

James P. Scanlan
Attorney at Law
1527 Wisconsin Avenue, NW
Washington, D.C. 20007
(202) 337-3927
jps@jpscanlan.com

April 12, 2018

ELECTRONICALLY TRANSMITTED

The Honorable Gene L. Dodaro
Comptroller General of the United States
GAO Headquarters
441 G Street, NW
Washington, DC 20548

Subj: Obligation of GAO to Explain to Congress and Executive Branch Agencies That, Contrary to the Belief Reflected in the March 2018 GAO Report on Public School Discipline Disparities, Generally Reducing Discipline Rates Tends to Increase, Not Reduce, the Proportions Blacks and Other More Susceptible Groups Make Up of Disciplined Students

Dear Mr. Dodaro:

The main purpose of this letter is to urge the Government Accountability Office (GAO) to immediately explain to Congress and the Departments of Education, Health and Human Services, and Justice that the understanding reflected in the March 2018 GAO report [*K-12 Education, Discipline Disparities for Black Students, Boys, and Students with Disabilities*](#)¹ that generally reducing public school discipline rates will tend to reduce the proportions blacks and other more susceptible groups make up of suspended students is the opposite of reality. A second purpose is to urge GAO to advise the Departments of Education, Health and Human Services, and Justice that they have an obligation to explain to the public and school administrators that the agencies' prior guidance regarding the effects of reducing adverse discipline outcomes on the measures of racial and other demographic differences on which the agencies have principally relied was also the opposite of reality.

In many documents and by many actions the three agencies have promoted the belief that relaxing standards and otherwise generally reducing adverse public school discipline outcomes

¹ To facilitate consideration of issues raised in documents such as this I include links to referenced materials in electronic copies of the documents, in some cases, for the reader's convenience, providing the links more than once. Such copies are available by means of the [Measurement Letters](#) page of jpscanlan.com. If the online version of the letter is amended, such fact will be noted on the first page of that version.

will tend to reduce (a) relative racial and other demographic differences in rates of experiencing the outcomes and (b) the proportions blacks and other more susceptible group make up of students experiencing the outcomes. The GAO report, among other ways, by suggesting that positive behavioral intervention and support (PBIS) and other programs that tend to reduce overall discipline rates will tend to reduce the measure of racial disparity on which the report relies, reflects the same belief.

In fact, however, generally reducing any outcome tends to increase, not reduce, both (a) and (b) as to the outcome. That is, reducing an outcome and thereby increasingly restricting it to those most susceptible to it, while tending to reduce relative differences in rates of avoiding the outcome (*i.e.*, experiencing the opposite outcome), will tend to increase relative differences in the outcome itself; correspondingly, reducing the outcome, while tending to increase the proportions groups more susceptible to the outcome make up of persons avoiding the outcome, will tend also to increase the proportions such groups make up of persons experiencing the outcome itself. By way of a simple example, lowering a test cutoff, while tending to reduce relative differences between pass rates of higher- and lower-scoring groups, tends to increase relative differences between the groups' failure rates; correspondingly, lowering the cutoff, while tending to increase the proportion the lower-scoring group makes up of persons who pass the test, tends also to increase the proportion the group makes up of persons who fail the test.²

I attach my December 8, 2017 written [testimony](#) explaining this issue to the U.S. Commission on Civil Rights (CCR). I also call your attention to my July 17, 2017 [letter](#) to the three agencies just mentioned, explaining the statistical issues and advising the agencies that they have an obligation to correct the mistaken understandings they have promoted. I also call your attention to [materials](#) I provided at a March 22, 2018 meeting with Department of Education staff, explaining the issues and recommending corrective actions for the agency to take once its leadership understands the issues.

Succinct explanations of the pertinent statistical patterns, with reference to government civil rights policies involving lending, school discipline, or criminal justice, may be found in my "[Misunderstanding of Statistics Leads to Misguided Law Enforcement Policies](#)," *Amstat News* (Dec. 2012), "[The Paradox of Lowering Standards](#)," *Baltimore Sun* (Aug. 5, 2013), "[Things government doesn't know about racial disparities](#)," *The Hill* (Jan. 28, 2014), and "[Things DoJ doesn't know about racial disparities in Ferguson](#)," *The Hill* (Feb. 22, 2016). Very recent succinct explanations of the patterns – with reference to the government's failure to understand that diversion programs contemplated by the Baltimore Police consent decree tend to increase, not reduce, the proportion blacks make up of persons with criminal records, and to the government's longstanding failure to understand that relaxing lending standards tends to

² When the population examined is made up of only two groups, the pattern of changes in the proportion the more susceptible group makes up of persons experiencing the two outcomes as an outcome is restricted to those most susceptible to it is directly related to the pattern of changes in the two relative differences. The matter is more complicated when the population is made up of several groups. But the complication is not of consequence to the principal issues addressed here.

increase, not reduce, relative racial differences in adverse borrower outcomes – may be found in my [“The misunderstood effects of the Baltimore police consent decree,”](#) *The Daily Record* (Feb. 15, 2018), and [“What the government gets wrong about fair lending,”](#) *American Banker* (Apr. 9, 2018). The latter matter was also the subject of my September 9, 2014 [letter](#) to Matthew J. Scirè GAO Director, Financial Markets and Community Investment, urging GAO to examine the ways misunderstandings of statistics undermine federal enforcement of fair lending laws.

