	0.9	1.0	p = .878*
Race			
Black	10.9	13.1	p = .007*
Hispanic	26.8	27.9	p = .319*
White	44.2	40.6	p = .005*
Asian	8.6	8.3	p = .637*
Other	4.1	5.4	p = .013*
Multi-racial	5.5	4.6	p = .137*
Unreported	2.3	2.6	p = .389*

Education			
Less than high school	0.2	0.2	p = .065**
High school diploma / GED	3.1	2.6	
Some junior college	10.3	10.2	
Associate's degree	19.7	20.5	
More than two years college	17.9	14.5	
Bachelor's degree	40.0	42.8	
Some graduate courses	3.5	3.4	
Graduate degree	4.6	5.4	
Unreported	0.6	0.5	p = .779*
Military experience			
No	88.2	87.7	p = .452*
Yes	11.2	11.8	
Unreported	0.6	0.5	p = .557*
Total N			
Background section skipped			

* t test

** Mann-Whitney U test

Note: The percentages exclude respondents who skipped the last section of the survey, on background and demographic characteristics. Length of service was reported in years, rounded to the lower integer for analysis, with frequency counts reported in Table A in terms of categories.

Appendix B

Table B-1. Estimates of Implicit Bias Treatment Effect from the Hussey and Hughes Model, and Model Extensions (N = 120 Experimental Block Time Periods)

	Equation 1		Equation 2		Equation 3	
Stops	Coefficient	St. Error	Coefficient	St. Error	Coefficient	St. Error
Intervention Crossover	-.013	.059	-.013	.059	.033	.070
Marijuana (MJ) Policy Change	--	--	.008	.100	.048	.110
Intervention Crossover x MJ Policy Change	--	--	--	--	-.087	.087
Intercept	4.90**	.067	4.90	.067	4.89	.066
Summonses	Coefficient	St. Error	Coefficient	St. Error	Coefficient	St. Error
Intervention Crossover	.072	.036	.072	.036	.161**	.043
Marijuana (MJ) Policy Change	--	--	-.392**	.077	-.301**	.078
Intervention Crossover x MJ Policy Change	--	--	--	--	-.180**	.046
Intercept	6.44**	.050	6.44**	.050	6.44**	.050
Arrests	Coefficient	St. Error	Coefficient	St. Error	Coefficient	St. Error
Intervention Crossover	.015	.017	.015	.017	.059**	.019
Marijuana (MJ) Policy Change	--	--	-.143**	.031	-.106**	.032
Intervention Crossover x MJ Policy Change	--	--	--	--	-.081**	.021
Intercept	7.26**	.022	7.26**	.022	7.25**	.022

*p < .05; **p<.01

Table B-2. Estimates of Implicit Bias Treatment Effect from the Hussey and Hughes Model, and Model Extensions: Police Stops – Frisks, Searches, and Use of Force (N = 120 Experimental Block Time Periods)

	Equation 1		Equation 2		Equation 3	
Frisks	Coefficient	St. Error	Coefficient	St. Error	Coefficient	St. Error
Intervention Crossover	-.014	.065	-.014	.065	.029	.097
Marijuana (MJ) Policy Change	--	--	.048	.105	.087	.109
Intervention Crossover x MJ Policy Change	--	--	--	--	-.083	.099
Intercept	4.44**	.105	4.44**	.105	4.43**	.104
Searches						
Intervention Crossover	-.054	.077	-.054	.077	.119	.108
Marijuana (MJ) Policy Change	--	--	.215	.126	.348*	.127
Intervention Crossover x MJ Policy Change	--	--	--	--	-.307**	.108
Intercept	3.71**	.067	3.71**	.067	3.68	.067
Use of Force						
Intervention Crossover	.167	.187	.167	.187	.286**	.103
Marijuana (MJ) Policy Change	--	--	-.097	.153	.019	.179
Intervention Crossover x MJ Policy Change	--	--	--	--	-.237	.136
Intercept	3.16**	.088	3.16**	.088	3.15**	.088

