

# JUDICATURE

Published by the Duke Law Center for Judicial Studies.  
Reprinted with permission. © 2017 Duke University School of Law.  
All rights reserved. [www.law.duke.edu/judicature](http://www.law.duke.edu/judicature)

## can science save justice?

**Psychologists and neuroscientists say cognitive forces beyond our control are negatively affecting the criminal justice system and thwarting the promise of justice for all. Incremental efforts to address these damaging biases often come up short.**

**Is it time for bold reform?**

BY ADAM BENFORADO

### ***“Know thyself.”***

Inscribed on the Temple of Apollo at Delphi, and echoed down the halls of time by Plato, Pope, Franklin, and Emerson, there may be no more fundamental maxim to describe the human project. It has been our work — our puzzle — for as long as we have been on this earth. And, though progress can be hard to discern in the cacophony of modern life, the truth is that we are now in an unrivaled position to answer the call of history.

In the last few decades, advances in the mind sciences, data collection, and experimental design have greatly increased our understanding of human behavior. Yet, these near-miraculous developments have not engendered the collective epiphany that might be expected. Why? Much of what psychologists and neuroscientists

have discovered about us is deeply unsettling. As Aristophanes warned in *The Clouds*, to really “know yourself” is to know “how ignorant and stupid you are.”<sup>1</sup>

We are not the people we have long assumed ourselves to be. While we feel like rational deciders, directing our actions through reason and deliberation, we are frequently guided by intuitive processes beyond our awareness or control. We quickly spot flaws in others, while ignoring the same shortcomings in ourselves. Our memories — highly malleable and subject to misattribution and suggestibility — are more like ever-changing collages than photographs. Rather than objectively sorting through the facts, we jump to conclusions and then search for evidence that confirms what we already believe to be true. To make matters worse, many of our rules, policies,

and institutions are based on myths about how people think and act. The upshot is that we are far less safe, less fair, less in control, less effective, and less just than we purport to be.

Our legal system is not immune from the critique. The latest scientific research suggests that the great edifice of law is grounded on incorrect and damaging notions about human cognition that have gone uncontested for centuries. We have assured ourselves that when an eyewitness comes forward and identifies a suspect with confidence, we can rest easy that we have the right guy. But in real line-ups, one third of the time when witnesses pick someone out they select an innocent filler.<sup>2</sup> We have assumed that when a suspect confesses to the rape and murder of a child, we can close the case. But, again, the emerging data has shaken our



foundations. Twenty-eight percent of the known American DNA exoneration cases involved a false confession.<sup>3</sup> And fault appears to lie with the leading method of interrogation, which advises detectives to rely on dubious demeanor cues for identifying deception and encourages the police to utilize maximization and minimization techniques shown in experiments to lead people to admit to offenses they did not commit.<sup>4</sup>

Consider what we've learned about judges. Conventional wisdom holds that there are two kinds of judges: umpires and activists. Umpires are the ones who choose to be objective: They call "balls and strikes," dispassionately applying the law to the facts at hand. Activists, by contrast, are ideologues who choose to pursue their own agendas. In this conception — engrained in many judicial codes of

conduct — bias is subject to introspection and is simply a matter of self-control. But the latest research suggests that all judges — just like all humans — are biased.

Part of the issue is that judges aren't able to put their backgrounds to the side when they sit on the bench. Justice Sonia Sotomayor's controversial claim that "[p]ersonal experiences affect the facts that judges choose to see" is strongly supported by available psychological data.<sup>5</sup> We all look at the world through lenses colored by our identities and histories. It's not surprising to learn that judges appointed by Democrats tend to decide matters in ways more favorable toward minorities, immigrants, and convicts and uphold agency decisions brought by labor unions and public-interest groups.<sup>6</sup> As you might expect, those appointed by Republicans tend to favor business and the prosecu-

tion. But judges' lenses are tinted by far more than ideology: Age, gender, race, and countless other factors all cast their shadow. In one recent study, judges who had a daughter rather than a son were 16 percent more likely to decide in favor of women's rights in civil rights cases involving issues of gender.<sup>7</sup>

It is worrisome to consider not only that our cultural backgrounds matter, but also that our judiciary has lacked diversity for so long. For hundreds of years, the Anglo-American legal system amassed judicial opinions defined by the viewpoints of a narrow subgroup of citizens — white, elite, older, Christian, educated men.<sup>8</sup> Today, Justice Sotomayor may bring a fresh outlook to the Supreme Court as the first Hispanic in the Court's 220-year history, but she is powerfully constrained by this precedent. ▶

Unfortunately, the problem of judicial bias extends beyond a judge's identity: Court decisions appear to be shaped by a whole host of situational elements that aren't supposed to matter. Studies on sentencing have shown that judges are influenced by the race of the defendant, legally prejudicial and irrelevant facts, subtle reminders of their own mortality, random numerical anchors generated by rolling a pair of dice, and even the time of day.<sup>9</sup> When researchers looked at a set of Israeli judges, they found that the judges were far more likely to grant prisoners parole at the start of the morning or after one of two breaks than they were at the end of the day or just before a break.<sup>10</sup> A separate study published last year showed that the shift to daylight savings time itself matters: On the Monday after losing an hour of sleep, judges hand out harsher punishments.<sup>11</sup>

Of course, judges don't feel biased at all because things like confirmation bias are hidden behind the elaborate legal frameworks we've constructed: canons of statutory interpretation, three-prong tests, amicus briefs, and the like. A judge searching for the answer to whether fleeing from the police in a car is a "violent felony" may believe she is engaging in a neutral assessment of the evidence to reach a conclusion. But the data suggests otherwise: Judges tend to start with the conclusion based on gut instinct and then search for the research that supports it, discarding and dismissing counter-evidence along the way.<sup>12</sup>

Judges are not uniquely vulnerable to bias. In my book, *Unfair: The New Science of Criminal Injustice*, I show how all our legal actors — including jurors, witnesses, defendants, experts, lawmakers, and prisons guards — are affected by unappreciated forces around them and within them. The skew in our system runs far and wide.

There can be no doubt, then, that empiricism presents a major threat. To embrace evidence is to show that the emperor is wearing no clothes. It undermines the legitimacy of our existing structures, and we would expect to see a backlash to research-derived policy

across disciplines, including law. In many ways, this is the central battle of our time: between societal ordering based on evidence and ordering based on opinion, faith, and "feel."