More comprehensive treatments of these patterns, and other patterns by which measures of demographic differences tend to be affected by the prevalence of an outcome, may be found in my [Comments for Commission on Evidence-Based Policymaking](#) (Nov. 14, 2016) (First CEP Comments), [“The Mismeasure of Health Disparities,”](#) *Journal of Public Health Management and Practice* (July/Aug. 2016), [“Race and Mortality Revisited,”](#) *Society* (July/Aug. 2014), [“The Perverse Enforcement of Fair Lending Laws,”](#) *Mortgage Banking* (May 2014), [“Measuring Health and Healthcare Disparities,”](#) Proceedings of the Federal Committee on Statistical Methodology 2013 Research Conference (Mar. 2014), [amicus curiae brief](#) in *Texas Department of Housing and Community Development, et al. v. The Inclusive Communities Project, Inc.*, Supreme Court No. 13-1731 (Nov. 17, 2014), [“The Mismeasure of Discrimination,”](#) Faculty Workshop, University of Kansas School of Law (Sept. 20, 2013). Each of the first and final two items has a section (Sections C, I.B., and C, respectively) addressing the impossibility of quantifying demographic difference by comparing the proportion a group makes up of persons potentially experiencing an outcome (the pool) and the proportion the group makes up of persons experiencing the outcome, as was done in the GAO report on discipline disparities.

Graphical and tabular illustrations of the patterns by which measures tend to be affected by the prevalence of an outcome may be found in the methods workshop [“Rethinking the Measurement of Demographic Differences in Outcome Rates,”](#) Maryland Population Research Center of the University of Maryland (Oct. 10, 2014),³ a workshop to which I invited GAO personnel, including in the letter to Mr. Scirè. Slides 96 to 108 of the workshop address the impossibility of quantifying a demographic difference based on the aforementioned comparison of the proportion a group makes up of the pool and the proportion the group makes up of persons experiencing an outcome. See also the [IDEA Data Center Disproportionality Guide](#) subpage of the [Discipline Disparities](#) page of [jpscanlan.com](#),⁴ which, by email of August 15, 2014, I brought to the

³ Similar illustrations may be found in workshops given at [University of Massachusetts Medical School](#) (2015), [University of California, Irvine](#) (2015), [George Mason University](#) (2014), [University of Minnesota](#) (2014), [Harvard University](#) (2012), and [American University](#) (2012).

⁴ The Department of Education-funded measurement guide discussed on that page recommends measuring differences between (a) the proportion a group makes up of the students and (b) the proportion the group makes up of students experiencing an outcome in either absolute (percentage point) or relative terms. The GAO discipline disparities report employs the former approach. Under that approach, at schools having only black and white students and, for example, black and white suspension rates of 15% and 5%, the difference between (a) and (b) would be 8.6 percentage points at schools where blacks make up only 5% of students. As the proportion blacks make up of students increases, the difference between (a) and (b) would rise to a maximum of about 26.7 percentage points when blacks make up 40% students. As the proportion blacks make up of students continues to increase, the difference would drop, reaching 3.3 percentage points when blacks make up 95% of students. Thus, the approach

attention of GAO Director George Scott in connection with a proposal for a like workshop for analysts in the GAO Education, Workforce, and Income Security Issues group.⁵

Recent substantial treatments of the government's failure to understand the ways measures of demographic differences tend to be affected by the prevalence of an outcome, the deleterious consequences of leading the public and entities covered by civil rights law to believe that actions will tend to reduce measures of racial disparity when in fact the actions will tend to increase the measures, and the government's obligation to advise affected entities (including the government of the United Kingdom) regarding the way the government has misled them, as well as the uncertain prospects that the government will be able to understand these issues in the foreseeable future, may be found in my "[The Government's Uncertain Path to Numeracy](#)," Federalist Society Blog (July 21, 2017), "[Innumeracy at the Department of Education and the Congressional Committees Overseeing It](#)," Federalist Society Blog (Aug. 24, 2017), "[EEOC, OMB, and the Collection of Data That Can't Be Analyzed](#)," Federalist Society Blog (Sept. 7, 2017), "[The Pernicious Misunderstanding of Effects or Policies on Racial Differences in Criminal Justice Outcomes](#)," Federalist Society Blog (Oct. 12, 2017), "[United States Exports Its Most Profound Ignorance About Racial Disparities to the United Kingdom](#)," Federalist Society Blog (Nov. 2, 2017), "[The Misunderstood Relationship Between Racial Differences in Conduct and Racial Differences in School Discipline and Criminal Justice Outcomes](#)," Federalist Society Blog (Dec. 20, 2017), and "[The Dubious Research on the Adverse Effects of Stringent School Discipline Policies](#)," Federalist Society Blog (April __, 2018) (forthcoming)

I also call your attention to Section B (at 3-4) of my November 28, 2016 [Comments for the Commission on Evidence-Based Policymaking](#) (Second CEP Comments), a brief follow-up to the First CEP Comments mentioned above. The section discusses funding by the Departments of Education and Health and Human Services of activities that contribute to the mistaken understanding of the effects of programs like PBIS on standard measures of racial disparity in adverse school discipline outcomes. Similar points could be made about most or all of the activities funded by the Departments of Education, Health and Human Services, and Justice that are treated as useful resources in Appendix III (at 66-68) of the GAO discipline disparities report. For it should be evident that programs that promote understandings of the effects of policies on measures of racial disparity that are the opposite of reality ought not be funded by the government.

used in the report would yield vastly different disparities at school that have exactly the same black and white suspension rates. On the other hand, the relative difference between (a) and (b) would consistently decline as the black proportion of students increases. Thus, that approach would also result in substantial anomalies, though the anomalies would differ greatly from those yielded by the absolute difference. More pertinent to the principal issue addressed in this letter, however, regardless of the proportion blacks make up of students, reducing the frequency of suspensions will tend to increase both the absolute difference and the relative difference between the proportion blacks make up of students and the proportion they make up of suspended students.

⁵ I originally contacted Mr. Scott regarding the workshop proposal by email of July 24, 2014, at which time I understood him to be the GAO Director for Education, Workforce, and Income Security Issues. By email of August 18, 2014, Mr. Scott advised me that he had changed positions but would forward the proposal to the appropriate persons. I did not hear further from GAO on the matter.