*p < .05; **p<.01

Table B-3. Estimates of Implicit Bias Treatment Effect from the Hussey and Hughes Model, and Model Extensions: Arrests – Use of Force (N = 120 Experimental Block Time Periods)

	Equation 1		Equation 2		Equation 3	
	Coefficient	St. Error	Coefficient	St. Error	Coefficient	St. Error
Intervention Crossover	-.023	.023	-.023	.023	.022	.030
Marijuana (MJ) Policy Change	--	--	.045	.044	.085*	.045
Intervention Crossover x MJ Policy Change	--	--	--	--	-.086*	.036
Intercept	5.51	.034	5.51	.034	5.50	.034

*p < .05; **p<.01

Table B-4. Estimates of Implicit Bias Treatment Effect from the Hussey and Hughes Model, and Model Extensions: Citizen Complaints (N = 120 Experimental Block Time Periods)

	Equation 1		Equation 2		Equation 3	
	Coefficient	St. Error	Coefficient	St. Error	Coefficient	St. Error
Intervention Crossover	.047	.050	.047	.050	.015	.060
Marijuana (MJ) Policy Change	--	--	.128	.090	.097	.098
Intervention Crossover x MJ Policy Change	--	--	--	--	.063	.068
Intercept	3.70**	.068	3.70**	.068	3.70**	.068

*p < .05; **p<.01

Table B-5. Contemporaneous Estimates of Implicit Bias Treatment Effect from the Hussey and Hughes Model, and Model Extensions: Stops by Race (N = 120 Experimental Block Time Periods)

	Equation 1		Equation 2		Equation 3	
Black Only Stops	Coefficient	St. Error	Coefficient	St. Error	Coefficient	St. Error
Intervention Crossover	.037	.070	.037	.070	.095	.086
Marijuana (MJ) Policy Change	--	--	-.054	.113	.007	.121
Intervention Crossover x MJ Policy Change	--	--	--	--	-.113	.108
Intercept	4.41**	.082	4.41**	.082	4.40**	.081
White Only Stops						
Intervention Crossover	.116	.128	.116	.128	.047	.137
Marijuana (MJ) Policy Change	--	--	.012	.229	-.025	.257
Intervention Crossover x MJ Policy Change	--	--	--	--	.106	.182
Intercept	1.76**	.153	1.76**	.153	1.78**	.153
Black and Hispanic Stops						
Intervention Crossover	.010	.070	.010	.070	.085	.082
Marijuana (MJ) Policy Change	--	--	-.024	.112	.048	.122
Intervention Crossover x MJ Policy Change	--	--	--	--	-.147	.103
Intercept	4.55**	.084	4.55**	.084	4.54**	.082
Hispanic Only Stops						
Intervention Crossover	-.127	.079	-.127	.079	-.057	.100
Marijuana (MJ) Policy Change	--	--	.120	.131	.183	.142
Intervention Crossover x MJ Policy Change	--	--	--	--	-.132	.124
Intercept	3.78**	.076	3.78**	.076	3.77**	.075

*p < .05; **p<.01

Table B-6. Contemporaneous Estimates of Implicit Bias Treatment Effect from the Hussey and Hughes Model, and Model Extensions: Summonses by Race (N = 120 Experimental Block Time Periods)