There are genuine reasons to be cautious with evidence-based approaches to policy. Research can turn out to be flawed, for one. But in the long run, embracing empiricism bests all the alternatives because it is grounded in testing and updating. When the reliability of an existing protocol for eyewitness lineups is undermined by subsequent nonreplications and real-world datasets, there is a ready solution: Revise the protocol to conform to the latest research. Progress happens by design with an evidence-based approach, not luck. Failures are expected events to be learned from, not embarrassments to be hidden from view.

Moreover, it's worth noting that those who argue against evidence-based change on the grounds that the underlying scientific evidence is not yet ripe are often arguing for a status quo that is based on *no* science. So, to the critics who suggest that the research on false confessions or racial bias is not robust enough to merit changes: Where are the peer-reviewed studies to support the accuracy of current approaches? There is a long history of trying to ensure inaction on a variety of issues by attacking existing findings as "unsettled."<sup>13</sup> The cigarette industry managed to avoid regulation for years by casting doubt on the scientific data that suggested smoking causes cancer, just as the fossil fuel industry has sought to disrupt the scientific consensus that man-made climate change is real. The trick is that you don't have to win the debate, all you have to do is make it look like one exists and you can count on Americans' natural inertia to ensure that the status quo is maintained.

### **WHY IGNORING EMPIRICISM ISN'T AN OPTION**

In the face of such significant challenges, we must be resolute. Ignoring empiricism is simply not an option for the law. The legitimacy of our legal system is grounded

in its strict adherence to the facts. A system that disregards the evidence is not a system based on the rule of law at all. Justice is predicated on truth. It matters, for example, whether African Americans actually get longer sentences than white Americans, regardless of what people believe to be the case. It matters whether three-strikes laws deter effectively. It matters if existing copyright protections don't have much of an effect on encouraging people to create works. We can build beautiful models based on assumptions about how rational individuals should behave, but what really matters is if people actually buy insurance when it is in their best interests or breach mortgage contracts when they're underwater.

Even if it were justifiable to ignore the behavioral evidence, hiding our heads in the sand would be foolish because sophisticated parties are already cataloguing our biases and using them to their advantage. The trial consulting industry was started by social scientists committed to leveraging their understanding of human behavior and the tools of their trade to ensure fairness in the selection of juries. But in the intervening decades, trial consulting has turned into big business, and the game has changed.<sup>14</sup> Clients pay thousands of dollars to win, not play fair, and the goal has shifted to studying the proclivities of jurors, witnesses, and others to ensure as skewed an outcome in the client's favor as possible. New firms have recently emerged that specialize in predicting judicial behavior.<sup>15</sup> Bloomberg Law Litigation Analytics, for example, sorts through legal data to offer individualized analysis on questions like the probability that a specific judge will grant a motion to dismiss, how well specific firms fare before that judge, and how often firms are successful in appealing.<sup>16</sup> Though the concern was once with snake-oil salesmen offering dubious data to desperate clients, the major danger today is that trial consultants are likely to become more and more effective. Members of the industry are not primarily lawyers; they are scientists, many with PhDs.<sup>17</sup> The evidence revolution is coming one way or another.

**Those who argue against evidence-based change on the grounds that the underlying scientific evidence is not yet ripe are often arguing for a status quo that is based on no science.**

The only reasonable answer, then, is to commit ourselves to evidence-based justice. First, we must continue to collect and analyze data. That means expanding our capture of real-world information. There can be a natural reluctance to expose one's work to quantification and assessment, but it is the only way to recognize hidden partiality and unfairness. Though a judge may handle hundreds or thousands of cases in a given year, the work of a judge is generally focused narrowly on the individual matter at hand, and it can be hard to see broader patterns. We must also increase funding for experimental work and foster collaboration between researchers and police departments, courts, and prisons. We learn the most when we can run experiments with real jurors, real judges, and real police officers, in the field.

Second, having collected data, we must adopt empirically-grounded best practices. Our goal should be to have all aspects of our legal machinery — how we train police cadets to handle the mentally ill, how we conduct *voir dire*, how we interview witnesses — backed by strong evidence. By its very nature, this will be an iterative process as we continue to collect data and hone our processes.

If embracing evidence-based justice is our path forward, there is still the question of just how bold we should be in our reforms. Most people who support empiricism have come down in the “incremental change” camp. For these innovators, the aim has been to figure out what is plausible in the current climate. What proposals can bring together broad coalitions? What ideas seem generally unthreatening? What can be done without much disruption? The result has been largely accepting the foundational legal myths as they are and looking to what minor changes are feasible today. So, we have seen calls for more oversight of plea bargains, allowing

jurors to take notes during trial, requiring open file discovery, and video recording all interrogations, among other reforms.<sup>18</sup> The work with eyewitness identification procedures is a prime example of the incremental approach. Over a period of decades, we have seen small-scale changes in certain jurisdictions based on research on blind administration, lineup construction, and other subtopics.

One of the problems with incrementalism is that it may make our core legal myths harder to dispel. Focusing on getting police departments to use a more diagnostic photo-array procedure with eyewitnesses ends up reifying the notion that eyewitness memory is sufficiently reliable to act as a primary guide for detectives, judges, and jurors. Incrementalism may also lead us to settle on rules that amount to political compromises and aren't defensible based on science. The push to end solitary confinement for juveniles is a great example. The research on solitary confinement suggests that it can produce deep psychological suffering and damage.<sup>19</sup> Young people appear to be particularly vulnerable. Yet, there are other populations — most notably, those with existing mental illness — who may be at even greater risk. And there is no scientific basis for using the age 18 as a break point. Someone who is one day short of his 18<sup>th</sup> birthday appears indistinguishable from someone a day older. Most importantly, even if certain groups may fare worse in solitary than others, the evidence suggests that such confinement presents a significant harm for almost everyone. In other words, the science implies a broad ban, but incrementalism says “we can't.”

So, while many of the changes being advanced by incrementalists — including banning solitary confinement for juveniles — are valuable and should be pursued, they can only be part of the solution. In

some cases, they may not get us where we ultimately want to go. In other cases, the pace of change is simply too slow and the injustice in the present is simply too great. We must concurrently engage in a broader reimagining and reformulating of our practices, procedures, and laws.

