Written Testimony of Barry Friedman

**Jacob D. Fuchsberg Professor of Law and
Associated Professor of Politics;
Faculty Director, Policing Project
New York University School of Law**

Before the House Committee on the Judiciary
Subcommittee on Crime, Terrorism, and Homeland Security
Facial Recognition Technology: Examining Its Use by Law Enforcement
July 13, 2021

Chairwoman Jackson Lee, Chairman Nadler, Ranking Member Jordan, Ranking Member Biggs and Distinguished Members of the Subcommittee: Thank you for inviting me to testify on the topic of the use of facial recognition technology by law enforcement.

In my remarks today, I would like to make three overarching points:

(1) Facial recognition technology (FRT), especially in the hands of law enforcement, is an extremely powerful tool; and it can be powerful for good and for bad;
(2) The present lawless status quo around law enforcement's use of FRT is unacceptable: there is an urgent need for comprehensive, nuanced regulation that would allow obtaining the benefits of this technology for public safety while eliminating the harms, in particular harms to marginalized communities—and especially Black communities—which bear the brunt of FRT harms at present; and
(3) I want to highlight key considerations that should be included in any regulatory approach to FRT, including steps Congress can take right now to rein in the risks posed by law enforcement's ongoing use of this technology.

I. Background for testimony

I am the Jacob D. Fuchsberg Professor of Law and Affiliated Professor of Politics at New York University School of Law. For over thirty years I have taught a number of courses relevant to this hearing, including Constitutional Law, Criminal Procedure, and Democratic Policing. I also am the author of numerous publications, in both the scholarly and public realm, about regulating policing, including *Unwarranted: Policing Without Permission*.

Perhaps most germane, I also am the Faculty Director of the Policing Project at NYU Law School. Our mission is to "partner with communities and police to promote public safety through transparency, equity, and democratic engagement."[1] We conduct research, but also do work on the ground all over the country, both with policing agencies and the communities they serve, to promote democratically-accountable policing. Ours is an all-stakeholders approach. Everywhere we work, we endeavor to do so both with communities affected by policing, and with the police themselves. That is reflected in our Advisory Board, which surely is unique in including public officials, activists, policing leaders, and civil liberties and racial justice advocates, among others. By listening to, and working with, everyone we hope to move the needle toward greater public safety that is just, non-discriminatory, and effective. If you are interested in the full scope of our work, you can learn more at our website, www.policingproject.org.

One of our primary areas of study and work is the use of emerging technologies by law enforcement. And core to that is facial recognition. I think it is fair to say I have spent hours and hours in discussion about these issues with law enforcement, with members of impacted communities, with racial justice and civil liberties advocates, and with representatives of the technology companies that make these products. Always looking for a sensible way forward to what is a very real and pressing social problem.

[1] *Our Mission*, Policing Project, https://www.policingproject.org/our-mission

II. The need for "front-end accountability"

Hearings like this are at the core of the Policing Project's mission. To explain why that is, I would like to draw an important distinction between what we refer to as "front-end" and "back-end" accountability.

There has been a great deal of concern in the country over the last few years about the impact of policing. Some of that concern has been about uses of force and coercion, be it police shootings or pedestrian and traffic stops. But it also has been about the kinds of technology you seek to address here, like facial recognition. And when these issues are discussed, the word "accountability" often is used.

But there are two kinds of accountability and they are very different. Most of what we hear about in policing is "back-end accountability." The police have done something that people feel is wrong, and they want to assign responsibility and see that there is responsive action taken. Examples include proceedings in court to exclude evidence that is obtained unlawfully, or the prosecution of officers, federal investigations or civil rights suits, and the like. All these are aimed at accountability after-the-fact, after something has happened. Back-end accountability is essential, but because it only targets misconduct, there is a limit to what it can accomplish to guide policing before it goes awry.

What has been almost entirely missing from policing is accountability of a very, very different sort: front-end accountability. By that I mean to say that the public has a voice in setting transparent, ethical, and effective policing policies *before* the police or government act. The very work you as legislators do daily. Although we expect this in most areas of government, policing is a striking exception.

In my view, the lack of democratic front-end accountability explains many of the problems with policing today. Take facial recognition as an example. Policing agencies have acquired and used this technology in secretive ways, without adequate guardrails around use cases or means of use. People are distrustful when surveillance technologies are used without transparency and rules. This is only natural. And nowhere is the mistrust higher than in Black and brown and marginalized communities, which already feel the brunt of many unfortunate policing practices. The wise policing leaders I know understand they cannot fight crime or achieve public safety without community trust. And if there is going to be trust around surveillance technologies—especially in the communities understandably most distrustful of policing at present, there has to be great sensitivity to the issues that lead to such mistrust. That means complete transparency, and a regulatory approach that recognizes and addresses harms.