	Equation 1		Equation 2		Equation 3	
Black Only Summonses	Coefficient	St. Error	Coefficient	St. Error	Coefficient	St. Error
Intervention Crossover	.051	.044	.051	.044	.130**	.051
Marijuana (MJ) Policy Change	--	--	-.368**	.079	-.279**	.07886
Intervention Crossover x MJ Policy Change	--	--	--	--	-.168**	.063
Intercept	5.71**	.056	5.71**	.056	5.70**	.055
White Only Summonses						
Intervention Crossover	-.044	.082	-.044	.082	.069	.117
Marijuana (MJ) Policy Change	--	--	-.331**	.127	-.250**	.133
Intervention Crossover x MJ Policy Change	--	--	--	--	-.194	.140
Intercept	3.92**	.096	3.92**	.096	3.92**	.096
Black and Hispanic Summonses						
Intervention Crossover	.060	.045	.060	.045	.131**	.054
Marijuana (MJ) Policy Change	--	--	-.367**	.084	-.284**	.092
Intervention Crossover x MJ Policy Change	--	--	--	--	-.153**	.060
Intercept	5.92**	.061	5.92**	.061	5.92**	.061
Hispanic Summonses						
Intervention Crossover	.121	.154	.121	.154	.210**	.072
Marijuana (MJ) Policy Change	--	--	-.327**	.109	-.226**	.113
Intervention Crossover x MJ Policy Change	--	--	--	--	-.190**	.082
Intercept	5.39**	.073	5.39**	.073	5.37**	.072

*p < .05; **p<.01

Table B-7. Contemporaneous Estimates of Implicit Bias Treatment Effect from the Hussey and Hughes Model, and Model Extensions: Arrests by Race (N = 120 Experimental Block Time Periods)

	Equation 1		Equation 2		Equation 3	
Black Only Arrests	Coefficient	St. Error	Coefficient	St. Error	Coefficient	St. Error
Intervention Crossover	.028	.018	.028	.018	.063**	.022
Marijuana (MJ) Policy Change	--	--	-.128**	.033	-.095**	.034
Intervention Crossover x MJ Policy Change	--	--	--	--	-.070**	.025
Intercept	6.55**	.025	6.55**	.024	6.54**	.024
White Only Arrests						
Intervention Crossover	-.050	.032	-.050	.032	-.047	.048
Marijuana (MJ) Policy Change	--	--	-.100	.069	-.099	.071
Intervention Crossover x MJ Policy Change	--	--	--	--	-.003	.054
Intercept	4.69**	.049	4.69**	.049	4.69**	.049
Black and Hispanic Arrests						
Intervention Crossover	.040	.017	.040	.017	.077**	.020
Marijuana (MJ) Policy Change	--	--	-.140**	.033	-.104**	.036
Intervention Crossover x MJ Policy Change	--	--	--	--	-.073**	.024
Intercept	6.76**	.023	6.76**	.023	6.76**	.024
Hispanic Arrests						
Intervention Crossover	.022	.025	.022	.025	.083**	.030
Marijuana (MJ) Policy Change	--	--	-.166**	.043	-.109**	.043
Intervention Crossover x MJ Policy Change	--	--	--	--	-.118**	.032
Intercept	6.34**	.033	6.34**	.038	6.33**	.034

*p < .05; **p<.01

Table B-8. Contemporaneous Estimates of Implicit Bias Treatment Effect from the Hussey and Hughes Model, and Model Extensions: Stops - Searches by Race (N = 120 Experimental Block Time Periods)

	Equation 1		Equation 2		Equation 3	
Black Only Searches	Coefficient	St. Error	Coefficient	St. Error	Coefficient	St. Error
Intervention Crossover	.018	.098	.018	.098	.148	.141
Marijuana (MJ) Policy Change	--	--	.138	.148	.243	.150
Intervention Crossover x MJ Policy Change	--	--	--	--	-.235	.147
Intercept	3.14**	.084	3.14**	.084	3.13**	.083
White Only Searches						
Intervention Crossover	-.206	.200	-.206	.200	.231	.236
Marijuana (MJ) Policy Change	--	--	.324	.316	.532	.332
Intervention Crossover x MJ Policy Change	--	--	--	--	-.646*	.257
Intercept	0.89**	.240	0.89**	.240	0.82**	.246
Black and Hispanic Searches						
Intervention Crossover	.005	.095	.005	.095	.153	.135
Marijuana (MJ) Policy Change	--	--	.146	.145	.270	.144
Intervention Crossover x MJ Policy Change	--	--	--	--	-.271	.141
Intercept	3.30**	.082	3.30**	.082	3.29**	.082
Hispanic Only Searches						
Intervention Crossover	-.177	.109	-.177	.109	.015	.136
Marijuana (MJ) Policy Change	--	--	.361*	.180	.521**	.190
Intervention Crossover x MJ Policy Change	--	--	--	--	-.350*	.156
Intercept	2.71**	.111	2.71**	.111	2.68**	.111

