

**James P. Scanlan
Attorney at Law
1527 30th Street, NW
Washington, D.C. 20007
jps@jpscanlan.com**

June 13, 2018

ELECTRONICALLY TRANSMITED

Sally Simpson, Professor and Chair
Department of Criminology and Criminal Justice
College of Behavioral & Social Sciences
University of Maryland
College Park, MD 20742

Re: Mistaken Premise of the Baltimore Police Consent Decree That Generally Reducing Adverse Criminal Justice Outcome Will Tend to Reduce, Rather Than Increase, (a) Relative Racial Differences in Rates of Experiencing the Outcomes and (b) the Proportion African Americans Make Up of Persons Experiencing the Outcomes

Dear Professor Simpson:

The purpose of this letter is to urge the faculty of University of Maryland Department of Criminology and Criminal Justice (CCJS) to explain to the parties and the court handling the consent decree covering practices of the Baltimore Police Department that the premise of the decree regarding its effect on measures of racial disparity in adverse criminal justice outcomes is incorrect.

The decree is premised on the belief that generally reducing adverse criminal justice outcomes (including adverse interactions with police officers) will tend to reduce (a) relative racial differences in rates of experiencing the outcomes and (b) the proportion African Americans make up of persons experiencing the outcomes. In fact, the opposite is the case.

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), tends 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, tends also to increase the proportions such groups make up of persons experiencing the outcome itself.

I attach my June 8, 2018 [letter](#)¹ to the Maryland Equity Project (MEP) and the Measurement, Statistics and Evaluation Program of the College of Education of the University of Maryland discussing two MEP studies that fail to understand the described pattern with respect to either school discipline or criminal justice outcomes. The letter attaches (a) my December 8, 2017 [testimony](#) to the U.S. Commission on Civil Rights, (b) my July 17, 2017 [letter](#) to the Departments of Education, Health and Human Services, and Justice, and (c) my [handout](#) for a March 22, 2018 meeting Department of Education staff. Those attachments principally address the federal government's mistaken belief that relaxing school discipline standards tends to reduce relative demographic differences in adverse discipline outcomes and the proportions more susceptible groups make up of persons experiencing the outcomes. But the mistaken belief about the effects of policies on measures of demographic differences in school discipline outcomes is the same as that underlying federal government actions regarding racial differences in criminal justice outcomes.

The key statistical pattern as it bears on misperceptions about the effects of policies on demographic differences in criminal justice or school discipline (or borrower) outcomes is explained briefly in my "[The misunderstood effects of the Baltimore police consent decree](#)," *The Daily Record* (Feb. 15, 2018), "[Things DoJ doesn't know about racial disparities in Ferguson](#)," *The Hill* (Feb. 22, 2016), "[The Paradox of Lowering Standards](#)," *Baltimore Sun* (Aug. 5, 2013), and "[Misunderstanding of Statistics Leads to Misguided Law Enforcement Policies](#)," *Amstat News* (Dec. 2012). This pattern and other patterns by which measures tend to be affected by the prevalence of an outcome, and the ways the failure to understand the patterns undermines analyses of demographic differences in the law and the social and medical sciences, are explained more fully in my "[Race and Mortality Revisited](#)," *Society* (July/Aug. 2014), and my [Comments for Commission on Evidence-Based Policymaking](#) (Nov. 14, 2016). Many graphical and tabular illustrations of the patterns may be found in the October 10, 2014 methods workshop I gave at the Maryland Population Research Center of the University of Maryland titled "[Rethinking the Measurement of Demographic Differences in Outcome Rates](#)" ([abstract](#)).²

¹ To facilitate consideration of issues raised in documents such as this I include links to referenced materials in electronic copies of the documents. 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.