What if we had the ability to start anew? How would we build our legal structures if we began with a correct understanding of human behavior, and we were not constrained by the existing frameworks? We need to focus our attention on this ideal, envisioning its architecture, however infeasible it may seem in the present. By anticipating a bold evidence-based future now, we can better design and control outcomes in a rapidly changing environment.

### **BOLD IDEAS**

In the United States — and in many countries around the world — to be legally objective is to place a “veil” between yourself and those you judge. Lady Justice is carved with her eyes covered by a shroud. In the words of William Penn, “Justice is justly represented blind, because she sees no difference in the parties concerned. She has but one scale and weight, for rich and poor, great and small.”<sup>20</sup>

But real justice is not blind at all. Your appearance matters at every step of the way. With the recent deaths of Philando Castile, Alton Sterling, Michael Brown, Eric Garner, and other African Americans at the hands of the police, public attention has been particularly drawn to race. Black people are more likely to be stopped by the cops and more likely to have force used against them.<sup>21</sup> Castile, for example, was pulled over at least 49 times in the 13 years before his fatal encounter.<sup>22</sup> African Americans also receive harsher treatment when it comes to charging, bail, and sentencing.<sup>23</sup> In one study of interracial murder, black men with the most stereo- ▶

**The guiding principle is simple:  
If a factor like race isn't relevant to determining  
guilt, jurors shouldn't know the defendant's race.**

typically African-American features were more than twice as likely to receive a capital sentence as those with lighter skin and thinner lips.<sup>24</sup>

But our eyes deceive us even when race is placed to the side. Research shows that people are more likely to believe more attractive witnesses.<sup>25</sup> Skinny female defendants receive more lenient treatment than obese women.<sup>26</sup> People who avert their gaze or fidget end up tagged as deceitful by both police interrogators and jurors, even though research shows that neither “tell” is accurate.<sup>27</sup>

When Brock Turner, the Stanford swimmer convicted of sexually assaulting an unconscious woman behind a dumpster, received a sentence of just six months in prison last year, the district attorney lamented that the “punishment does not fit the crime.”<sup>28</sup> But it was no surprise: As the prosecutor conceded in her closing argument, Turner — with his broad smile, bright eyes, and pale, boyish face — doesn't “look like a rapist.”<sup>29</sup>

The legal system's primary defense against these biases has been to instruct judges, jurors, and others simply to turn them off. But the latest psychological research suggests that much of the skew is not susceptible to conscious control. There is no magic switch to erase a lifetime of exposure to damaging stereotypes that link the concepts of blackness and violence, or beauty and honesty.

Rather than telling people to ignore what they see, psychologists and legal scholars have begun to consider a number of promising interventions. Some incrementalists have pushed for educating judges, police officers, and prosecutors on their implicit biases. Others have suggested better constraining the discretion of our legal actors. Perhaps most encouraging have been efforts to make the metaphorical “blindfold” into a real one.

Other fields have shown what's possi-

ble. For a long time, orchestras were dominated by men. That changed when major philharmonics started auditioning musicians behind screens. Once this practice was put in place, the proportion of women among new hires doubled quickly.<sup>30</sup> In an attempt to address diversity problems in business, the Silicon Valley software company GapJumpers began screening job applicants so that employers could see only the individual data that was relevant to the position. According to the company, replacing traditional résumé review has boosted the proportion of people of color, women, non-Ivy League students, and those with disabilities landing a first-round interview from roughly 20 percent to more than 60 percent.<sup>31</sup> More broadly, one of the reasons that medicine has made such leaps in safety and effectiveness in recent decades is that it is now settled practice that neither the administrators nor the participants know the critical aspects of the research study. You can't be biased when you have no opportunity to be.

Applying these insights to the realm of criminal justice, a number of urgent reforms jump out. Prosecutors should not be told the race of the defendant when they are deciding the terms to offer in a plea bargain, because implicit processes lead to less favorable treatment of African Americans. Crime-lab analysts should be blind to all details about a case when they are assessing a forensic sample, because studies show that knowing a fingerprint was provided by someone who has already confessed or was picked out of a lineup significantly increases their likelihood of finding a match to the perpetrator.<sup>32</sup> And we certainly should not instruct jurors — as judges often do — that in assessing witness credibility they should focus on demeanor.

All of this should be standard practice across the United States. But what if we

were bolder and tried to blind the whole process of adjudication?

**TRULY BLIND JUSTICE**

Consider a future trial that takes place entirely virtually. Jurors watch proceedings remotely, with each participant represented by a neutral avatar designed to remove variations in attractiveness, body shape, skin color, mannerisms, and voice inflection. The guiding principle is simple: If a factor like race isn't relevant to determining guilt, jurors shouldn't know the defendant's race. The goal is to remove the typical shortcuts that allow jurors to jump to conclusions and provide decision-makers with only the information that is diagnostic to the task at hand.

The benefits of the “blinding” technology would extend to everyone in the courtroom. Today, an attorney can easily strike a juror based on race or gender as long as he comes up with a nondiscriminatory explanation for the exclusion.<sup>33</sup> (“She looked bored chewing her gum.”) But when a prosecutor doesn't know whether a juror is black or white, male or female, he can't discriminate.

With proceedings moving to the virtual arena, it becomes possible to record and screen everything before it comes before the jury. And this could help us address another type of bias. Currently, when presented evidence is subsequently deemed inadmissible, jurors are told to forget what they just saw or heard. But in experiments, mock jurors can't do that — and neither can many sitting judges. In one seminal study of a civil suit, three times the number of jurors found the defendant liable when exposed to evidence that they were later told to disregard as when they were not exposed to the evidence at all.<sup>34</sup>

In the criminal context, once you've learned that the defendant was previously accused of an assault, you can't help but

have that bit of information affect your current determination of whether he's guilty of a completely unrelated murder. With prerecording, the jury never learns of the impermissible material. Advance screening also allows us to cut out the dead time. Culling bench colloquies, overruled motions, swearing-in witnesses, administrative delays, and the like, we can greatly streamline trials, potentially reducing juror fatigue and improving attention.