*p < .05; **p<.01

Table B-9. Contemporaneous Estimates of Implicit Bias Treatment Effect from the Hussey and Hughes Model, and Model Extensions: Stops - Frisks by Race (N = 120 Experimental Block Time Periods)

	Equation 1		Equation 2		Equation 3	
Black Only Frisks	Coefficient	St. Error	Coefficient	St. Error	Coefficient	St. Error
Intervention Crossover	.060	.075	.060	.075	.106	.107
Marijuana (MJ) Policy Change	--	--	-.049	.116	-.000	.117
Intervention Crossover x MJ Policy Change	--	--	--	--	-.091	.113
Intercept	3.94**	.104	3.94**	.104	3.93**	.104
White Only Frisks						
Intervention Crossover	.071	.193	.071	.193	.254	.204
Marijuana (MJ) Policy Change	--	--	-.140	.294	-.062	.317
Intervention Crossover x MJ Policy Change	--	--	--	--	-.261	.261
Intercept	1.08**	.220	1.08**	.220	1.05**	.226
Black and Hispanic Frisks						
Intervention Crossover	.029	.072	.029	.072	.088	.105
Marijuana (MJ) Policy Change	--	--	-.018	.111	.039	.112
Intervention Crossover x MJ Policy Change	--	--	--	--	-.116	.109
Intercept	4.14**	.102	4.14**	.102	4.11**	.102
Hispanic Only Frisks						
Intervention Crossover	-.181	.097	-.181	.097	-.179	.128
Marijuana (MJ) Policy Change	--	--	-.298	.180	.300	.189
Intervention Crossover x MJ Policy Change	--	--	--	--	-.003	.146
Intercept	3.29**	.131	3.29**	.131	3.29**	.131

*p < .05; **p<.01

Table B-10. Contemporaneous Estimates of Implicit Bias Treatment Effect from the Hussey and Hughes Model, and Model Extensions: Stops – Use of Force in Stops by Race (N = 120 Experimental Block Time Periods)

	Equation 1		Equation 2		Equation 3	
Black Only Stops – Force Used	Coefficient	St. Error	Coefficient	St. Error	Coefficient	St. Error
Intervention Crossover	.327**	.108	.327**	.108	.454**	.111
Marijuana (MJ) Policy Change	--	--	-.242	.164	-.108	.201
Intervention Crossover x MJ Policy Change	--	--	--	--	-.261	.173
Intercept	2.66**	.096	2.66**	.096	2.64**	.096
White Only Stops – Force Used						
Intervention Crossover	.027	.208	.027	.208	.358	.315
Marijuana (MJ) Policy Change	--	--	-.027	.294	.160	.300
Intervention Crossover x MJ Policy Change	--	--	--	--	-.518	.355
Intercept	0.57*	.250	0.57*	.250	0.52*	.263
Black and Hispanic Stops – Force Used						
Intervention Crossover	.276*	.103	.276*	.103	.418**	.112
Marijuana (MJ) Policy Change	--	--	-.160	.161	-.009	.181
Intervention Crossover x MJ Policy Change	--	--	--	--	-.292	.160
Intercept	2.80**	.103	2.80**	.103	2.78**	.103
Hispanic Only Stops – Force Used						
Intervention Crossover	.038	.134	.038	.134	.136	.176
Marijuana (MJ) Policy Change	--	--	-.004	.239	.097	.248
Intervention Crossover x MJ Policy Change	--	--	--	--	-.199	.205
Intercept	1.92**	.169	1.92**	.169	1.90**	.169