² Similar workshops at arms of other universities include: "[The Mismeasure of Health Disparities in Massachusetts and Less Affluent Places](#)," Quantitative Methods Seminar, Department of Quantitative Health Sciences, University of Massachusetts Medical School (Nov. 18, 2015) ([abstract](#)); "[The Mismeasure of Discrimination](#)," Center for Demographic and Social Analysis, University of California, Irvine (Jan. 20, 2015); "[The Mismeasure of Demographic Differences in Outcome Rates](#)," Public Sociology Association of George Mason University (Oct. 18, 2014); "[The Mismeasure of Association: The Unsoundness of the Rate Ratio and Other Measures That Are Affected by the Prevalence of an Outcome](#)," Minnesota Population Center and Division of Epidemiology and Community Health of the School of Public Health of the University of Minnesota (Sept. 5, 2014); "[The Mismeasure of Group Differences in the Law and the Social and Medical Sciences](#)," Institute for Quantitative Social Science at Harvard University (Oct. 17, 2012); "[The Mismeasure of Group Differences in the Law and the Social and Medical Sciences](#)," Department of Mathematics and Statistics of American University (Sept. 25, 2012).

Recent discussions of the misunderstanding of the effects of policies on measures or racial disparity in criminal justice outcomes and the consequences of that misunderstanding, often with reference to the Baltimore Police consent decree, may also be found in my [“United States Exports Its Most Profound Ignorance About Racial Disparities to the United Kingdom,”](#) Federalist Society Blog (Nov. 2, 2017), [“The Pernicious Misunderstanding of Effects or Policies on Racial Differences in Criminal Justice Outcomes,”](#) Federalist Society Blog (Oct. 12, 2017), [“The Government’s Uncertain Path to Numeracy,”](#) Federalist Society Blog (July 21, 2017), [“Racial Impact Statement Laws in New Jersey and Elsewhere,”](#) Federalist Society Blog (Mar. 20, 2017), [“Compliance Nightmare Looms for Baltimore Police Department,”](#) Federalist Society Blog (Feb. 8, 2017), [“Misunderstanding of Statistics Confounds Analyses of Criminal Justice Issues in Baltimore and Voter ID Issues in Texas and North Carolina,”](#) Federalist Society Blog (Oct. 3, 2016), and [“Things the President Doesn’t Know About Racial Disparities,”](#) Federalist Society Blog (Aug. 5, 2016).

Among the consequences of the misunderstanding is that when policies that are supposed to reduce measures of racial differences in fact increase those measures, observers who believe that racial bias plays a large role in a racial difference will tend to believe that bias must be increasing. And all observers will tend to believe that, whatever the nature of the problems leading to racial differences in outcomes, the problems must be getting worse. Also, as discussed in the February 17, 2017 Federalist Society Blog post and several other places, by earnestly complying with guidance or instructions to generally reduce adverse outcomes, individual actors increase the chances that they will be accused of discrimination.

The recent study of the racial disparities in criminal justice outcomes in the United Kingdom that is the subject of the November 2, 2017 Federalist Society Blog post presents an especially perverse anomaly arising from the mistaken understanding of the effects policies on measures of racial disparity. A key theme of the study is the importance of promoting trust in the criminal justice system among minority communities in order that members of those communities will avail themselves of the advantages of pleading guilty to provable charges. Yet, as a result of its misunderstanding of measurements issues, the study itself is promoting distrust in the criminal justice among those communities. The situation is likely to worsen now that the Ministry of Justice has largely adopted the study’s recommendations.

So far, neither counsel for the parties nor the court has shown an understanding of this issue notwithstanding my numerous efforts to explain it to them. It may be that few members of the CCJS faculty currently understand the matter, and many or most members may share the mistaken view that generally reducing an outcome will tend to reduce relative differences in rates of experiencing the outcome and the proportions more susceptible groups make up of persons experiencing the outcome. Presumably, however, all members of the CCJS will understand the matter once they think it through.

But, absent some authoritative body’s explaining this matter to the parties and the court handling the Baltimore police decree, they, and, equally or more important, the citizens of

Sally Simpson, Professor and Chair
Department of Criminology and Criminal Justice
Page 4
June 13, 2018

Baltimore will be expecting the decree to reduce the aforementioned (a) and (b), when in fact it is more likely to increase those measures.

I suggest that a university based in Maryland has an obligation to ensure that its own scholars understand this issue and to take actions to correct the mistaken understanding of all pertinent parties including the public. That holds especially for an arm of the university like CCJS that has a special focus on statistical methods and data analysis regarding criminal justice issues. In that regard, I suggest that the CCJS should also determine whether any of its own work, like that the MEP, may contribute to the mistaken understanding of policies on demographic differences in criminal justice or other outcomes.