To create a truly immersive experience, all courtroom participants could use virtual-reality headsets, which would provide everyone with the sense of being together in a real courtroom without anyone needing to leave her house or office to participate. This would not only make trials safer — reducing the possibility of violence — but also address the psychological strain entailed in providing in-court testimony or sitting on a jury. A rape victim would no longer have to be in the same room as her attacker. Plus, parties would no longer be burdened with paying travel expenses for experts and other witnesses, and courts would reap serious cost and scheduling benefits from reduced physical space requirements.

Just as important, remote access could radically increase the diversity of our juries and help achieve our goal of broad public engagement. Our current approach means that many Americans never serve as jurors because their lives don't allow for it. What if one day, a person could choose when to participate, fitting her service into her existing schedule of work, child care, and other responsibilities? With a much larger percentage of the populace in the mix, we might be able to greatly increase the number of jurors participating in each case to address the problem of outlier juries and ensure a truly representative cross-section of the population that reflected the particular demographics of the community.

That would require reimagining jury deliberations (perhaps drawing upon successful models of online collaboration in other fields). But the ultimate result could be vastly improved accuracy and consistency. Every defendant would face a true

jury of his peers. Every defendant would be tried in the exact same room by people who looked, sounded, and acted the same as in every other trial. And we'd be in a far better position to discover and address currently unknown biases — or ones introduced by the new format itself — because researchers could conduct experiments using the same virtual template and equipment used in real trials.

All of this might seem like science fiction, but is it? The basic technology for a virtual trial already exists. The first virtual-reality headset available to the public, the Oculus Rift, went on sale in 2016, and industry experts predict a rapid rise in applications and offerings in the coming decade. With recent advances, being physically present is no longer required in many high-stakes interactions. A surgeon can conduct a hernia operation on a patient hundreds of miles away. A soldier can target a pickup truck speeding across the Afghan desert while seated on another continent. And scientists and engineers around the world regularly cooperate on complex projects without ever meeting in person.

Our legal system is naturally more reluctant to embrace change, but much of the precedent for the virtual trial has already been established. We allow translators when a witness doesn't speak English. We cover up a defendant's swastika tattoo to avoid prejudice at trial. We permit digital re-creations of events to be played for jurors. In certain circumstances, we even allow people to testify through video-conferencing technology, as with child sex-abuse victims and witnesses who are too ill to travel.

That said, the virtual trial does raise certain questions. Does prerecording proceedings violate hearsay rules? When a defendant views an avatar rather than a real human, does that infringe his Sixth Amendment right "to be confronted with the witnesses against him"? Vigorous confrontation, of course, is still preserved in the form of cross-examination. What's missing is the ability of everyone to observe the witness's actual demeanor. In light of the scientific evidence, that seems

like a dubious foundation, particularly as the Sixth Amendment also ensures the right to an impartial jury. Leaving the system as it is means accepting pervasive violation of that arguably more fundamental clause.

Still, the prospect of virtual trials may make many uneasy, even if the legal hurdles can be overcome. Might we be losing something essential in moving to avatars? We've always had the ability to do trials with sworn affidavits — written testimony — but we've chosen to bring people together in a room to hash things out. Our approach rests upon incorrect myths about the value of such in-person interactions, but reducing the human element at trial may be hard for most people to accept. We have long believed in the common sense of the layman juror and the gut instincts of the judge, and we have fought hard to unfetter them. But the confidence in our human faculties has led us astray, and it's worth considering whether we might go one step further and get rid of the human element altogether.

### TRIAL BY ALGORITHM

Once we have the technological capacity, why not replace trials entirely? All of the testimony and evidence could simply be entered into a program that would analyze the accuracy and importance of each element and apply the relevant law. In the future, computers will be able to administer eyewitness identifications, assess forensic evidence, and consider the plausibility of alibis all without any human direction. Everything could be assigned precisely the weight warranted by existing scientific research.

Such a program could be specifically designed to avoid known human biases and to focus on what we actually think matters, disregarding the cues we've deemed irrelevant but can't seem to ignore. Every judge knows that jurors struggle to remember facts, understand expert testimony, and follow jury instructions. Indeed, in studies, jurors regularly pass over legal definitions with respect to things like rape and insanity in favor of their own lay definitions.<sup>35</sup> A computer, ▶

though, is capable of perfect recall of facts and law. Its understanding of a term like “mistaken belief in consent” can incorporate every case and law review article ever written on the subject.

That said, a computer can take into account whatever we’d like it to consider. Interesting new research suggests that people are reluctant to hand over moral decision-making to artificial intelligence because many of us don’t trust people — or computers — who use cold-cost benefit analysis to make ethical decisions.<sup>36</sup> Getting things right isn’t all that matters to us — we want decisions that reflect human emotions and seeming irrationalities. But these dynamics can be built into our legal algorithms just like anything else. If we want to acknowledge the diversity of viewpoints on a factual matter — that discrete subgroups of citizens can see things differently — we can program that in. Likewise, if we want murderers of pedophiles to get lighter sentences than the law implies or we want to allow for occasional nullification, we can direct the computer to that end. But to do so, we must take explicit action, acknowledging our departure from established principles and norms, and specifying exactly when departures are acceptable (e.g., nullification in instances like Peter Zenger’s seditious libel case against the Crown, but not for the murderers of civil rights leader Medgar Evers).<sup>37</sup> The result would be a far more transparent system. A person would know exactly what caused him to be convicted, and so would the rest of us.

With every individual’s case stored in a computer, we could constantly update the files. Whenever a new scientific breakthrough upended an established practice, leading to it being discredited, the computer would reweigh the evidence against every affected individual. Those who were no longer deemed guilty beyond a reasonable doubt would be released. We could use the same recalibration process for newly discovered evidence or changes in the law. Currently, the issue of applying new developments to old cases is extremely complicated, and that complexity can act to reinforce the status quo — at

the expense of fairness and accuracy. Why did the Department of Justice choose to ignore the scathing 2016 report by the President’s Council of Advisors on Science and Technology that concluded that many forensic disciplines — including bite-mark and shoe-tread analyses — lacked sufficient support to be scientifically acceptable?<sup>38</sup> In all likelihood it had nothing to do with the merits — the authors of the report had impeccable credentials and the findings were widely supported by the relevant scholarly community. The problem for Attorney General Loretta Lynch was that accepting the truth would throw thousands and thousands of cases into doubt, burdening an already over-extended judiciary and correctional system. With a computer in charge, everything can be updated automatically.