But frankly, it is asking too much of policing agencies to develop regulatory approaches to complex technologies on their own. That is the job of legislative bodies. It is your job. This hearing is a great example of front-end accountability in action, and I am grateful for the opportunity to participate.

III. There is an urgent need to regulate law enforcement use of FRT

Of all the new technologies used by policing agencies today, none has captured the public's attention as intently as facial recognition technology. This is not surprising – unlike some other modern crime-fighting tools such as cell-site simulators or even other biometrics such as DNA – its utility is intuitive and nightmare scenarios of government misuse are not hard to imagine. We don't even have to imagine them: one need only look to FRT's use for state-sanctioned mass surveillance and social control in China and Russia, including of racial minorities such as the Uyghurs.[2] Put simply, the power of this technology, and the threat it poses, is obvious and real.

At the same time, there is promise in the technology. That is why law enforcement has embraced it so quickly. It provides a way to identify people quickly including suspects in criminal offenses, or locating subjects of Amber and Silver alerts. Federal agencies have used FRT to identify individuals who entered the Capitol unlawfully on January 6, 2021.

Too often, the use of new technologies by police is debated as a matter of being "for" or "against" it. We become highly polarized. At the Policing Project, we believe the better approach is to figure out if society can benefit from a particular technology. Then, if there are real benefits to be had, the question becomes whether it is possible to minimize or eliminate any harm. And we have to pay special and close attention to where those harms fall. Too often, society sees a net benefit in some policing practice, and pays little attention to the fact that the harms fall on particular communities—especially BIPOC communities. That is flat out unacceptable. Addressing crime and achieving public safety only works if it is co-produced by police and the community, and that will not happen when the costs of policing fall disproportionately on some communities.

I often have heard proponents of this technology in the law enforcement community and beyond argue that FRT is just another tool in the police tool belt, or that it's just a more efficient way of doing the same thing police always do when trying to identify a suspect, victim or witness: i.e., sort through mugshots. This is not true. With FRT, police can (and do) search databases of millions of faces in a matter of seconds. AI-driven tools of this sort offer police the ability to sort, organize, and store unfathomably large quantities of our personal information, and to track people wherever they go. In short, as the Supreme Court and legal scholars have recognized, when it comes to surveillance technologies used by police, "digital is different."[3]

The problem we face is that today in too many jurisdictions the use of this technology is entirely unregulated, and with a lack of regulation comes harm. Although a few states recently have passed legislation addressing law enforcement use of this technology, the status quo for most of the country remains a regulatory vacuum.[4] This is untenable. Some agencies may use FRT wisely; some agencies are doing just the opposite. It is the Wild West out there, and you have the authority to address this problem. Until there is a real, comprehensive regulatory regime for FRT, there will be misuse, and mistrust. The technology simply should not be in use—other than perhaps controlled experiments from which we can learn—until a regulatory framework is in place.

There is no hiding from the fact that true regulation of FRT presents complicated issues. What we need is careful study and rational regulation. Indeed, it is imperative, because we should not kid ourselves: FRT is just one of many biometric technologies in the product development

[2] *See, e.g.*, Alfred Ng, *How China uses facial recognition to control human behavior*, CNET (Aug. 11, 2020), https://www.cnet.com/news/in-china-facial-recognition-public-shaming-and-control-go-hand-in-hand/; Robyn Dixon, *Russia's surveillance state still doesn't match China. But Putin is racing to catch up*, WASH. POST (Apr. 17, 2021), https://www.washingtonpost.com/world/europe/russia-facial-recognition-surveillance-navalny/2021/04/16/4b97dc80-8c0a-11eb-a33e-da28941cb9ac_story.html.

[3] *E.g.*, Andrew Guthrie Ferguson, *Facial Recognition and the Fourth Amendment*, 105 MINN. L. REV. 1105, 1128-29 (2021); Jennifer Stisa Granick, *SCOTUS & Cell Phone Searches: Digital is Different,* JUST SECURITY (June 25, 2014), https://www.justsecurity.org/12219/.