I leave the illustration the pertinent statistical patterns to the attached CCR testimony and the above references. I note, however, the GAO report itself, read in conjunction with a March 21, 2014 Department of Education report titled “[Data Snapshot: School Discipline](#),” shows that recent reductions in out-of-school suspension rates were accompanied by increased relative racial differences in such suspensions. Table 12 (at page 71) of the report shows that according to Department of Education data covering the 2013-14 school year, the ratio of the black out-of-school suspension rate to the white out-of-school suspension rate was 3.9 (14.1%/3.6%). This is up from a 3.2 ratio (16%/5%) for the 2011-12 school year indicated in the March 2014 Data Snapshot. In recent years, similar patterns of increasing relative racial/ethnic differences in discipline rates have been observed across the country, as states and local educational authorities have been generally reducing discipline rates while mistakenly believing that doing so will tend to reduce the ratio of the black suspension rate to the white suspension rate.⁶

Such pattern is something that all persons or entities analyzing demographic differences should understand but that virtually no person or entity analyzing demographic differences does understand. There will of course be departures from such pattern, for factors are at work apart from the effects of the prevalence of the outcomes. But that in no way affects the need for those analyzing demographic differences to understand the pattern and other patterns by which measures of differences between outcome rates tend to be affected by the prevalence of an outcome. For without understanding such patterns, it is impossible to draw inferences about the nature of the forces causing the outcome rates of advantaged and disadvantaged groups to differ and impossible to divine whether those forces are increasing or decreasing over time or otherwise are greater in one setting than another.

As discussed in the CCR testimony and many other items mentioned above, the mistaken understanding of the effects of reducing an adverse outcome on the measures of racial disparity the government typically employ is but part of a larger failure of understanding by the government and the social and medical science communities of the ways all standard measures of differences involving outcome rates tend to be affected by the prevalence of an outcome and, consequently, of the impossibility of usefully analyzing data on demographic differences without consideration of the extent to which observed patterns are simply functions of changes in

⁶ See the following subpages to the Discipline Disparities page of jpscanlan.com: [California Disparities](#), [Colorado Disparities](#), [Connecticut Disparities](#), [Florida Disparities](#), [Maryland Disparities](#), [Minnesota Disparities](#), [Oregon Disparities](#), [Rhode Island Disparities](#), [Utah Disparities](#), [Beaverton, OR Disparities](#), [Denver Disparities](#), [Henrico County, VA Disparities](#), [Kern County \(CA\) Disparities](#), [Milwaukee Disparities](#), [Los Angeles SWPBS](#), [Loudoun County \(VA\) Disparities](#), [Minneapolis Disparities](#), [Montgomery County, MD Disparities](#), [Portland, OR Disparities](#), [St. Paul Disparities](#), [South Bend Disparities](#). See also the [DOE Equity Report](#), [Massachusetts Disparities](#), [Preschool Disparities](#), and [Suburban Disparities](#) subpages regarding the way that relative racial differences in discipline rates tend to be comparatively large in areas where discipline rates are comparatively low. This pattern is also reflected in the GAO report, which shows (Table 14, at 74) that as proportion of students who are low income decreases, the black and white out-of-school suspension rates also decrease, but the ratio of the black rate to the white rate increases.

the prevalence of an outcome and the extent to which they may reflect something else. As a result of that larger failure of understanding, the government has spent many billions of dollars conducting and funding research that rarely provides useful information for formulating policy, but commonly provides much that is misleading in that regard. See the recommendation at pages 2 to 4 of the July 17, 2017 to the Departments of Education, Health and Human Services, and Justice that they halt all funding of research into demographic differences that fails to consider these issues. See also the recommendations at pages 45 to 46 of the First CEP Comments.

As reflected in the First CEP Comments, as well as in “[The Mismeasure of Health Disparities](#),” “[Race and Mortality Revisited](#),” and “[Measuring Health and Healthcare Disparities](#),” [The Mismeasure of Health Disparities](#),” the government has devoted enormous sums to unsound health and healthcare disparities research. In that regard, it warrants note that more than a decade ago the National Center for Health Statistics (NCHS) recognized that as health and healthcare improve – and, thus, as favorable health and healthcare outcomes like survival and receipt of appropriate care generally increase and the corresponding adverse outcomes like mortality and non-receipt of appropriate care generally decrease – relative differences in the (increasing) favorable outcomes tend to decrease while relative differences in the corresponding (decreasing) adverse outcomes tend to increase. No other arm of the government has ever showed an understanding that it is even possible for the relative difference in the favorable outcome and the relative difference in the corresponding adverse outcome to change in opposite directions as the prevalence of the two outcomes changes, much less that NCHS has recognized that this tends to occur systematically. See Section A (at 1-3) of the Second CEP Comments regarding the failure of cancer disparities researchers to recognize that it is even possible for patterns of changes in relative differences in cancer survival to be the opposite of patterns of changes in relative differences in cancer mortality and the commonplace discussion of survival differences when the researchers have in fact examined mortality differences. The Patient Centered Outcomes Research Institute that GAO oversees, which itself funds health and healthcare disparities research, does not understand these issues any better than the government agencies that fund such research.⁷

I will likely address the larger issues at length with GAO soon. I note, however, that the First CEP Comments already provide substantial guidance for actions GAO ought now to take

⁷ As also discussed in these references, NCHS has not addressed the matter in a way that reflects an understanding of the purpose of health and healthcare disparities research. See “[The Mismeasure of Health Disparities](#)” regarding the agency’s recent reversal of its recommendation that healthcare disparities be measured in terms of relative differences in adverse outcomes, thus repudiating the research that relied on the agency’s earlier guidance including a decade of National Healthcare Disparities Reports (NHDRs). That is not to say that any of the earlier research into healthcare, or health, disparities, including that in the NHDRs, had been sound. See my July 1, 2015 [letter](#) to the Agency for Healthcare Research and Quality regarding the way that the agency’s confusions over measurement issues caused it to highlight as some of the largest decreases in disparities over a particular period situations where the agency would also regard the disparities to be substantially larger at the end of the period than at the beginning of the period. See also my August 29, 2017 [letter](#) to the National Quality Forum regarding the inability of that organization to provide anything useful with regard to the measurement of health and healthcare disparities, even though the organization is deemed a preeminent authority on measurement.