While the misunderstanding of the effects of the Baltimore police consent decree is the principal concern of this letter, I also note that racial disparities in public school discipline outcomes have been a prominent issue in Maryland for some years. As reflected in the attached materials and a number of the above references, that issue, too, has been approached from the mistaken perspective that generally reducing suspensions rates will tend to reduce relative racial differences in suspension rates and the proportion African Americans make up of suspended students.³ Thus, school discipline is another area where faculty members who deal with demographic differences and other quantitative issues should recognize an obligation to understand the matter and see that government officials and the public understand it as well.

Sincerely,

/s/ James P. Scanlan

James P. Scanlan

Attachment

³ As discussed in the attached letter to arms of the College of Education, a 2015 Maryland Equity Project study found that general reductions in suspensions in Maryland between the 2008-09 and 2013-14 school years were accompanied by increases in relative racial differences in suspension rates in Maryland generally and in 21 of 23 Maryland school districts on which data were available. See pages 2-3 of the letter regarding the way like patterns are being observed across the country.

**James P. Scanlan
Attorney at Law
1527 30th Street, NW
Washington, D.C. 20007
jps@jpscanlan.com**

June 8, 2018

ELECTRONICALLY TRANSMITED

Gail L. Sunderman, Director
Robert Croninger, Research Associate
Maryland Equity Project
Gregory R. Hancock, Director
Measurement, Statistics and Evaluation Program
College of Education
University of Maryland
College Park, MD 20742

Re: Mistaken Understanding That Generally Reducing Adverse Discipline and Criminal Justice Outcome Will Tend to Reduce, Rather Than Increase, (a) Relative Racial Differences in Rates of Experiencing the Outcomes and (b) the Proportion Blacks Make Up of Persons Experiencing the outcomes

Dear Director Sunderman and Research Associate Croninger of the College of Education's Maryland Equity Project and Director Hancock of the College's Measurement, Statistics and Evaluation Program:

This letter principally concerns two studies of the Maryland Equity Project (MEP) of the College of Education of the University of Maryland. I include Director Hancock among the recipients because the letter addresses measurement issues that pertain to many activities of the College's Measurement, Statistics and Evaluation Program (MSEP).

I just reviewed a November 2015 MEP Policy Brief by Matthew Henry titled "[Out-of-School Suspensions in Maryland Public Schools](#),"¹ and a June 2018 MEP Data Brief by Director Sunderland and Erin Janulis titled "[When Law Enforcement Meets to School Discipline: School-Related Arrests in Maryland 2015-16](#)." Both studies reflect the view, which has been promoted by the U.S. Departments of Education, Justice, and Health and Human Services, as well as many members of the social science community, that generally reducing adverse school

¹ To facilitate consideration of issues raised in documents such as this I include links to referenced materials in electronic copies of the documents. 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.

Gail L. Sunderman, Director (MEP)
Robert Croninger, Research Associate (MEP)
Gregory R. Hancock, Director (MSEP)
June 8, 2018
Page 2

discipline or criminal justice outcomes will tend to reduce (a) relative racial differences in rates of experiencing the outcomes and (b) the proportion blacks make up of persons experiencing the outcomes. In fact, exactly the opposite is the case.

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), tends 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, tends also to increase the proportions such groups make up of persons experiencing the outcome itself.

I attach my December 8, 2017 [testimony](#) explaining the issue to the U.S. Commission on Civil Rights, my July 17, 2017 [letter](#) explaining the issue to the U.S. Departments Education, Justice, and Health and Human Services, and the [handout](#) I used to explain the issue to Department of Education staff at a March 22, 2018 meeting.