[4] *E.g.*, An Act Relative to Justice, Equity and Accountability in Law Enforcement in the Commonwealth, 2020 Mass. Acts ch. 253; An Act to Increase Privacy and Security by Prohibiting the Use of Facial Surveillance by Certain Government Employees and Officials, 2020 Me. Laws ch. 394; An Act Relating to the Use of Facial Recognition Services, 2020 Wash. Sess. Laws 1847.

pipeline. In the balance of my remarks, I will apply this benefit-cost approach to argue for nuanced regulation. As you will see, we are a long way off from that.

IV. Use cases

I want to distinguish among several very different use cases for FRT, as I think the analysis may differ across them. My remarks are directed at one particular use case: FRT for retrospective investigative identifications. By that I mean when law enforcement uses FRT to identify a person (or persons) by searching a still image—often called a "probe image"—against a database of other still images to return a list of candidates for comparison.[5] (I will refer to the database against which a probe image is searched as the "enrollment database.") I focus on the identification use case for two reasons: (1) it appears to be the most common way that law enforcement uses this technology at present; and (2) I believe it is the use case for which nuanced, comprehensive legislation is most apt.[6]

I distinguish this use case from two others – face verification and face surveillance. Face verification, the process of authenticating a person's identity by comparing two images, also generally is used by law enforcement, in particular by U.S. border enforcement agencies to authenticate travelers' identities.[7] This use case is not without risks – and indeed should be subject to many of the same guardrails that I detail below for face identification. Yet two conditions make me less nervous about verification if subject to proper controls: first, leading face verification algorithms generally are more accurate and less biased than leading face identification algorithms.[8] This is due at least in part to the fact that verification only requires comparing two images rather than images across an entire enrollment database, which is when the error rate gets compounded.[9] Second, because face verification does not require creating databases that store annotated images of our faces, it can raise fewer privacy concerns.

On the other hand, FRT for face surveillance – i.e., using real time or stored video footage to track people as they pass by public or private surveillance cameras, allowing their whereabouts to be traced – raises such serious privacy and liberty concerns, I continue to think the risks of allowing it outweigh the benefits. At the least I would not consider allowing it unless and until we can prove our ability to regulate and control the use of FRT for face identification. And we certainly have not done that.

V. Evaluating the benefits and costs of FRT for face identification

It is not difficult to articulate the (at least intended) benefits of this use case: the ability to solve serious crimes more quickly and efficiently. Proponents of this technology are quick to offer

[5] I use the term "identification" to reflect common parlance. In truth, a more precise way to conceive of the underlying process is that the algorithm actually is making a prediction about the likelihood that two images do or do not represent the same person.

[6] Kristin Finklea et al., *Federal Law Enforcement Use of Facial Recognition Technology*, CONG. RESEARCH SERV. 5 (Oct. 27, 2020), https://fas.org/sgp/crs/misc/R46586.pdf.

[7] *Id.* at 6.

[8] *Compare* NIST FRVT 1:1 Verification, Verification Performance, https://pages.nist.gov/frvt/html/frvt11.html *with* NIST FRVT 1:N Identification, Identification Performance, https://pages.nist.gov/frvt/html/frvt1N.html; *see also generally* Patrick Grother et al., Nat'l Inst. Stds. & Tech., *NISTIR 8280: Face Recognition Vendor Test (FRVT) Part 3: Demographic Effects* (2019).

[9] Handbook of Face Recognition, 644 (Stan Z. Li & Anil K. Jain, eds., 2d ed. 2011).

up anecdotes of FRT used to aid investigations of child kidnapping or sex trafficking. These examples are real and I take seriously the claim about the value of FRT to advance public safety. We all should. Police officers face a laborious and often fruitless task when they try to match photos of crime suspects to mug shots of people who have already been arrested. And I do not have to belabor with this audience the utility that face recognition can offer in investigations of a mass attack, such as occurred in your halls on January 6th. There is also the potential benefit that, properly structured, FRT offers an opportunity to make the eyewitness identification process more objective and auditable. According to the Innocence Project, mistaken eyewitness identifications account for 69% of wrongful convictions that have been overturned by DNA evidence.[10] FRT has the potential to help.

The problem with these benefits is we have little idea of their scope. We are left, as we are too often with policing, with nothing but anecdote. How many policing agencies are using FRT? How often has it been used successfully or unsuccessfully? How many identifications would have been made without FRT and at what expenditure of time? An economist would tell you that in doing benefit-cost analysis there is no need even to delve into costs until you are certain there are benefits. It simply is unacceptable that we plunge into the world of biometric face identification with no attempt whatsoever to know how valuable it is.