respecting those issues. See also the recommendations at pages 2 and 15-16 of my April 13, 2017 [letter](#) to Attorney General Jeff Sessions. There is no reason for GAO to await further communication from me to begin to address the longstanding innumeracy of the federal government with respect to essentially all matters pertaining to differences at which advantaged and disadvantaged groups experience favorable or adverse outcomes.⁸

And there is no reason whatever for GAO to fail to move immediately to correct the mistaken understanding communicated by the recent discipline disparities report. Already observers, including the individual members of Congress to whom the report specifically responds,⁹ are relying on the report, and its mistaken view as to the effects of generally reducing discipline rates on measures of racial disparity, to question actions being considered by the Departments of Education and Justice regarding their January 2014 Dear Colleague letter on public school discipline issues. Thus, GAO's correction of the mistaken understanding in its report is a matter of considerable urgency. GAO should also immediately recommend that the Departments of Education, Health and Human Services, and Justice take steps to inform the public and school administrators regarding the way the agencies' prior guidance and actions have misled them.

GAO should also expeditiously explain to all federal agencies, and to Congress, that the belief reflected in policies and legislation that generally reducing an outcome will tend to reduce relative demographic differences in rates of experiencing the outcome (or the proportions more susceptible groups make up persons experiencing the outcome) is incorrect, and that agencies that have promoted such belief have an obligation to correct it. That other governmental entities have so far been unable to understand this matter, and on their own could continue to fail to understand it for years or decades, heightens the obligation of GAO, an entity that is essentially the government's auditor, to understand the matter fully and to ensure that other arms of the government fully understand it as well.

Finally, implicit in the above discussion is that the GAO discipline disparities report should be immediately withdrawn since it will otherwise continue to mislead those who read it. But GAO should also review all of its publications to identify situations where the publication either contains a mistaken understanding of the type in the discipline disparities report or discusses measures of demographic differences in a way that misleadingly suggests to readers that the

⁸ The failures of understanding of the government and the social and medical science communities pertain not only to analyses involving different demographic groups, but also to analyses involving treated and control groups in clinical trials. See pages 41-43 of the November 14, 2016 CEP Comments. See also my [Comment on FDA Proposed Subgroup Regulations](#) (May 16, 2014) and [Comment on European Medicines Agency Subgroup Guidelines](#) (July 31, 2014). In that regard, I suggest that you will find that guidance the Patient Centered Outcomes Research Institute provides on the reporting of subgroups effects shows no understanding, for example, that an intervention that increases cancer survival will usually show a larger proportionate increase in survival among older subjects, but a larger proportionate decrease in mortality among younger subjects, or that it is not possible for a factor to cause equal proportionate changes in an outcome for groups with different baseline rates for the outcome while at the same time causing equal proportionate changes in the groups' rates of experiencing the opposite outcome.

⁹ See April 4, 2018 [statement](#) of Congressmen Jerrold Nadler and Bobby Scott.

The Honorable Gene L. Dodaro
Comptroller General of the United States
April 12, 2018
Page 8

measures are sound. The GAO report [*Multi-Agency Plan Need to Continue Progress Addressing High and Disproportionate Discipline Rates*](#) (Feb. 9, 2017) falls into the former category. The GAO report [*Standards Needed to Improve Identification of Racial and Ethnic Overrepresentation in Special Education*](#) (Mar. 29, 2013) that prompted my contacting George Scott in 2014 falls into the latter category. A longstanding problem in attention to racial differences in assignment to special education programs is the failure to recognize that generally reducing assignment rates, including by adding circumspection to the assignment process, will tend to increase the ratio of the black assignment rate to the white assignment rate. See the previously mentioned [“Innumeracy at the Department of Education and the Congressional Committees Overseeing It,”](#) Federalist Society Blog (Aug. 24, 2017). The report’s discussion of that ratio as a potentially sound means of measuring disproportionality, and the evident assumption that the higher the ratio in a particular jurisdiction the greater is the need for scrutinizing assignment procedures in the jurisdiction, contributes to a crucial failure of understanding of an subject to which states devote vast resources.¹⁰ Many GAO publications may raise similar issues and most of them should probably be withdrawn.

Sincerely,

/s/ James P. Scanlan

James P. Scanlan

Attachment

¹⁰ That holds also for the Department of Education regulations issued in response to the report, though implementation of those regulations has been delayed.

Measuring Discipline Disparities

James P. Scanlan

(Statement Prepared for U.S. Commission on Civil Rights Briefing “The School to Prison Pipeline: The Intersection of Students of Color with Disabilities” (Dec. 8, 2017)

Federal government policy regarding racial differences in school discipline outcomes has been consistently based on the belief that relaxing discipline standards and otherwise reducing adverse discipline outcomes will tend to reduce (a) relative (percentage) racial differences in rates of experiencing the outcomes and (b) the proportions African Americans and other racial minorities make up of persons experiencing the outcomes. In fact, exactly the opposite is the case.

By way of clarification, if the minority suspension rate is 15% and the white rate is 5%, the ratio of the minority rate to the white rate would be 3.0. That is, the minority rate is 200% greater than the white rate. The 200% figure is the relative, or percentage, difference. In the same situation, assuming minorities are 20% of students, they would be 43% of suspended students.

Federal policy has been based on the belief that activities that generally reduce suspensions (like Positive Behavioral Interventions & Support (PBIS) programs) will tend to reduce the 3.0 ratio and the 43% proportion figures. In fact, such activities will tend to increase those figures.

Test Score Illustration

Table 1 provides a simple illustration of why this is the case. The table is based on hypothetical test scores of higher- and lower- scoring groups (which are denominated AG for advantaged group and DG for disadvantaged group).