*p < .05; **p<.01

Table B-11. Contemporaneous Estimates of Implicit Bias Treatment Effect from the Hussey and Hughes Model, and Model Extensions: Use of Force in Arrests by Race (N = 120 Experimental Block Time Periods)

	Equation 1		Equation 2		Equation 3	
Black Only Use of Force	Coefficient	St. Error	Coefficient	St. Error	Coefficient	St. Error
Intervention Crossover	-.039	.027	-.039	.027	-.004	.034
Marijuana (MJ) Policy Change	--	--	-.136**	.046	.171**	.049
Intervention Crossover x MJ Policy Change	--	--	--	--	-.071	.043
Intercept	4.82	.035	4.82	.035	4.81	.034
White Only Use of Force						
Intervention Crossover	-.080	.054	-.080	.054	-.072	.081
Marijuana (MJ) Policy Change	--	--	.018	.102	.023	.105
Intervention Crossover x MJ Policy Change	--	--	--	--	-.012	.087
Intercept	2.77	.081	2.77	.081	2.77	.084
Black and Hispanic Use of Force						
Intervention Crossover	-.021	.028	-.021	.028	.032	.034
Marijuana (MJ) Policy Change	--	--	.084	.055	.140**	.056
Intervention Crossover x MJ Policy Change	--	--	--	--	-.109*	.044
Intercept			5.05	.040	5.05	.039
Hispanic Use of Force						
Intervention Crossover	.032	.041	.032	.041	.122**	.046
Marijuana (MJ) Policy Change	--	--	-.060	.069	.029	.072
Intervention Crossover x MJ Policy Change	--	--	--	--	-.181**	.056
Intercept	4.56	.057	4.56	.057	4.54	.056

*p < .05; **p<.01

Table B-12. Contemporaneous Estimates of Implicit Bias Treatment Effect from the Hussey and Hughes Model, and Model Extensions: Stops by Officer Race/Ethnicity (N = 120 Experimental Block Time Periods)

	Equation 1		Equation 2		Equation 3	
White Officer Stops	Coefficient	St. Error	Coefficient	St. Error	Coefficient	St. Error
Intervention Crossover	-.008	.065	-.008	.065	-.009	.080
Marijuana (MJ) Policy Change	--	--	.086	.127	.085	.138
Intervention Crossover x MJ Policy Change	--	--	--	--	.001	.101
Intercept	4.05**	.078	4.05**	.078	4.05**	.078
Black Officer Stops						
Intervention Crossover	-.051	.119	-.051	.119	-.039	.137
Marijuana (MJ) Policy Change	--	--	.177	.218	.190	.250
Intervention Crossover x MJ Policy Change	--	--	--	--	-.025	.197
Intercept	2.79**	.138	2.79**	.138	2.79**	.138
Hispanic Officer Stops						
Intervention Crossover	-.025	.090	-.025	.090	.087	.107
Marijuana (MJ) Policy Change	--	--	-.054	.145	.060	.157
Intervention Crossover x MJ Policy Change	--	--	--	--	-.227	.135
Intercept	3.87**	.114	3.87**	.114	3.85**	.111
Other Officer Stops						
Intervention Crossover	.007	.116	.007	.116	.078	.126
Marijuana (MJ) Policy Change	--	--	-.274	.187	-.205	.213
Intervention Crossover x MJ Policy Change	--	--	--	--	-.140	.173
Intercept	2.32**	.138	2.32**	.138	2.31**	.140

*p < .05; **p<.01

Table B-13. Contemporaneous Estimates of Implicit Bias Treatment Effect from the Hussey and Hughes Model, and Model Extensions: Summonses by Officer Race/Ethnicity (N = 120 Experimental Block Time Periods).