On the other hand, there are very real costs, especially social costs, to using the technology. I will review a few of them briefly; I have written about them extensively elsewhere, and others testifying today such as my friend Sakira Cook from the Leadership Conference will probe them at length.[11] First, there are significant accuracy and bias concerns that can have very serious racial impacts. This is not hypothetical. FRT is not now, nor will it ever be, perfect and many of even the best algorithms do not perform equally well across demographic groups. At present, many algorithms have higher error rates when attempting to identify individuals from certain marginalized groups, including people of color and women. Racial disparities are not unique to the algorithms; rather, they infect the entire process.[12] Disparities in arrest rates mean that the criminal databases that many agencies use for face recognition searches have a disproportionate number of Black, Latinx, and immigrant faces.[13] In addition, enforcement disparities exist across many levels of the criminal system – from the rates of traffic and pedestrian stops, to arrest and incarceration rates. In particular, studies continue to show that enforcement of low-level and drug offenses disproportionately targets Black Americans.[14] Comprehensive data on law enforcement use of FRT do not exist, but if its use tracks typical enforcement patterns, then communities of color likely are subject to more FRT searches. Indeed, before ending its program, San Diego found

[10] *Eyewitness Identification Reform*, Innocence Project, https://innocenceproject.org/eyewitness-identification-reform.

[11] Barry Friedman, Unwarranted: Policing Without Permission 29–233 (2017).

[12] Like anyone who studies issues of accountability and racial bias in AI systems, we all owe a debt to the trailblazing work of Joy Buolamwini, Timnit Gebru, and Inioluwa Deborah Raji, among others. Their research, including Gender Shades, Datasheets for Datasets and Model Cards for Model Reporting, has taught us a great deal about the biases in FRT technologies and what must be done about them. *See, e.g.*, Joy Buolamwini & Timnit Gebru, *Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification*, PMLR (2018); Timnit Gebru et al., *Datasheets for Datasets*, PMLR (2018); Margaret Mitchell et al., *Model Cards for Model Reporting*, ACM (2019).

[13] Jennifer Lynch, *Face Off: Law Enforcement Use of Face Recognition Technology*, EFF 10 (2020), https://www.eff.org/files/2020/04/20/face-off-report-2020_1.pdf.

[14] *See generally* Elizabeth Hinton et al., *An Unjust Burden: The Disparate Treatment of Black Americans in the Criminal Justice System*, VERA INSTITUTE (2018), https://www.vera.org/downloads/publications/for-the-record-unjust-burden-racial-disparities.pdf.

that police used FRT up to 2.5 more times on communities of color as compared to their presence in the general population.[15]

Second, FRT poses a very real threat to privacy and autonomy. It should take no lengthy discussion to establish this. Whether it's our faces ending up in FRT databases solely because we want to operate a vehicle or fears of cameras that can track and identify us, the risks to individual privacy and autonomy have been written about extensively. The warm embrace of FRT by authoritarian regimes as a means of social control should give us pause.

Third, the availability of this technology can chill First Amendment freedoms. Both historical and current law enforcement practice shows this hardly is hypothetical, whether one refers to the notorious COINTELPRO efforts of law enforcement agencies during the civil rights era, or the recent conduct of various South Florida police agencies that ran images from Black Lives Matter protests and a Juneteenth Block Party through a massive, unregulated face recognition database.[16] There is a persistent inclination of those in power to investigate and tamp down dissent. The Framers of our Constitution understood this all too well. And, unfortunately, these uses of surveillance tools also too frequently are aimed at marginalized communities.

Finally (although I am skipping over other harms), unlike almost every other form of forensic evidence, defendants rarely are informed that FRT was used in their case. This lack of disclosure could hamper the defense's ability to conduct a thorough investigation. For example, although it is standard to disclose other matches produced by traditional police lineups or photo arrays, this practice apparently does not apply to FRT results. And there remain open questions about how this lack of disclosure could affect a defendant's due process rights, such as whether some FRT results could constitute *Brady* evidence that must be disclosed to the defense.