The first row of the table shows the pass rates for the two groups at a particular cutoff. The pass rates are 80% for AG and 63% for DG. Thus, AG’s pass rate is 1.27 times (27% greater than) DG’s pass rate.¹

¹ While I commonly refer to patterns of relative differences in this statement, the table actually presents rate ratios (also termed risk ratios or relative risks). The relative difference is the rate ratio minus 1 where the rate ratio is above 1 and 1 minus the rate ratio where the rate ratio is below one. In the former case, the larger the rate ratio, the larger the relative difference; in the latter case, the smaller the rate ratio, the larger the relative difference. It is more common to employ the disadvantaged group’s rate as the numerator for the favorable as well as the adverse outcome, which is the approach as to favorable outcomes of the “four-fifths” or “80 percent” rule for identifying disparate impact under the Uniform Guideline for Employee Selection Procedures. I have sometimes employed this approach, as in “Can We Actually Measure Health Disparities?,” *Chance* (Spring 2006) (http://www.jpscanlan.com/images/Can_We_Actually_Measure_Health_Disparities.pdf). More recently, however, I have usually used the larger figure as the numerator for both rate ratios, in which case, as to both favorable and adverse outcomes, the larger the rate ratio, the larger the relative difference. Choice of numerator in the rate ratio, however, has no bearing the patterns described here whereby measures tend to be affected by the prevalence of an outcome.

Table 1. Illustration of effect of lowering test cutoff on relative difference between pass rates of advantaged group (AG) and disadvantaged group (DG)

Cutoff	AG Pass Rate	DG Pass Rate	AG/DG Pass Ratio
1 High	80%	63%	1.27
2 Low	95%	87%	1.09

The second row shows what would happen if the cutoff is lowered to the point where AG’s pass rate is 95%. Assuming normal test score distributions, DG’s pass rate would be about 87%. With the lower cutoff AG’s pass rate would be only 1.09 times (9% greater than) DG’s pass rate. The fact that lowering a cutoff tends to reduce relative differences in pass rates is the reason why lowering a test cutoff is universally regarded as reducing the disparate impact of tests on which some groups outperform others.

At this point it may seem that I have contradicted my point at the outset. But, whereas lowering a cutoff tends to reduce relative differences in pass rates, it tends to increase relative differences in failure rates. This pattern is illustrated in Table 2. The table presents the same information as Table 1, but with the failure rates of the two groups added, along with the ratio of DG’s failure rate to AG’s failure rate (in the final column). The column with the rate ratios for test passage is highlighted in blue and the column with the rate ratios for test failure is highlighted in red.

Table 2. Illustration of effect of lowering test cutoff on (a) relative difference between pass rates and (b) relative difference between failure rates of advantaged group (AG) and disadvantaged group (DG)

Cutoff	AG Pass Rate	DG Pass Rate	AG Fail Rate	DG Fail Rate	AG/DG Pass Ratio	DG/AG Fail Ratio
1 High	80%	63%	20%	37%	1.27	1.85
2 Low	95%	87%	5%	13%	1.09	2.60

The final (red highlighted) column shows that with the initial cutoff DG’s failure rate was only 1.85 times (85% greater than) AG’s pass rate. With the lower cutoff, DG’s failure rate is 2.60 times (160% greater than) AG’s failure rate.

That is, as the prevalence of test passage and test failure generally changed as a result of lowering the cutoff, the relative difference in the increasing side of the dichotomy (test passage) decreased and the relative difference in the decreasing side of the dichotomy (test failure) increased.

As suggested at the outset, appraisals of discipline disparities issue sometimes focus on the proportions racial minorities make up of persons disciplined (compared with the proportions such groups make up of students). Patterns of changes in the proportions groups make up of persons experiencing either of the two outcomes as the prevalence of the outcomes changes are corollaries to the patterns shown in Table 2.

Table 3 is the same as Table 2, but with two more columns added on the right. These columns show the proportions DG makes up of persons who pass the test (highlighted in blue) and persons who fail the test (highlighted in red) in circumstances where DG makes up 50% of persons who take the test.

Table 3. Illustration of effect of lowering test cutoff on (a) relative difference between pass rates and (b) relative difference between failure rates of advantaged group (AG) and disadvantaged group (DG) and proportion DG makes up of (c) persons who pass the test and (d) persons who fail the test (where DG makes up 50% of test takers)

Cutoff	AG Pass Rate	DG Pass Rate	AG Fail Rate	DG Fail Rate	AG/DG Pass Ratio	DG/AG Fail Ratio	DG Prop of Pass	DG Prop of Fail
1 High	80%	63%	20%	37%	1.27	1.85	44%	65%
2 Low	95%	87%	5%	13%	1.09	2.60	48%	72%

The penultimate column shows that lowering the cutoff causes the proportion DG makes up of persons who pass the test to increase from 44% to 48%. That would reduce the difference between the proportion DG makes up of persons who take the test and the proportion it makes up of persons who pass the test.

But the final column shows that lowering the cutoff also increased the proportion DG makes up of persons who fail the test, from 65% to 72%. That would increase the difference between the proportion DG makes up of persons who take the test and the proportion DG makes up of persons who fail the test.

These patterns are not peculiar to test score data or the numbers I used to illustrate them. Rather, changing the frequencies of virtually any outcome and its opposite tends to cause the relative difference in the increasing outcome to decrease and the relative difference in the decreasing outcome to increase (with related effects on the proportions groups more susceptible to the outcomes make up of persons who experience the increasing outcome and the decreasing outcome).

This will not invariably happen with the consistency that will be observed with hypothetical test score data. For many factors are at work. But it will typically happen, especially when the changes in the prevalence of an outcome are substantial. In the school discipline context in particular, generally reducing discipline rates, while tending to reduce relative racial differences in rates of avoiding discipline (analogous to test passage), will tend to increase relative racial differences in rates of being disciplined (analogous to test failure). And in fact that is being observed all across the country as school districts have been generally reducing discipline rates while mistakenly believing that doing so should reduce relative racial differences in discipline rates (or the proportions racial minorities make up of student who are disciplined).²

² See page 8 of my July 17, 2017 letter to the Departments of Education, Health and Human Services, and Justice. http://www.jpscanlan.com/images/Letter_to_Departments_of_Education,_HHS,_and_Justice_July_17,_2017_.pdf

It is important to recognize that the situation is not one where the government has reasoned that, while the above-described patterns will be found in test score data, there are reasons why the patterns will not ordinarily be found in other situations. Rather, despite dealing with issues about demographic differences in test outcomes for half a century, the government has failed even to understand that lowering a test cutoff tends to increase relative differences in failure rates.