One of the greatest benefits of replacing conventional trials would be to allow us to rigorously assess each and every case. We simply do not have the resources to do that now, and so we rely on dangerous shortcuts, most notably, plea bargains. With trial by computer, ending plea bargaining is a realistic possibility because it requires no public participation and the process can be completed nearly instantaneously, at a fraction of the cost.

No change is without risk, and we should be aware of the threats entailed in such major innovation. Algorithms designed by humans can reflect human biases — a lesson that Google and other companies have learned the hard way, as their search tools have been criticized for reproducing forms of discrimination.<sup>39</sup> In the realm of criminal law, we must do more to eliminate the racial skew present in certain reoffending risk assessment and crime forecasting tools.<sup>40</sup> But biases in such rule-based algorithms are far easier to detect and eradicate than biases in flesh-and-blood human beings. Likewise, although fears of hacking are legitimate, they are not unique to this context. We rely on computers for almost every other sensitive area of our lives — from air traffic control to counterterrorism operations. In the hospital, an algorithm already processes the electrical activity of your

heart to identify atrial fibrillation.<sup>41</sup> And companies like IBM’s Watson Health are working to develop the capacity to diagnose a range of conditions by analyzing patients’ medical records with the aid of algorithms that can instantly search through all known medical literature.<sup>42</sup> Alzheimer’s, broken bones, lung cancer, concussions — all will be identified by a computer in the future. Why not guilt and innocence?

The truth is that a system of justice without human control is likely to be more just than a system with human control. And there is reason to think that our reticence may itself be a bias — what researchers have dubbed “algorithm aversion.”<sup>43</sup> In experiments, even when people observe that algorithms significantly outperform humans in predicting the future — as is generally the case across a wide range of contexts — they prefer to put their faith in human forecasters.<sup>44</sup> The reason appears to be that when we see an algorithm err we judge it far more harshly than a human making the same mistake, because we think humans will get better with practice.<sup>45</sup>

But, of course, algorithms can be made to learn — and much more effectively than humans. Indeed, in the coming decades, the greatest advances are likely to be made when we can augment simple rule-based algorithms with deep-learning-based ones. That’s already happening in medicine, with scientists recently training a computer to diagnose skin cancer by showing it thousands and thousands of lesion images along with the corresponding classifications.<sup>46</sup> The researchers didn’t just provide the existing rules to sort out malignant from benign; they let the computer figure out its own system. When matched up against real human dermatologists, the algorithm caught more melanomas and made fewer false positives.

Such innovation is possible in law as well. Provided with all the factual details for cases where guilt or innocence had been established with near-certainty (including those in the DNA exoneration database), a computer could learn the telltale signs of guilt or innocence that currently evade

## It is not a coincidence that more than half of those in prison have had a traumatic brain injury and nearly a quarter of the correctional population suffers from severe mental illness.

us. It could develop its own classification tools, and whenever it made a mistake it could immediately update its approach.

### A PUBLIC-HEALTH MODEL

Of course, no matter how bold we are, redesigning our trial process doesn't really get at the root of the problem in our criminal justice system. To do that, you've got to work to reduce crime itself. The reason we've come up short, despite investing a staggering amount of resources, is that we've been employing a flawed model of why people commit crimes. We think offending comes down to corrupt dispositions — evil hearts — and bad choices. In recent times, we've generally adhered to the view that criminals are rational creatures who can be deterred simply by adding more downside — longer sentences, worse incarceration conditions, and the like — into their calculus. Prison should be a miserable place not only because it further disincentivizes offending, but also because those who commit crimes deserve it. Years of solitary confinement, rape, brutalization — this is all acceptable because the people behind bars chose a life of crime with open eyes.

Although widely held, this account of criminality is not supported by the scientific literature. The more insight we gain into the genetic and environmental influences on criminal behavior, the harder it is to justify a world in which bad acts are assumed to reflect the freely made choices of evil people and offenders are treated with scorn. Every year, we learn more about the role of toxic substances and nutritional deficiencies in cognitive dysfunction linked to criminal behavior. It is not a coincidence that more than half of those in prison have had a traumatic brain injury and nearly a quarter of the correctional population suffers from severe mental illness.<sup>47</sup> It is not a coinci-

dence that those who are incarcerated are disproportionately uneducated, poor, and survivors of childhood abuse and neglect.

Based on the latest research, incrementalists have advanced a number of promising initiatives to reduce crime rates. Some have lobbied for abolishing mandatory minimums, repealing three-strikes laws, and reducing prison assaults by increasing monitoring. Since having a job correlates with a decrease in recidivism, others have pushed for new job-training programs for convicts and fought to end “check-the-box” rules that prevent those with criminal records from gaining employment.<sup>48</sup> As data shows that increasing the perception that perpetrators will be caught is a much more effective deterrent than increasing the harshness of punishment, advocates also have pushed to shift resources away from prisons and toward bolstering police presence.<sup>49</sup>

Many of these efforts are likely to play an important role in making our justice system more just and our cities safer. But, ultimately, they can feel somewhat arbitrary and overly cautious. If the evidence shows that our experiment with mass incarceration has been an abject failure, is reducing sentence lengths by a small amount going to do much? There's also something problematic about constantly shifting the boundary between who is legally responsible for their crimes and who is not. While we already acknowledge that some harmful acts are not the product of free will — a man whose sudden seizure causes him to drop his baby cannot be said to have chosen to assault his child — the lines we draw between compelled behavior and intentional conduct are a convenient fiction. They simply reflect the divide between the unmistakable, documented influences on human actions and the determinants that remain hidden. The fact that it is very difficult to figure out the partic-

ular nexus of factors that led a person to pull that trigger, kick in that door, or write that bad check does not mean that he freely chose to commit a crime.

Why not follow the evidence to its logical conclusion and reconceive efforts to reduce crime in terms of public health? What if we quit wasting time trying to sort out who deserves blame and got out of the payback business? What if we focused instead on remedying the harm, rehabilitating the criminal, discouraging others from taking similar actions, and treating the conditions that precipitated the crime in the first place?