With these costs in mind, it is easy to understand why many reasonable minds argue we simply must ban law enforcement use of FRT. I have wrestled with this conclusion myself. But ultimately, I still think it is possible to obtain the benefits of this technology while ensuring that we eliminate to the greatest degree possible the harms. And the way to do that is through sound and comprehensive regulation. This is not a revelatory concept – one need only look to the way Congress has approached the use of another invasive biometric in the criminal justice system—DNA—for regulatory inspiration.[17] Of course there are substantive and procedural differences between these two biometrics and their use by law enforcement, but the caution with which Congress and other legislative bodies has approached DNA stands in stark contrast to what has happened with FRT. I cannot say this enough times: technologies like FRT simply should not be used without a regulatory framework in place that imposes strict controls around their use and ensures that they serve communities, particularly historically marginalized communities and specifically Black communities. Anything else is simply irresponsible.

[15] Alex Holdor-Lee, *Privacy expert Clare Garvie explains why your face is already in a criminal lineup*, DOCUMENT JOURNAL (Oct. 12, 2020), https://www.documentjournal.com/2020/10/privacy-expert-clare-garvie-explains-why-your-face-is-already-in-a-criminal-lineup/.

[16] Joanne Cavanaugh Simpson & Marc Freeman, *South Florida police quietly ran facial recognition scans to identify peaceful protestors. Is that legal?*, SOUTH FLORIDA SUNSENTINEL (June 26, 2021), https://www.sun-sentinel.com/local/broward/fl-ne-facial-recognition-protests-20210626-7sll5uuaqfbeba32rndlv3xwxi-htmlstory.html

[17] Emily J. Hanson, *The Use of DNA by the Criminal Justice System and the Federal Role: Background, Current Law, and Grants*, CONG. RESEARCH SERV. (Jan. 29, 2021), https://fas.org/sgp/crs/misc/R41800.pdf.

V. Regulatory role for Congress

I want to turn now to what are some of the essential aspects of an adequate regulatory regime. Over the past many months, I've made it a point to have conversations with a diverse set of stakeholders, including technology vendors, law enforcement, civil rights and racial justice advocates and privacy experts to discuss the issue of law enforcement use of FRT. I owe the insight I bring to you today to these conversations. And although people I speak with disagree vehemently about such fundamental matters as whether FRT should be used at all, I find remarkable agreement about what is *not* okay – about what shoddy practices around the use of FRT look like, and what must not be permitted. I also have learned of holes in our knowledge about FRT that need to be filled, and ideally before it is used at any greater scale. These are the matters regulation must address before we think of using FRT responsibly.

The points I am about to make are not comprehensive.[18] The Policing Project hopes in the not-too-distant future to release a detailed enumeration of necessary provisions and steps to regulate FRT. But the points I make below show the level of nuance and attention I think is missing from the debate over FRT at present. And they reinforce the fact that many, many agencies are using FRT today in ways that are not coming close to meeting these most basic requirements.

A. Key considerations for regulation[19]

- Law enforcement use of FRT should be transparent and accountable – to the general public and criminal defendants alike

First and foremost, no government agency ever should be using biometric identification technologies such as face recognition, that were not appropriated for, and approved by, a democratically-responsible body. Unfortunately, too often, agencies acquire these technologies in other, less transparent and accountable ways.[20] All this does is breed mistrust of the agency and the technology. Nothing is lost by disclosure and democratic approval, and everything is gained.

Second, all agencies that use or access FRT should be required to draft and disclose use policies. Often the best way to maximize the benefits of a technology, while minimizing harms, is by setting clear rules on how the technology is used and disclosing the use policies to the public. Among other things, these policies should include provisions describing authorized uses and users, training requirements, privacy protections, internal oversight mechanisms, audit processes, and penalties for misuse. They should be attentive to the harms of the technology, especially privacy and racial harms. They also should identify which vendors and software programs are being used.

[18] Many of these ideas originate from collaboration with my colleague, Prof. Andrew Ferguson, from whom I've learned a great deal. *See* Barry Friedman & Andrew Guthrie Ferguson, *Here's a Way Forward on Facial Recognition*, N.Y. TIMES (Oct. 31, 2019), https://www.nytimes.com/2019/10/31/opinion/facial-recognition-regulation.html; *see also generally* Ferguson, *supra* note 3 (providing an in-depth analysis of how the Fourth Amendment intersects – and doesn't – with law enforcement use of FRT and suggestions for a regulatory framework).

[19] As I was putting the finishing touches on this written testimony, I discovered U.S. GAO's recent report on forensic technology, which includes a discussion of FRT use by law enforcement. U.S. Gen. Accounting Office, *GAO-21-435SP Forensic Technology* (July 6, 2021). From my brief review, it appears that many of their findings (e.g., lack of specialized training for human reviewers; algorithm versions tested by NIST are not necessarily the versions procured by agencies) and their policy recommendations (increased training, standards, and transparency) are in line with my own views as stated here.