It is also important to understand that an increase in the relative difference in the adverse outcome does not mean that a disparity has increased in some meaningful sense any more than the reduction in the relative difference in the favorable outcome means that a disparity has decreased in a meaningful sense. Rather, the problem is that neither relative difference is a useful indicator of the strength of the forces causing the outcome rates of two groups to differ (or, as we might otherwise put it, the size of the difference in the circumstances of two groups reflected by their outcome rates). That is quite important to recognize as we endeavor to understand the causes of disparities and determine whether they are growing larger or smaller over time or are larger in one setting than another.

Still focusing on either Table 2 or Table 3 (though the former is somewhat simpler), one may think of the pass and fail rates as reflecting any favorable and adverse outcome rates that result from decisions of individual decision-makers. In the school discipline context, consider the failure rates as if they are the suspension rates of minorities and whites and the pass rates as if they are the groups' rates of rates of avoiding suspension. To the extent that bias on the part of decision-makers contributes to the differences between rates, any actions that reduce that bias will tend to reduce all measures of racial differences between favorable or adverse outcomes.

At the same time, however, simple reductions in adverse discipline outcomes, such as those resulting from PBIS programs, will tend to change the measures of difference in the manner reflected in the tables. Thus, in consequence of general reductions in discipline rates, a school district that substantially reduces suspension rates will tend to show a pattern of changing measures of differences in outcome rates akin to that found in movement from the first row to the second row of the two tables.

In circumstances where decision-makers, including teachers and administrators, are being encouraged to generally reduce suspension rates, all other things being equal, the results for decision-makers who do not try very hard to reduce suspension rates will tend to look more like the first row than the second row. The results for decision-makers who try very hard to reduce suspension rates will tend to look more like the second row than the first row.

Thus, consider a situation where the two rows reflect the results of actions of two different decision-makers and an effort is made to determine which decision-maker is more likely to have made racially biased decisions. One would reach opposite conclusions depending on whether one examined relative differences in the favorable outcome or relative differences in the adverse outcome. In fact, however, there is no rational basis for distinguishing between the two rows with regard to the question of which is more likely to reflect the results of biased decisions.

It should be evident that it is essential for school administrators endeavoring to address discipline disparities issues, and those monitoring those efforts and otherwise attempting to ensure equal

treatment for all groups, to understand these patterns. Yet the situation is not simply that virtually no one involved in such efforts understands these patterns; rather, virtually everyone involved in such efforts proceeds on a belief about the effects of generally reducing discipline rates on the measures most commonly employed in quantifying racial and other demographic disparities that is the opposite of reality.

Illustration of the Effects of Substituting a Reprimand for What Would Otherwise Be a First Suspension on Proportions More Susceptible Groups Make up of Persons Suspended

Data made available in Department of Education reports provide other simple illustrations of the effects of generally reducing adverse discipline outcomes rates on measures of racial or other demographic differences in discipline outcomes.

Tables 4 and 5 are based on data from a March 21, 2014 Department of Education report titled “Data Snapshot: School Discipline.”³ The data in the report enable one to determine the proportions demographic groups make up of K-12 and preschool students who are suspended (a) one or more times and (b) two or more times.⁴

Table 4. Illustration of effect of giving all students a reprimand instead of their first suspension on proportion African Americans make up of K-12 and preschool students receiving one or more suspensions

Setting	Number of Suspensions	AA Proportion of Students Experiencing the Outcome
K-12	One or more	37%
K-12	Two or more	43%
Preschool	One or more	44%
Preschool	Two or more	48%

Table 4 provides that information with regard to African American students in K-12 and preschool. The first row of the first set of two rows shows the proportion African Americans make up of K-12 students suspended one or more times (37%) and the second of those rows shows the proportion they make up of K-12 students suspended two or more times (43%). Suppose, then, that in every situation that otherwise would have resulted in a first suspension, the students were given a reprimand rather than a suspension. In such case, the figure in the second row would tend to become the figure for one or more suspensions. Thus, the 37% figure for the proportion African Americans make up of K-12 students suspended one or more times would tend to rise to 43%.

³ <https://www2.ed.gov/about/offices/list/ocr/docs/crdc-discipline-snapshot.pdf>

⁴ The document provided information on the proportions demographic groups made up of K-12 and preschool students suspended one time and suspended multiple times. From the information provided in the report, one can then determine the proportions the groups made up of persons suspended (a) one or more times and (b) two or more times.

The second two rows of the table provide a similar illustration for preschool. In this setting, giving students a reprimand instead of their first suspension would tend to cause the proportion African Americans make up of students suspended one or more times to increase from 44% to 48%.

Table 5 presents the same type of information for boys, who commonly have higher suspension rates than girls and thus commonly make up a larger proportion of suspended students than the approximately 50% that they make up of all students. Here, too, the Department of Education data show that in both K-12 and preschool, giving students a reprimand rather than what would otherwise be their first suspension would tend to increase the proportion boys (the group more susceptible to suspension) make up of students suspended one or more times.

Table 5. Illustration of effect of giving all persons a reprimand instead of their first suspension on proportion boys make up of K-12 and preschool students receiving one or more suspensions

Setting	Number of Suspensions	Male Proportion of Students Experiencing the Outcome
K-12	One or more	70%
K-12	Two or more	72%
Preschool	One or more	80%
Preschool	Two or more	82%

Illustration of Effects of the Prevalence of Adverse Discipline Outcomes in Different Settings on Measures of Racial Disparity in Those Settings

I often describe the statistical pattern at work in the discipline context (and essentially every other context where disparities are quantified in terms of relative differences or measures that are functions of relative differences) as that whereby the rarer and outcome, the greater tends to be the relative difference in experiencing it and the smaller tends to be the relative differences in avoiding it. One important, though universally misunderstood, manifestation of that pattern is that in settings (or among subpopulations) where adverse discipline outcomes are comparatively uncommon, relative racial differences in rates of experiencing those outcomes will tend to be larger, while relative differences in the corresponding favorable outcome will tend to be smaller, than in settings where the outcomes are comparatively common.