The key points are explained fairly succinctly 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 DoJ doesn’t know about racial disparities in Ferguson](#),” *The Hill* (Feb. 22, 2016), and “[The misunderstood effects of the Baltimore police consent decree](#),” *The Daily Record* (Feb. 15, 2018). The failure to understand the aforementioned pattern and other patterns by which measures tend to be affected by the prevalence of an outcome, and the implication of that failure in the analysis of demographic differences in the law and the social and medical sciences, are explained more fully in my “[Race and Mortality Revisited](#),” *Society* (July/Aug. 2014), and my [Comments for Commission on Evidence-Based Policymaking](#) (Nov. 14, 2016). Many graphical and tabular illustrations of the patterns may be found in the October 10, 2014 methods workshop I gave at the University of Maryland’s Maryland Population Research Center titled “[Rethinking the Measurement of Demographic Differences in Outcome Rates](#)” ([abstract](#)).² A fair summary of the implications of the failure to understand such patterns with respect to the quantification of demographic differences in educational outcomes may be found in my “[Innumeracy at the Department of Education and the Congressional Committees Overseeing It](#),” *Federalist Society Blog* (Aug. 24, 2017).

In 2014, I created web pages discussing that recent reductions in suspensions in [Maryland](#) and in [Montgomery County, Maryland](#) were accompanied by increased relative racial differences in suspension rates. These patterns are being observed across the country, as reflected in the subpages to the [Discipline Disparities](#) page of [jpscanlan.com](#) involving the following states or local jurisdictions: [California Disparities](#), [Colorado Disparities](#), [Connecticut Disparities](#), [Florida](#)

² Professor Sangeetha Madhavan, Associate Director of the Maryland Population Research Center, who organized and attended the workshop, can provide information about the workshop.

Gail L. Sunderman, Director (MEP)
Robert Croninger, Research Associate (MEP)
Gregory R. Hancock, Director (MSEP)
June 8, 2018
Page 3

[Disparities](#), [Massachusetts Disparities](#), [Minnesota Disparities](#), [Oregon Disparities](#), [Rhode Island Disparities](#), [Utah Disparities](#), [Beaverton, OR Disparities](#), [Denver Disparities](#), [Henrico County, VA Disparities](#), [Kern County \(CA\) Disparities](#), [Los Angeles SWPBS](#), [Loudoun County \(VA\) Disparities](#), [Milwaukee Disparities](#), [Minneapolis Disparities](#), [Montgomery County, MD Disparities](#), [Portland, OR Disparities](#), [St. Paul Disparities](#), [South Bend Disparities](#), [Urbana Disparities](#).

The 2015 MEP Policy Brief, which relies on more recent data than my web pages on Maryland and Montgomery County, adds substantially to the body of evidence indicating that general reductions in adverse discipline outcomes will tend to increase, not reduce, relative racial differences in discipline rates. The document shows a continuation of the pattern whereby general reduction in out-of-school suspension in Maryland have been accompanied by an increase in the ratio of the black suspension rate to the white suspension rate. It also shows that in 21 of the 23 jurisdictions for which data were presented in an appendix, during the period of general reductions in suspensions between 2008 and 2014, overall suspension rates showed a larger percentage decrease than black suspension rates. That the overall percentage decrease was larger than black percentage decrease means that the ratio of the black suspension rate to the non-black suspension rate increased. While this ratio is something different from the ratio of the black suspension rate to the white suspension rate, the two ratios will commonly change in the same direction.

I often describe the statistical pattern most pertinent to interpretations of data on school discipline and criminal justice outcomes as that whereby the rarer an outcome the greater tends to be the relative difference in experiencing it and the smaller tends to be the relative difference in avoiding it. One manifestation of the pattern is that areas (or subpopulations) with comparatively low rates for adverse outcome (including among disadvantaged groups) tend to show comparatively large relative demographic differences in rates of experiencing the outcomes but comparatively small relative demographic differences in rates of avoiding the outcome. I gave substantial attention to the failure to understand this pattern in "Race and Mortality Revisited," and treat it with a focus on misinterpretations of demographic differences in Minnesota, Norway, and Sweden in "[It's easy to misunderstand gaps and mistake good fortune for a crisis](#)," Minneapolis Star Tribune (Feb. 8, 2014), and in Massachusetts in "[The Mismeasure of Health Disparities in Massachusetts and Less Affluent Places](#)," Quantitative Methods Seminar, Department of Quantitative Health Sciences, University of Massachusetts Medical School (Nov. 18, 2015) ([abstract](#)). See also the later paragraphs of "[United States Exports Its Most Profound Ignorance About Racial Disparities to the United Kingdom](#)," Federalist Society Blog (Nov. 2, 2017), and the [Massachusetts Disparities](#), [Loudoun County \(VA\) Disparities](#), [Suburban Disparities](#), [Preschool Disparities](#), [Restraint Disparities](#), and [DOE Equity Report](#) subpages of the Discipline Disparities page, as well as Table 2 to 5 of the March 22, 2018 Department of Education handout. Thus, it should not be surprising that the 2015 Policy Brief shows that Montgomery County, which has the lowest black suspension rate in Maryland, also has the largest ratio of the black rate to the white rate.