This may sound revolutionary, but it's really not so different from how we handle outbreaks of disease. When a dangerous virus overwhelms a town, causation is relevant, but blame isn't. We don't treat someone who has contracted Ebola or dengue fever as sinful. We get to work restoring the person's health, preventing new cases, and trying to eliminate root causes. When an individual poses a particular threat to the public, we quarantine him until he's no longer a danger, but we don't subject him to poor treatment and contempt on the grounds that he is a wicked person who deserves it.

Other countries are already showing us the path forward. The penal systems of Germany and the Netherlands are organized around resocialization and rehabilitation. Prisoners are treated with dignity and have their rights to vote, work, and receive benefits restored when they are released — things they need to become productive citizens.

In Norway's Halden prison, inmates — including rapists and murderers — are locked in their cells only in the evening and spend their days working, studying, cooking, exercising, or playing music.<sup>50</sup> Rather than being denied positive human contact as punishment, they are encour- ▶

aged to maintain their family connections. And the staff members at Halden act as mentors, not enforcers, helping prisoners to overcome their problems and prepare to reintegrate back into society. The logic is simple: Place people in monstrous conditions, and you'll create monsters.

In the United States, we can catch a glimpse of the public-health model at work in the form of problem-solving courts, which have been around since the 1990s and have shown real promise in terms of reduced recidivism and cost.<sup>51</sup> These courts explicitly reject harsh punishment in favor of focused treatment for underlying problems like mental illness and drug abuse. The key stakeholders — prosecutors, defenders, judges, and offenders — are not cast as adversaries, but as partners, working together to develop a path forward.

That stands in stark contrast to our current approach, which is dominated by powerful group divisions: criminals versus police officers, prosecutors versus defenders, prisoners versus guards. We assume that such bipolarities ensure balance and accuracy — that the truth naturally emerges from the meritorious battle of clashing positions. But the available evidence casts significant doubt on that proposition and suggests that these group divisions can produce dangerous dynamics that encourage serious misconduct, from *Brady* violations to prisoner abuse. Detectives end up focused on gaining confessions, rather than on gathering reliable information. Prosecutors end up focused on nailing down guilty pleas, and defenders end up focused on getting

clients off, rather than ensuring accuracy and fairness. With a public-health mindset, we are freed to replace our adversarial approaches with a system dedicated to the common goals of truth, equality, and justice. Partisan experts and trial consultants have no place in such a landscape. Their valuable knowledge ought to be used for the common good.

Abandoning conflict and blame as organizing principles allows us to focus on the needs of those harmed by crimes, who have long been shunted to the side of the criminal-justice process and treated as mere props in the effort to gain a conviction. Helping victims to heal should be a central aim of our system. In some cases, that may mean facilitating apologies and aiding victims in forgiving those who have committed crimes against them. Recent research suggests that such actions can be far more effective at repairing the harm than retributive punishment of the offender.<sup>52</sup> In fact, granting forgiveness may provide a victim with a heightened sense of justice, as well as improved psychological well-being. In other cases, catering to a victim's needs may mean figuring out how the perpetrator can provide restitution. Even if offenders are not treated as blameworthy, they ought to mitigate the impact of what they've done.

Most importantly, a public-health model of crime allows us to shift resources from punishment to prevention. A reactive criminal-justice system, like the one we have now, is doomed to always come up short. There is no execution that can compensate for a victim's murder. There is no appeal process that can restore the

years lost to a wrongful conviction.

In the future, our major tools for fighting crime will not be police officers, trials, and incarceration, but better prenatal intervention, improved schools, and widely available mental-health care. We'll screen children for risk factors — just as we already do for learning disabilities, head lice, and hearing — and address them before criminal behavior ever manifests itself. That will make for duller episodes of *Law & Order*, but it will leave us far safer and more just.

The biggest obstacle in our way is the granite slab of history — an imposing gray façade with the etched words, “We've always done it this way, so how could it be wrong?” The words have warned off many would-be reformers, but the evidence from psychology and neuroscience can steady our hand this time around. Let's get out our chisels and hammers. A fair system is reachable — we just need the courage to raise our tools and not fear that we might endanger our core institutions as we begin to chip away at the rock of injustice. -----▶



**ADAM BENFORADO** is a law professor at the Drexel Kline School of Law. This essay draws upon material in his book *Unfair: The*

*New Science of Criminal Injustice* and articles published previously in *The Washington Post* and *The Chronicle of Higher Education*.

PHOTO BY JOE CRAIG

<sup>1</sup> GREEK DRAMAS BY AESCHYLUS, SOPHOCLES, EURIPIDES, AND ARISTOPHANES 322 (William James Hickie trans., D. Appleton and Company 1904).

<sup>2</sup> Brandon L. Garrett, *Introduction: New England Law Review Symposium on ‘Convicting the Innocent,’* 46 NEW ENGLAND L. REV. 683 (2012).

<sup>3</sup> Innocence Project, *DNA Exonerations in the United States*, accessed May 12, 2017, <http://www.innocenceproject.org/dna-exonerations-in-the-united-states/>.

<sup>4</sup> Saul M. Kassin et al., *Police-Induced Confessions: Risk Factors and Recommendations*, 34 LAW & HUMAN BEHAVIOR 3 (2010).

<sup>5</sup> Sonia Sotomayor, *Lecture: “A Latina Judge’s Voice,”*

N.Y. TIMES, May 14, 2009, available at <http://www.nytimes.com/2009/05/15/us/politics/15judge.text.html>.

<sup>6</sup> See, e.g., Neil Vidmar, *The Psychology of Trial Judging*, 20 CURRENT DIRECTIONS IN PSYCH. SCI. 60 (2011); see also Cass R. Sunstein & Thomas Miles, *Depoliticizing Administrative Law*, 58 DUKE L. J. 2193 (2008).

<sup>7</sup> Adam N. Glynn & Maya Sen, *Identifying Judicial Empathy: Does Having Daughters Cause Judges to Rule for Women’s Issues*, 59 AM. J. OF POL. SCI. 37 (2015).

<sup>8</sup> Susan Navarro Smelcer, *Supreme Court Justices: Demographic Characteristics, Professional Experience,*

*and Legal Education*, 1789–2010, CRS Report R40802 (Washington, D.C.: Library of Congress, Congressional Research Service, April 9, 2010), 6–11, 30–31, available at <http://www.fas.org/sgp/crs/misc/R40802.pdf>.