Tables 6 and 7 are based on data from the Massachusetts and Loudoun County, Virginia. Both are areas where policymakers or others have expressed concern that, though the areas have comparatively low suspension rates, relative racial differences or other measures of racial differences in suspensions are comparatively high.

The two tables may be compared to Table 2 above (save that they do not show the rates at which the two groups avoid suspension, the equivalent of test passage) with columns reordered to be more consistent with the way the issues are commonly discussed (and with the same color-coding for the rate ratios for the adverse and favorable outcomes). But I have added an additional column at the end termed EES, for estimated effect size. This column presents a measure of the strength of the forces causing outcome rates of two groups to differ that is theoretically unaffected by the prevalence of an outcome. I describe it (and its strength and weaknesses) in my “Race and Mortality Revisited,” *Society* (July/Aug. 2014)⁵ and various other places.

Table 6: Out-of-school suspension rates for African American and white students in Massachusetts and nationally in 2012-2013, with measures of difference

Area	AA Rate	White Rate	AA/White Ratio-Susp	White/AA Ratio - No Susp	EES
Massachusetts	10.0%	2.7%	3.70	1.08	0.65
National	16.4%	4.6%	3.57	1.14	0.71

Table 6 shows the common patterns whereby the setting with comparatively low suspension rates (Massachusetts compared with national figures) shows larger relative differences in suspension rates, but smaller relative differences in rates of avoiding suspension, than are found nationally. The EES figures – .65 in Massachusetts and .71 nationally – indicate that the forces causing suspension rates of African American and white students to differ (whatever those forces may be) are weaker in Massachusetts than nationally.⁶

Table 7 presents similar information from schools in Loudoun, County Virginia (an affluent suburb of Washington, DC), where suspension rates are very low. In this case, the concern about large racial disparities was prompted by the comparatively high ratio of the proportion African Americans made up of suspended students to the proportion they made up of students.⁷

⁵ http://jpscanlan.com/images/Race_and_Mortality_Revisited.pdf

⁶ These data and similar data relating to students with disabilities are discussed more fully in my November 12, 2017 letter to the Boston Lawyers’ Committee for Civil Rights and Economic Justice. http://jpscanlan.com/images/Letter_to_Boston_Lawyers_Committee_Nov.12.2015.pdf

⁷ That areas with low African American representation among students tend to have higher such ratios than other areas even when the areas have same suspension rates for African American students and for other students is among a number of reasons beyond the statistical patterns addressed here that comparisons of the proportion a group makes up of persons potentially experiencing an outcome and the proportion the group makes up of persons actually experiencing the outcome cannot effectively quantify the forces causing outcome rates of advantaged and disadvantaged groups to differ. See references in the succeeding note. See also the IDEA Data Center Disproportionality Guide subpage of the Discipline Disparities page of jpscanlan.com. <http://jpscanlan.com/disciplinedisparities/ideadatacenterguide.html>

The ratio African American suspension rate to the white suspension rate is actually slightly lower in Loudoun County than nationally, while the relative difference in rates of avoiding suspension is much lower in Loudoun County than nationally. The EES figures – .55 in Loudoun County and .71 nationally – indicate that the forces causing suspension rates of African American and white students to differ are considerably weaker in Loudoun County than nationally.⁸

Table 7: Out-of-school suspension rates for African American and white students in Loudoun County (VA) Public Schools and nationally in 2012-2013, with measures of difference

Area	AA Rate	White Rate	AA/White Ratio-Susp	White/AA Ratio - No Susp	EES
LCPS	4.65%	1.3%	3.54	1.04	0.55
National	16.4%	4.6%	3.57	1.14	0.71

Neither Massachusetts nor Loudoun County has any idea that to the extent that racial disparities in school discipline can be effectively measured, their disparities are smaller, not larger, than nationally. Nor do they have any idea that the actions to generally reduce discipline rates that they see as means of reducing the measures of racial disparity that are causing them concern will in fact tend to increase those measures.

Table 8, which is based on Table 8 of the aforementioned "Race and Mortality Revisited," is similar to Tables 6 and 7. But rather than comparing figures from a particular geographic area with national figures, Table 8 compares figures in preschool (where suspensions are comparatively rare) with figures from K12 (where suspensions are much more common). The table presents figures on multiple suspensions, which is the outcome respecting which racial disparities received the greatest attention when racial disparities in preschool suspensions first received substantial attention in 2014.⁹

⁸ These data are discussed more fully in the Loudoun County (VA) Disparities subpage of the Discipline Disparities page of [jpscanlan.com](http://jpscanlan.com/disciplinedisparities/loudounctydisparities.html) (<http://jpscanlan.com/disciplinedisparities/loudounctydisparities.html>). That subpage also discusses data showing that between the 2009-2010 and the 2013-2014 school years general reductions in suspension rates were accompanied by an increase in the relative differences between African American and white suspension rates and a decrease in the relative difference between African American and white rates of avoiding suspension, with negligible change in the EES. See also my September 5, 2017 letter explaining this issue to the Loudoun County School Board. http://jpscanlan.com/images/Letter_to_Loudoun_County_Public_Schools_Sept._5,_2017_.pdf

⁹ The facts receiving special attention in coverage of the issue were that African Americans were 18% of preschool children but 48% of preschool students receiving multiple suspensions. The figures in Table 8 are the suspension rates that can be derived from data in the previously mentioned Department of Education March 2014 document "Data Snapshot: School Discipline." The 18% and 48% figures were also highlighted in a March 21, 2014 Department of Education report titled "Data Snapshot: Early Childhood Education." <https://www2.ed.gov/about/offices/list/ocr/docs/crdc-early-learning-snapshot.pdf>

Table 8. African American and white rates of multiple suspensions in preschool and K-12, with measures of difference

Level	AA Mult Susp Rate	White Mult Susp Rate	AA/Wh Ratio Mult Susp	Wh/AA Ratio No Mult Susp	EES
Preschool	0.67%	0.15%	4.41	1.01	.49
K12	6.72%	2.23%	3.01	1.05	.51

As will commonly be observed, Table 8 shows that in the setting where suspensions are less common (preschool) relative differences in multiple suspension rates are greater, while relative differences in rates of avoiding multiple suspensions are smaller, than in the setting where suspensions are more common (K-12). In this case, however, the EES figures are very similar suggesting that, whatever the forces causing African American and white suspension rates to differ, they are of approximately the same strength in the two settings.