[20] Matthew Guariglia & Dave Mass, *How Police Fund Surveillance Technology is Part of the Problem*, EFF (Sept. 23, 2020), https://www.eff.org/deeplinks/2020/09/how-police-fund-surveillance-technology-part-problem.

Third, we need careful data collection on, and disclosure of, FRT use. That is the only way we as a society will learn about best practices and the scope of any benefits. And such disclosure is of course essential to criminal defendants who are identified using FRT. For each use of FRT by law enforcement, the following information at a minimum should be collected and disclosed: which probe images are used and their quality, what matches were returned by the system, and investigative outcomes.

- Searches should not be limited to criminal databases

The choice of enrollment database—i.e., which repository of images that a probe image gets searched against—matters hugely when it comes to the impacts of law enforcement's FRT use. Some agencies search only criminal ("mugshot") databases; while others access statewide DMV databases. As I've written elsewhere, I believe that FRT should not be limited to criminal databases.[21] This may sound counterintuitive, but these databases are the product of decades of discriminatory policing for offenses like drug crimes; using those will continue us on this course. And if the goal is to identify suspects, witnesses, and victims, then using a wider dataset only makes more sense. To yield more equitable and effective results, the police should search databases that include all our faces. If we permit the use of this technology, we should all be in the pool.

- Accuracy and fairness of the technology

Accuracy and bias are headline news when it comes to FRT. And you've likely heard claims about accuracy – such as vendors that tout near perfect accuracy in percentage terms, such as 90% accurate. But it is meaningless to talk about accuracy percentages. What matters with technologies like FRT is false positives (people identified as matches who are not) and false negatives (when there is a match in the database that is not returned). Every FRT system has to make tradeoffs between these; no system can get it exactly right. If we are looking for a lost child, we want fewer false negatives; if we are trying not to turn innocent people into suspects, we want fewer false positives. And how well FRT performs in terms of false positives and false negatives is a result of a number of factors that are not getting enough attention.

To effectively evaluate false positive and false negatives rates, we need independent testing of FRT systems under real world conditions, on the right size databases, and on the types of images law enforcement actually searches. This testing also needs to report error rates by demographic group and evaluate the impact of human review. None of this is happening today – at least not in any way that's publicly reviewable.

The National Institute of Standards and Technology (NIST) currently runs the leading technical benchmark for FRT algorithms and its work in this area has been essential for improving the technology's accuracy and providing public information about these algorithms. But NIST's testing program, for all its value, tells us little about actual law enforcement use contexts:

(1) NIST doesn't test the algorithms that law enforcement actually uses. Rather, it tests whatever a vendor submits, which, depending on which vendor you ask, may be an algorithm that's similar to what it sells to law enforcement, or it may not be at all.

[21] Friedman & Ferguson, *supra* note 18.

(2) NIST's testing databases are too small. *The literature makes clear that enrollment database size affects algorithmic accuracy and demographic bias*.[22] As the Government Accountability Office's recent report revealed, NIST's largest testing database is way smaller than many law enforcement databases (in one case, 69 times smaller).[23] We can't know how well the technology performs in many law enforcement use contexts without evaluating them on larger databases.

(3) NIST hasn't evaluated demographic performance on the types of images most commonly used by law enforcement, e.g., surveillance camera images. Making matters worse, these images typically are lower quality than images like mugshots and NIST's testing shows that error rates rise exponentially on lower quality photos. Without demographic breakdowns on operationally representative images, we have no idea how these algorithms perform for different subpopulations when deployed; and

(4) Finally, NIST's testing program only evaluates algorithms; it doesn't evaluate the impact that human reviewers have on accuracy and bias. So its tests don't tell us anything about full system performance.

These deficiencies in NIST's testing are a huge problem because when it comes to publicly available independent, expert testing, NIST is the only game in town.

To solve this problem, Congress should give NIST the necessary resources to build a better technical benchmark, i.e., one that mirrors law enforcement use contexts. Congress also needs to invest in research for developing full-system operational testing that can evaluate both quantitative metrics (like error rates, broken out by demographics) and more qualitative metrics (like fairness and privacy).

- Human-in-the-loop

Evaluating and validating the tech itself is a necessary condition for regulating FRT, but it is not sufficient. We need to get the human part of this analysis right too.