<sup>9</sup> See, e.g., Josh Salman, Emily Le Coz, & Elizabeth Johnson, *Florida’s Broken Sentencing System*, HERALD TRIB., available at <http://projects.heraldtribune.com/bias/sentencing/>; A. J. Wistrich, Chris Guthrie, & Jeffrey J. Rachlinski, *Can Judges Ignore Inadmissible Information? The Difficulty of Deliberately Disregarding*, 153 U. OF PENN. L. REV. (2005); Brite English, Thomas Mussweiler, & Fritz Strack, *Playing Dice With Criminal Sentences: The Influence of Irrelevant Anchors on*

- Experts' Judicial Decision Making, 32 PERSONALITY & SOC. PSYCH. BULLETIN 197 (2006); Abram Rosenblatt et al., *Evidence for Terror Management Theory I: The Effects of Mortality Salience on Reactions to Those Who Violate or Uphold Cultural Values*, 57 J. PERSONALITY & SOC. PSYCH. 681 (1989).
- <sup>10</sup> Shai Danziger, Jonathan Levav, & Liora Avnaim-Pesso, *Extraneous Factors in Judicial Decisions*, 108 PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES (2010), available at <http://www.pnas.org/content/108/17/6889>.
- <sup>11</sup> Kyoungmin Cho, Christopher M. Barnes, & Cristiano L. Guanara, *Sleepy Punishers Are Harsh Punishers: Daylight Saving Time and Legal Sentences*, 28 PSYCH. SCI. (2017), available at <http://journals.sagepub.com/doi/full/10.1177/0956797616678437>.
- <sup>12</sup> Allison Orr Larsen, *Confronting Supreme Court Fact Finding*, 98 VIRGINIA L. REV. 1255 (2012).
- <sup>13</sup> Jon D. Hanson & Douglas A. Kysar, *Taking Behavioralism Seriously: Some Evidence of Market Manipulation*, 112 HARVARD L. REV. 1420 (1999).
- <sup>14</sup> Nicole LeGrande & Kathleen Mierau, *Witness Preparation and the Trial Consulting Industry*, 17 GEO. J. OF LEGAL ETHICS 947 (2004).
- <sup>15</sup> *New Bloomberg Law Litigation Analytics Solution Provides Insights into Judicial Behavior*, PR NEWSWIRE, Oct. 18, 2016, available at <http://www.prnewswire.com/news-releases/new-bloomberg-law-litigation-analytics-solution-provides-insights-into-judicial-behavior-300346705.html>
- <sup>16</sup> Bloomberg Law, *Litigation Analytics: Technology Snapshot*, available at [https://www.bna.com/uploadedFiles/BNA\\_V2/Legal/Pages/Custom\\_Trials/BLAW/21573%20BLW%20Litigation%20Analytics%20Tech%20Snapshot%20\(007\)%20Final\(1\).pdf](https://www.bna.com/uploadedFiles/BNA_V2/Legal/Pages/Custom_Trials/BLAW/21573%20BLW%20Litigation%20Analytics%20Tech%20Snapshot%20(007)%20Final(1).pdf)
- <sup>17</sup> Joel D. Lieberman & Bruce D. Sales, *SCIENTIFIC JURY SELECTION* (Washington, DC: American Psychological Association, 2007), 9.
- <sup>18</sup> Alex Kozinski, *Criminal Law 2.0*, 44 GEO. L.J. ANN. REV. CRIM. PROC. iii (2015).
- <sup>19</sup> Daniel P. Mears, *Supermax Prisons: The Policy and the Evidence*, 12 CRIMINOLOGY AND PUB. POL'Y 681 (2013).
- <sup>20</sup> WILLIAM PENN, *SOME FRUITS OF SOLITUDE IN REFLECTIONS AND MAXIMS*, 80 (1901).
- <sup>21</sup> James W. Buehler, *Racial/Ethnic Disparities in the Use of Lethal Force by US Police, 2010–2014*, 107 AM. J. OF PUB. HEALTH 295, (Feb. 2017).
- <sup>22</sup> Sharon LaFraniere & Mitch Smith, *Philando Castile Was Pulled Over 49 Times in 13 Years, Often for Minor Infractions*, N.Y. TIMES, July 16, 2016.
- <sup>23</sup> Ian Ayres & Joel Waldfogel, *A Market Test for Race Discrimination in Bail Setting*, 46 STANFORD L. REV. 992 (1994); Jawjeong Wu, *Racial/Ethnic Discrimination and Prosecution*, CRIMINAL JUSTICE & BEHAVIOR (2016); David B. Mustard, *Racial, Ethnic, and Gender Disparities in Sentencing: Evidence from the U.S. Federal Courts*, 44 J. OF LAW & ECONOMICS 300 (2001).
- <sup>24</sup> Jennifer L. Eberhardt et al., *Looking Deathworthy: Perceived Stereotyping of Black Defendants Predicts Capital-Sentencing Outcomes*, 17 PSYCH. SCI. 383 (2006).
- <sup>25</sup> Maureen O'Sullivan, *The Fundamental Attribution Error in Detecting Deception: The Boy-Who-Cried-Wolf Effect*, 29 PERSONALITY & SOC. PSYCH. BULLETIN 1316, 1320, 1323–24 (2003).
- <sup>26</sup> Natasha Schvey, Rebecca Puhl, Katherine Levandoski, & Kelly Brownell, *The Influence of a Defendant's Body Weight on Perceptions of Guilt*, 37 INTERNATIONAL J. OF OBESITY 1279 (2013).
- <sup>27</sup> Dan Simon, *The Limited Diagnosticity of Criminal Trials*, 64 VANDERBILT L. REV. 175–76, n.140 (2011).
- <sup>28</sup> Sam Levin, *Ex-Stanford Swimmer Gets Six Months in Jail and Probation for Sexual Assault*, GUARDIAN, June 2, 2016, available at <http://www.theguardian.com/us-news/2016/jun/02/stanford-swimmer-sexual-assault-brock-allen-turner-palo-alto>.
- <sup>29</sup> Tracey Kaplan, *Stanford Sex Case: Brock Turner Found Guilty of Assault on Unconscious Woman*, MERCURY NEWS, Mar. 30, 2016, available at [http://www.mercurynews.com/breaking-news/ci\\_29705492/jury-finds-former-stanford-athlete-guilty-sex-assault](http://www.mercurynews.com/breaking-news/ci_29705492/jury-finds-former-stanford-athlete-guilty-sex-assault)