Table 9 is based on data from a 2012 Department of Education report titled “Helping to Ensure Equal Access to Education: Report to the President and Secretary.”¹⁰ Data were provided only on the proportion African Americans make of students and expelled students overall and in zero tolerance schools. The actual expulsions rates were not available. But based on the data available, one can present those two proportions in each setting and derive therefrom the relative difference between the African American rate and the rate for all other students.

Table 9: Proportions African Americans make up of students and expelled students overall and in schools with zero tolerance policies, with ratio of the African American expulsion rate to the white expulsion rate

Setting	AA Proportion of Students	AA Proportion of Expulsions	AA/Non-AA Expulsion Ratio
Overall	18%	39%	2.91
Zero Tolerance Schools	19%	33%	2.10

In accordance with the pattern described above, the ratio of the African American expulsion rate to the expulsion rate of other students was higher where expulsions were presumably less common (overall) than in the setting where expulsions were presumably more common (zero tolerance schools). (I do not present an EES figure because one needs the actual expulsion rates to derive such figure.) There nevertheless continues to be a near universal belief that zero tolerance policies lead to larger relative racial differences in adverse disciplines outcomes (and larger African American proportions or persons experiencing those outcomes) than more lenient policies.

An understanding of these patterns is also essential to drawing sound inferences about processes based on the comparative size of disparities. Relative racial differences in suspension rates are commonly greater, while relative differences in rates of avoiding suspension are commonly smaller, among girls (where suspensions are less common) than among boys (where suspensions

¹⁰ <http://www2.ed.gov/about/reports/annual/ocr/report-to-president-2009-12.pdf>

are more common). Correspondingly, relative gender differences in suspension are commonly greater, while relative gender differences in rates of avoiding suspension are commonly smaller, among whites (where suspensions are less common) than among African Americans (where suspensions are more common). See the Discipline Disparities page of jpscanlan.com.¹¹

Similarly, relative racial differences in suspensions will commonly be greater, while relative differences in avoiding suspensions will commonly be smaller, among students without disabilities (where suspensions are less common) than among students with disabilities (where suspensions are more common). Correspondingly, relative differences between the suspension rates of students with and without disabilities will commonly be greater, while relative differences between rates at which such groups avoid suspension will commonly be smaller, among whites (where suspensions are less common) than among African Americans (where suspensions are more common).

One cannot draw inferences about processes on the basis that one of these disparities is larger than another, or otherwise usefully hypothesize about why any disparity is larger than another, without understanding the above-described and other patterns by which measures tend to be affected by the prevalence of an outcome.

Conclusion

The failure to understand the ways the prevalence of an outcome affects relative differences in rates of experiencing an outcome and relative differences in rates of avoiding the outcome is but part of a larger failure of the government (and the social science and statistical communities) to understand the ways standard measures of differences between outcome rates of advantaged and disadvantaged group tend to be affected by the prevalence of an outcome. For more extensive treatment of that issue with regard to all analyses of demographic differences in outcome rates in the law and the social and medical sciences, see the aforementioned "Race and Mortality Revisited," my November 14, 2016 Comments for Commission on Evidence-Based Policymaking,¹² and my October 8, 2015 letter to the American Statistical Association.¹³ With regard to the way the larger failure has undermined Department of Education analyses of demographic differences regarding student outcomes apart from discipline, see my "Innumeracy at the Department of Education and the Congressional Committees Overseeing It," Federalist Society Blog (Aug. 24, 2017).¹⁴ See also the July 17, 2017 letter to the Departments of Education, Health and Human Services, and Justice mentioned in note 2 *supra*, which, in addition to advising the agencies of their obligations to correct prior guidance to school administrators as to the likely effects of generally reducing discipline rates on measures of discipline disparities, suggests that the agencies halt all funding of research into demographic

¹¹ <http://jpscanlan.com/disciplinedisparities.html>

¹² <https://www.regulations.gov/document?D=USBC-2016-0003-0135>

¹³ http://jpscanlan.com/images/Letter_to_American_Statistical_Association_Oct._8,_2015_.pdf

¹⁴ <http://www.fed-soc.org/blog/detail/innumeracy-at-the-department-of-education-and-the-congressional-committees-overseeing-it>

differences that fails to consider implications of the ways the measures employed tend to be affected by the prevalence of an outcome.

But the mistaken belief that generally reducing an adverse outcome should tend to reduce, rather than increase, relative differences in rates of experiencing the outcome (and the proportions groups more susceptible to the outcome make up of persons experiencing it) – which informs federal civil rights policies regarding criminal justice, lending, employment, and voter qualification, as well as school discipline – is an extreme example of the larger failure of understanding. And it has pernicious consequences. These include the many anomalies where by complying with government encouragements to relax standards and otherwise reduce adverse outcomes, entities covered by civil rights law increase the chances that the government will accuse them of discrimination. Similar anomalies exist in situations where individual actors who comply with their employers' instruction to reduce adverse outcomes increase the chances that their employees will accuse them of discrimination. Further, in contexts where actions that are supposed to be reducing measures of racial disparity are followed by increases in those measures, observers will conclude that the forces causing outcome rates to differ must be growing stronger, often prompting increasing distrust in the fairness of systems.

Such conclusions will not have a sound statistical basis. But so far very few people understand that.