You've probably heard the term "human-in-the-loop" offered up as a meaningful safeguard: we are told not to worry about FRT accuracy and bias, because in law enforcement situations a human checks the machine's outputs to confirm them. But once again, how this phrase is used is too simplistic and overlooks real issues. It ignores simple facts like that people tend to believe too deeply in, and be persuaded by, what technology tells them. In order to make the human-algorithm pairing work well, several things need to be done.

Human examiners should receive specialized training on both face comparison and the software.

Agencies should require that any officer who reviews FRT results is trained in the latest science on face comparison. Research has shown that forensically trained face examiners are better at recognizing faces than untrained individuals.[24] Although consensus standards for face

[22] *E.g.*, Li & Jain *supra* note 9 at 644.

[23] U.S. Gen. Accounting Office, *GAO-21-518, Facial Recognition Technology: Federal Law Enforcement Agencies Should Better Assess Privacy and Other Risks* 16, Fig. 3 (2021).

[24] P. Jonathan Phillips et al., *Face recognition accuracy of forensic examiners, superrecognizers, and face recognition algorithms*, PNAS (2018), https://www.pnas.org/content/115/24/6171.

comparison still are under development, there are available resources that agencies can draw on today to train their officers.[25] In addition, examiners should receive vendor-supported training on the technical requirements and specifications of the software so they understand the ins-and-outs of how these systems function and how they can fail. This means explanations on things like how image quality and variations in an individual's appearance (glasses; bangs; hats) and other environmental factors can affect error rates. It also should include comprehensive training on the software's logging capabilities to ensure audit trails are created.

Research is needed to set standards on human-algorithm interaction.

We should treat face recognition like we do other forensics. For example, in criminal investigations, the best practice for witness identification of suspects is to "blind" the process. The law enforcement individual conducting a photo array does not know which of the photos is the suspect, and that way cannot unintentionally signal correct answers to the witness. We need to think about processes to "blind" the human in the loop so that in evaluating results from the algorithm, they are able to conduct an independent evaluation rather than over-trust the machine output.

Unfortunately, we don't know enough about the best practices in this area. Understanding how humans and machines work best together is an area of ongoing research. Some research shows that face recognition is most accurate when humans and machines work together.[26] On the other hand, automation bias (i.e., humans over-rely on machine output) is a well-studied phenomenon.[27] Congress should spearhead the development of consensus standards and best practices for the human component of FRT analysis and develop research-backed processes and protocols for conducting these examinations in ways that maximize accuracy and minimize bias.[28]

In the meantime, however, there are some best practices we can implement right now: namely, agencies only should allow trained officers who are uninvolved in the underlying investigation to review FRT results. In addition, FRT results should be subject to multiple levels of independent human review before any action is taken, to reduce the impact of automation bias.

Develop a gold standard process for investigative corroboration.

While we're on the lookout out for potentially meaningless terms, we also should view with skepticism any assurances that face recognition results are "not considered positive identifications" and that "further investigation is required."[29] Just as there are no consensus

[25] *See, e.g.*, *Facial Identification Subcommittee*, https://www.nist.gov/osac/facial-identification-subcommittee (listing proposed and passed standards for face comparison); *Facial Identification Scientific Working Group*, https://www.fiswg.org/documents.html (providing training, comparison and evaluation resources).

[26] Philips et al., *supra* note 21.

[27] Daniel E. Ho et al., *Evaluating Facial Recognition Technology: A Protocol for Performance Assessment in New Domains*, Stanford Instit. for Human-Centered AI 14 (Nov. 2020), https://hai.stanford.edu/sites/default/files/2020-11/HAI_FacialRecognitionWhitePaper.pdf; *see also* John J. Howard, et al., *Human-algorithm teaming in face recognition: How algorithm outcomes cognitively bias human decision-making*, PLOS ONE (2020), https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0237855.

[28] I acknowledge that some of this work is underway, see, e.g., entities like OSAC's Facial Identification Subcommittee. *E.g.*, *Facial Identification Subcommittee*, https://www.nist.gov/osac/facial-identification-subcommittee. My remarks here are directed at emphasizing the need to support—not supplant—this type of work.

[29] *E.g.*, Det. Police Dep't, *Directive No. 307.5 Facial Recognition* § 307.5-5.4, https://detroitmi.gov/sites/detroitmi.localhost/files/2019-

standards for human review of FRT results, there is no common best practice for what constitutes sufficient investigative follow up to confirm an FRT lead. I have no doubt that many agencies and individual officers take this requirement to heart and engage in thorough, independent investigative efforts after receiving an FRT lead. Still, Mr. Robert Williams' story, whose testimony you also will hear today, provides a painfully real reminder that others do not.