	Equation 1		Equation 2		Equation 3	
White Officer Summonses	Coefficient	St. Error	Coefficient	St. Error	Coefficient	St. Error
Intervention Crossover	.049	.043	.049	.043	.133**	.052
Marijuana (MJ) Policy Change	--	--	-.360*	.086	-.284**	.088
Intervention Crossover x MJ Policy Change	--	--	--	--	-.160**	.057
Intercept	5.64**	.057	5.64**	.057	5.63**	.055
Black Officer Summonses						
Intervention Crossover	.019	.061	.019	.061	.071	.058
Marijuana (MJ) Policy Change	--	--	-.538	.151	-.474**	.168
Intervention Crossover x MJ Policy Change	--	--	--	--	-.116	.094
Intercept	4.56**	.098	4.56**	.098	4.55**	.090
Hispanic Officer Summonses						
Intervention Crossover	.132*	.055	.132*	.055	.234	.072
Marijuana (MJ) Policy Change	--	--	-.314**	.119	-.210	.120
Intervention Crossover x MJ Policy Change	--	--	--	--	-.206	.082
Intercept	5.24**	.096	5.24**	.096	5.22**	.094
Other Officer Summonses						
Intervention Crossover	.044	.074	.044	.074	.146	.093
Marijuana (MJ) Policy Change	--	--	-.627	.187	-.484*	.190
Intervention Crossover x MJ Policy Change	--	--	--	--	-.244*	.114
Intercept	3.95**	.094	3.95**	.094	3.93**	.098

*p < .05; **p<.01

Table B-14. Contemporaneous Estimates of Implicit Bias Treatment Effect from the Hussey and Hughes Model, and Model Extensions: Arrests by Officer Race/Ethnicity (N = 120 Experimental Block Time Periods).

	Equation 1		Equation 2		Equation 3	
White Officer Arrests	Coefficient	St. Error	Coefficient	St. Error	Coefficient	St. Error
Intervention Crossover	.019	.021	.019	.021	.049	.029
Marijuana (MJ) Policy Change	--	--	-.159	.044	-.136**	.044
Intervention Crossover x MJ Policy Change	--	--	--	--	-.053	.029
Intercept	6.31**	.032	6.31**	.032	6.31**	.032
Black Officer Arrests						
Intervention Crossover	-.017	.030	-.017	.030	.048	.037
Marijuana (MJ) Policy Change	--	--	-.063	.045	-.001	.050
Intervention Crossover x MJ Policy Change	--	--	--	--	-.128**	.043
Intercept	5.34**	.030	5.34**	.030	5.33**	.030
Hispanic Officer Arrests						
Intervention Crossover	.028	.025	.028	.025	.084**	.031
Marijuana (MJ) Policy Change	--	--	-.125**	.040	-.071	.041
Intervention Crossover x MJ Policy Change	--	--	--	--	-.109**	.034
Intercept	6.27**	.024	6.27**	.024	6.26**	.025
Other Officer Arrests						
Intervention Crossover	.021	.029	.021	.029	.043	.034
Marijuana (MJ) Policy Change	--	--	-.128*	.051	-.108	.061
Intervention Crossover x MJ Policy Change	--	--	--	--	-.041	.048
Intercept	4.72**	.040	4.72**	.040	4.72**	.040

*p < .05; **p<.01

Table B-15. Contemporaneous Estimates of Implicit Bias Treatment Effect from the Hussey and Hughes Model, and Model Extensions: Use of Force in Arrests by Officer Race/Ethnicity (N = 120 Experimental Block Time Periods).