- <sup>30</sup> MAHZARIN R. BANAJI & ANTHONY G. GREENWALD, *BLINDSPOT: HIDDEN BIASES OF GOOD PEOPLE* 147 (2013).
- <sup>31</sup> Claire Cain Miller, *Is Blind Hiring the Best Hiring*, N.Y. TIMES, Feb. 25, 2016, available at [https://www.nytimes.com/2016/02/28/magazine/is-blind-hiring-the-best-hiring.html?\\_r=0](https://www.nytimes.com/2016/02/28/magazine/is-blind-hiring-the-best-hiring.html?_r=0).
- <sup>32</sup> Saul M. Kassin, Itiel E. Dror, & Jeff Kukucka, *The Forensic Confirmation Bias: Problems, Perspectives, and Proposed Solutions*, 2 J. OF APPLIED RESEARCH IN MEMORY & COGNITION (2013).
- <sup>33</sup> Equal Justice Initiative, *Illegal Racial Discrimination in Jury Selection: A Continuing Legacy* (2010).
- <sup>34</sup> Stephan Landsman & Richard F. Rakos, *A Preliminary Inquiry into the Effect of Potentially Biasing Information on Judges and Jurors in Civil Litigation*, 12 BEHAVIORAL SCIENCES AND THE LAW 113 (1994).
- <sup>35</sup> Dan M. Kahan, *Culture, Cognition, and Consent: Who Perceives What, and Why, in Acquaintance-Rape Cases*, 158 U. OF PA. L. REV. 729 (2010).
- <sup>36</sup> Jim A. Everett, David A. Pizarro, & M. J. Crockett, *Inference of Trustworthiness from Intuitive Moral Judgments*, 145 J. OF EXP. PSYCH.: GENERAL 772 (2016).
- <sup>37</sup> *When Jurors Ignore The Law*, N.Y. TIMES, May 27, 1997, available at <http://www.nytimes.com/1997/05/27/opinion/when-jurors-ignore-the-law.html>.
- <sup>38</sup> Jordan Smith, *FBI and DOJ Vow to Continue Using Junk Science Rejected by White House Report*, THE INTERCEPT, Sept. 23, 2016, available at <https://theintercept.com/2016/09/23/fbi-and-doj-vow-to-continue-using-junk-science-rejected-by-white-house-report/>.
- <sup>39</sup> Ben Guarino, *Google Faulted for Racial Bias in Image Search Results for Black Teenagers*, WASH. POST, June 10, 2016, available at [https://www.washingtonpost.com/news/morning-mix/wp/2016/06/10/google-faulted-for-racial-bias-in-image-search-results-for-black-teenagers/?utm\\_term=.a5bdfa00d1f3](https://www.washingtonpost.com/news/morning-mix/wp/2016/06/10/google-faulted-for-racial-bias-in-image-search-results-for-black-teenagers/?utm_term=.a5bdfa00d1f3)
- <sup>40</sup> Erica Posey, *When it Comes to Justice, Algorithms are Far From Infallible*, Brennan Center for Justice, March 27, 2017, available at <https://www.brennancenter.org/blog/when-it-comes-justice-algorithms-are-far-infallible>.
- <sup>41</sup> Siddhartha Mukherjee, *A.I. Versus M.D.*, NEW YORKER, April 3, 2017, available at <http://www.newyorker.com/magazine/2017/04/03/ai-versus-md>.
- <sup>42</sup> *Id.*
- <sup>43</sup> Berkeley J. Dietvorst, Joseph P. Simmons, & Cade Massey, *Algorithm Aversion: People Erroneously Avoid Algorithms After Seeing Them Err*, 114 J. EXP. PSYCHOL. 114 (2015).
- <sup>44</sup> *Id.*
- <sup>45</sup> *Id.*
- <sup>46</sup> Andre Esteva et al., *Dermatologist-Level Classification of Skin Cancer with Deep Neural Networks*, 542 NATURE 115 (2017); Mukherjee, *A.I. Versus M.D.*
- <sup>47</sup> Barbara H. Fried, *Beyond Blame*, BOSTON REV., Jun. 28, 2013, available at <http://bostonreview.net/forum/barbara-fried-beyond-blame-moral-responsibility-philosophy-law>.
- <sup>48</sup> JEREMY TRAVIS, *BUT THEY ALL COME BACK: FACING THE CHALLENGES OF PRISONER REENTRY* (2005).
- <sup>49</sup> National Institute of Justice, *Five Things About Deterrence*, Sept. 12, 2014, available at [http://nij.gov/five-things/Pages/deterrence.aspx?utm\\_source=eblastgovdelivery&utm\\_medium=eblast&utm\\_campaign=five+things+deterrence](http://nij.gov/five-things/Pages/deterrence.aspx?utm_source=eblastgovdelivery&utm_medium=eblast&utm_campaign=five+things+deterrence).
- <sup>50</sup> Amelia Gentleman, *Inside Halden, the Most Humane Prison in the World*, GUARDIAN, May 18, 2012, available at <http://www.guardian.co.uk/society/2012/may/18/halden-most-humane-prison-in-world>.
- <sup>51</sup> David DeMatteo et al., *Community-Based Alternatives for Justice-Involved Individuals with Severe Mental Illness: Diversion, Problem-Solving Courts, and Reentry*, 41 J. OF CRIMINAL JUSTICE 64 (2013).
- <sup>52</sup> Julie Juola Exline et al., *Forgiveness and Justice: A Research Agenda for Social and Personality Psychology*, 7 PERSONALITY & SOC. PSYCH. REV. 337 (2003); Michael Wenzel & Tyler G. Okimoto, *How Acts of Forgiveness Restore a Sense of Justice: Addressing Status/Power and Value Concerns Raised by Transgressions*, 40 EUR. J. OF SOC. PSYCH. 401 (2010); Charlotte van Oyen Witvliet et al., *Retributive Justice, Restorative Justice, and Forgiveness: An Experimental Psychophysiology Analysis*, 44 J. OF EXP. SOC. PSYCH. 10 (2008).