Federal agencies like the Federal Bureau of Investigation and the Department of Homeland Security are avid users of this technology for investigative identifications and likely have significant expertise in this area. There is both an opportunity and a need for federal agencies, in conjunction with defense and civil rights and racial justice experts, to develop gold standard processes for the investigative follow up that should be required to corroborate an FRT lead that then could serve as a model for local agencies.

- Law enforcement use of FRT for face identification should be limited and legitimate

Considering the various and serious harms that can result from law enforcement use of FRT, agencies need to have good reason to deploy this tool. To my mind, using FRT to investigate serious crimes, like murder, rape and child kidnapping, is a good reason. Using FRT to investigate petty crime or to search for unlawful immigrants, is not. Overcriminalization is an epidemic in our country and one that disproportionately affects people of color, particularly in Black and brown communities. Unless we limit face identification to serious offenses, we will exacerbate this problem.

We also need to have proof of officers' "predicate" for using the technology at all. In the past, I have advocated for a warrant requirement to ensure face identification is used only as permitted. I still believe this approach makes sense - without a requirement for neutral supervision, there simply is too much potential for abuse. (This belief also is reflected in Maine's recent bipartisan legislation regulating law enforcement use of FRT – one of the few states that has managed to take action on this issue). Before FRT is used, a judge should be persuaded that the probe photo represents a real suspect for an enumerated offense.

- Regulate Vendors

Congress unquestionably has power to regulate the vendors of FRT products and it should. The vendors of high-risk technology products like FRT bear responsibility for ensuring their products work as promised and do not perpetuate harmful bias. Regulators in Europe have recognized this; the European Commission's recent regulatory framework proposal on AI sets forth a series of pre-market conformity requirements that vendors must meet before they can sell "high-risk" AI systems.[30] In a similar vein, NIST recently issued a draft report on managing AI bias that includes a detailed set of recommendations directly targeting AI developers. Vendors have a role to play in ensuring the ethical operation of their systems – and Congress should require that they do so.

09/Revised%20facial%20recognition%20directive%20transmitted%20to%20Board%209-12-2019.pdf; *NYPD Questions and Answers Facial Recognition*, https://www1.nyc.gov/site/nypd/about/about-nypd/equipment-tech/facial-recognition.page.

[30] *Regulatory framework proposal on Artificial Intelligence, European Commission*, https://digital strategy.ec.europa.eu/en/policies/regulatory-framework-ai.

To give some examples, Congress should require that vendors:

⇒ Train their algorithms on datasets that are designed to mitigate harmful bias. This means building datasets that are sufficiently diverse and representative of the population in which it will be deployed and that don't perpetuate historical biases.

⇒ Consider factors other than raw accuracy when selecting which models to develop – in particular, vendors should have to consider accuracy for the relevant subpopulations that stand to be impacted by their technology.

⇒ Ensure their FRT products are self-auditing, i.e., that they are built with sufficient capabilities to log user actions and trace results, so that law enforcement agencies can audit themselves and make public disclosure.

⇒ Provide documentation of all internal testing and error rate reporting for their systems.

⇒ Provide comprehensive user training on the technical requirements and specifications of their products.

⇒ Set criteria around things like image quality specifications required to ensure system accuracy and ensure that users cannot submit or modify images that don't meet these specifications.

- Facilitate fact-finding to guide development of best practices

Finally, as I noted above, the discourse around the benefits and harms of FRT suffers from a massive information gap. We don't know how many agencies are using this technology; we don't know how many cases it helped solve; or how often its use has led to misidentifications. Congress can help fill these knowledge gaps by providing federal funding for the local data gathering that is necessary to evaluate the use and impact of this tool.

Conclusion

As I said, this is not a complete list. But I hope I have convinced you of two things:

First, there is a crying need for regulation of FRT technology, now. No matter where you are on the pro-con debate over FRT, you must recognize—as so many with whom I have spoken do—that permitting its use without regulation of some of these key elements is risky, unacceptable, certain to lead to racial bias, and undemocratic.

Second, the proper use of FRT is a complicated matter, and it requires regulation that matches that complexity.

I want to thank you for the opportunity to testify today. The matter you are considering is extremely consequential. We all should want optimal outcomes and rational solutions.

We would of course be willing to provide any other information that could be of use.