	Equation 1		Equation 2		Equation 3	
White Officer UoF	Coefficient	St. Error	Coefficient	St. Error	Coefficient	St. Error
Intervention Crossover	.012	.029	.012	.029	.044	.037
Marijuana (MJ) Policy Change	--	--	.027	.063	.051	.065
Intervention Crossover x MJ Policy Change	--	--	--	--	-.055	.046
Intercept	4.48**	.044	4.48**	.047	4.48**	.044
Black Officer UoF						
Intervention Crossover	-.011	.056	-.012	.056	.028	.060
Marijuana (MJ) Policy Change	--	--	.094	.090	.141	.103
Intervention Crossover x MJ Policy Change	--	--	--	--	-.087	.080
Intercept	3.67**	.074	3.67**	.074	3.67**	.072
Hispanic Officer UoF						
Intervention Crossover	-.051	.045	-.051	.045	.044	.059
Marijuana (MJ) Policy Change	--	--	.076	.070	.167*	.070
Intervention Crossover x MJ Policy Change	--	--	--	--	-.186**	.067
Intercept	4.55	.064	4.55	.064	4.53	.065
Other Officer UoF						
Intervention Crossover	-.034	.056	-.034	.056	-.057	.059
Marijuana (MJ) Policy Change	--	--	.024	.094	-.004	.117
Intervention Crossover x MJ Policy Change	--	--	--	--	.052	.102
Intercept	3.25**	.063	3.25**	.063	3.25**	.063

*p < .05; **p<.01

Table B-16. Contemporaneous Estimates of Implicit Bias Treatment Effect from the Hussey and Hughes Model, and Model Extensions: Arrests by Arrestee Race/Ethnicity and Officer Race/Ethnicity (N = 120 Experimental Block Time Periods).

	Equation 1		Equation 2		Equation 3	
White Officer Arrests of Blacks and Hispanic Blacks	Coefficient	St. Error	Coefficient	St. Error	Coefficient	St. Error
Intervention Crossover	.054*	.023	.054*	.023	.089**	.027
Marijuana (MJ) Policy Change	--	--	-.154**	.056	-.123*	.058
Intervention Crossover x MJ Policy Change	--	--	--	--	-.065*	.030
Intercept	5.77**	.036	5.77**	.036	5.76**	.036
White Officer Arrests of Whites						
Intervention Crossover	-.055	.041	-.055	.041	-.043	.067
Marijuana (MJ) Policy Change	--	--	-.096	.082	-.090	.083
Intervention Crossover x MJ Policy Change	--	--	--	--	-.01	.065
Intercept	3.83**	.069	3.83**	.069	3.82**	.068
Black Officer Arrests of Blacks and Hispanic Blacks						
Intervention Crossover	-.023	.035	-.023	.035	.029	.042
Marijuana (MJ) Policy Change	--	--	-.076	.059	-.016	.068
Intervention Crossover x MJ Policy Change	--	--	--	--	-.113*	.057
Intercept	4.96**	.039	4.96**	.039	4.95**	.039
Black Officer Arrests of Whites						
Intervention Crossover	-.040	.096	-.043	.096	-.117	.140
Marijuana (MJ) Policy Change	--	--	-.130	.152	-.176	.165
Intervention Crossover x MJ Policy Change	--	--	--	--	.123	.167
Intercept	2.77	.114	2.77	.114	2.79	.0117

Hispanic Officer Arrests of Blacks and Hispanic Blacks						
Intervention Crossover	.058*	.028	.058*	.028	.095*	.039
Marijuana (MJ) Policy Change	--	--	-.125**	.044	-.085	.046
Intervention Crossover x MJ Policy Change	--	--	--	--	-.076	.044
Intercept	5.79**	.025	5.79**	.025	5.78**	.026
Hispanic Officer Arrests of Whites						
Intervention Crossover	-.077	.054	-.077	.054	-.005	.066
Marijuana (MJ) Policy Change	--	--	-.039	.0123	.001	.130
Intervention Crossover x MJ Policy Change	--	--	--	--	-.114	.074
Intercept	3.56**	.064	3.56**	.064	3.55**	.066

*p < .05; **p<.01