Dear Attorney General Bonta,

My name is Maya Dave, and I am a rising 9th grader at Stanford Online High School. I write on behalf of Mind on Trial, a youth-led nonprofit publication that uses rigorous neuroscience and legal analysis to advocate for a fairer, evidence-based justice system at the intersection of neuroscience, behavior, and law. I am writing to urge you to take action against the use of racially biased risk assessment tools in California's criminal sentencing system.

Every single day, California courts increasingly rely on algorithmic tools as a factor in deciding the sentence length for people. These tools, called risk assessment instruments, are designed to predict whether someone will commit a crime in the future. The most widely used of the many instruments currently available in the market include COMPAS, LSI-R, and the Public Safety Assessment. These tools have been in use in American courts since the 1990s, and their adoption has grown sharply over the past decade. Admittedly, California is not alone in this. As of 2023, more than half of the U.S. states use some form of algorithmic risk assessment at sentencing or parole.[1] Their scores are exerting a growing and largely unaudited influence on sentencing decisions across the state. These instruments do not make final decisions. They should not. The fact that they are used at all is evidence of witting or unwitting judicial prejudgment. These tools currently produce risk scores that judges consider alongside other factors at sentencing. But the weight those scores carry is disproportionately substantial, and the gap in how they are being validated, audited, and regulated is creating serious and documented problems of racial inequity.

Risk assessment tools, unfortunately, heavily rely on factors that are statistically correlated with race, including neighborhood location, employment history, family criminal history, and the level of education for the defendant. These are the proxies for structural disadvantages that have followed Black and Latino communities for generations as a direct result of discriminatory housing policy, unequal school funding, and over-policing.

When an algorithm is trained on data produced by a biased system, the output naturally also inherits that bias. A 2016 ProPublica investigation of COMPAS found that Black defendants were nearly twice as likely as white defendants to be falsely flagged as high risk for future crime.[2] These errors were not random. They followed a racial pattern: Black defendants were consistently over-flagged as high risk, while white defendants with similar profiles were more often labeled low risk. This disparity has since been replicated by independent researchers,

[1] Stevenson, M. & Doleac, J. (2021). Algorithmic risk assessment in the hands of humans. SSRN. https://ssrn.com/abstract=3489440; see also Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016). Machine bias. ProPublica. https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing
[2] Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016). Machine bias. ProPublica. https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

including a 2019 study published in Science Advances confirming that widely used risk tools perform unequally across racial groups.[3]

Beyond the statistical evidence of bias, neuroscience reveals why environmental disadvantage cannot be treated as a neutral proxy for individual risk. Brain development research, specifically the body of adolescent neuroscience research that led the Supreme Court to ban the death penalty for juveniles in *Roper v. Simmons* (2005), shows that behavior is shaped by environment, stress, trauma, and structural disadvantage in ways that are measurable and real.[4] A young person who grew up in a high-poverty neighborhood with limited access to mental health resources did not choose those circumstances. Punishing them more severely because of those circumstances, through an algorithm that converts inequality into a risk score, is the automation of inequity.

The constitutional problem is more direct than many realize, and it operates on two levels simultaneously. The first is equal protection: the Fourteenth Amendment does not permit a state to impose harsher sentences on the basis of race, regardless of whether the mechanism is a human decision or an algorithmic one. The second is due process: a defendant has the right to challenge the evidence used against them in court. When that evidence is a proprietary algorithm whose internal logic is shielded from disclosure, that right becomes meaningless. In *Daubert v. Merrell Dow Pharmaceuticals* (1993), the Supreme Court established that scientific evidence used in legal proceedings must be reliable, tested, and subject to peer review.[5] Risk assessment tools used in sentencing today do not consistently meet this standard. A 2021 report by the Partnership on AI found that most commercial risk assessment tools lack transparent validation data and are not tested for racial bias before deployment in courts.[6] Their racial error rates are documented in peer-reviewed literature. Their use continues anyway.

Your office has already recognized that racial disparities in California's criminal justice system demand urgent, undivided attention. The Racial Justice Bureau you established within the Department of Justice reflects exactly the kind of institutional commitment this issue requires. Now, we are asking you to extend that commitment to the sentencing tools currently operating in California courts.

Specifically, we strongly urge you to:

---

[3] Dressel, J., & Farid, H. (2018). The accuracy, fairness, and limits of predicting recidivism. Science Advances, 4(1). https://doi.org/10.1126/sciadv.aao5580

[4] American Psychological Association. (2004). Roper v. Simmons. https://www.apa.org/about/offices/ogc/amicus/roper

[5] Daubert v. Merrell Dow Pharmaceuticals, Inc., 509 U.S. 579 (1993). https://supreme.justia.com/cases/federal/us/509/579/

[6] Partnership on AI. (2021). Report on algorithmic risk assessment tools in the U.S. criminal justice system. https://partnershiponai.org/wp-content/uploads/2021/08/Report-on-Algorithmic-Risk-Assessment-Tools-in-the-U.S.-Criminal-Justice-System.pdf

(1) Direct the California Department of Justice to conduct an independent audit of all risk assessment tools currently used in criminal sentencing for racial bias and disparate impact;
(2) Issue guidance to California courts requiring full transparency and public disclosure of any algorithm used at sentencing, because no proprietary tool should determine a person's prison term without public accountability;
(3) Suspend the use of any tool found to produce racially disparate outcomes until it has been independently validated on California's population; and
(4) Support legislation requiring that sentencing frameworks center on individual circumstances and human judgment, informed by neuroscience evidence, rather than algorithmic prediction alone.

The use of these tools is expanding. Every year this continues is another year in which structural inequality, encoded in data, shapes the length of a person's sentence in a state that has committed itself to doing better.

Mind on Trial is a publication led by students of neuroscience and law who are also students of history. We know that change in this country has always required people naming the problem clearly, in public, and with evidence. We are doing our best to fulfill that here.

Thank you for your leadership on racial justice in California. We can be contacted at info@mindontrial.org. We look forward to your response.

With appreciation for your attention to this matter,
Maya Dave
Founder, Mind on Trial
info@mindontrial.org
mindontrial.org
Decoding the mind. Questioning the law.