Gail L. Sunderman, Director (MEP)
Robert Croninger, Research Associate (MEP)
Gregory R. Hancock, Director (MSEP)
June 8, 2018
Page 4

The 2018 MEP Data Brief, in the context of discussion of the variation in arrest rates and disproportionality by school district, notes as an unexpected pattern that “Arundel County has a relatively low arrest rate (0.5 per 1000 students), but the [ratio of the] risk of arrest for black students [to non-black students] (7.79) and SWD to [non-SWD students] (6.86) is high.” But a correlation (though an imperfect one) between low overall rates³ and high risk ratios is something to be expected rather than be unexpected.

There are many pernicious consequences of leading observers to believe that policies that general reduce adverse school discipline and criminal outcomes will tend to reduce relative demographic differences in rates of experiencing the outcomes. I discuss some of these in “[The Pernicious Misunderstanding of Effects or Policies on Racial Differences in Criminal Justice Outcomes](#),” Federalist Society Blog (Oct. 12, 2017). I treat the problems facing the city of Baltimore as a result of the entry of a Consent Decree covering police practices that is premised on the mistaken belief that generally reducing adverse interactions between the police and the public will reduce the relative racial difference in rates of experiencing the interactions and the proportion blacks make up of persons experiencing the interactions, among other places, in “[Compliance Nightmare Looms for Baltimore Police Department](#),” Federalist Society Blog (Feb. 8, 2017), and “[The Government’s Uncertain Path to Numeracy](#),” Federalist Society Blog (July 21, 2017), as well as the recent *Daily Record* commentary discussed at the outset.

I suggest that the University of Maryland should be taking an affirmative role in correcting the mistaken belief regarding the likely effects of the Baltimore decree on measures of racial differences. But it certainly should not be contributing to that mistaken belief.

Finally, I emphasize that the patterns I describe by which measures tend to be affected by the prevalence of an outcome will not be observed in every case, since many other factors are at work.⁴ The second paragraph of the 2015 MEP Policy Brief show that between 1974 and 2010 the black suspension rate showed a larger percentage increase than the overall rate. That is contrary to the pattern I have described and would seem to indicate that factors other than the general increase in suspension were playing a large role in the matter. But departures from the patterns I describe do not detract from the need to understand the ways measures tend to be affected by the prevalence of an outcome in analyzing demographic differences and the need for analyses of demographic differences to attempt to determine the extent to which patterns of changes in measures of such differences (or the comparative size of the differences in different

³ Efforts to identify such correlations – and, indeed, any analyses of demographic differences – should focus on rates of advantaged and disadvantaged groups rather than overall rates. For the overall rate will be affected by the proportion the two groups (and other groups) make up of the overall population.

⁴ One of the two jurisdictions where the percentage decrease in the black suspension rates between 2008 and 2014 was larger than the percentage decrease in the overall suspension rate was Montgomery County (where the black rate decreased by 34.8% while the overall rate decreased by 32.0%). The difference between this pattern and the pattern discussed on my web page regarding Montgomery County may be related to the different time frames examined. But it may also reflect something else.

Gail L. Sunderman, Director (MEP)
Robert Croninger, Research Associate (MEP)
Gregory R. Hancock, Director (MSEP)
June 8, 2018
Page 5

settings) are solely functions of changes in the prevalence of an outcome (or the differing prevalence of the outcome in the different settings) and the extent to which the patterns reflect something meaningful about underlying processes.

I am continuing to press the U.S. Department of Education and other federal agencies to halt all funding of research into demographic difference involving favorable and adverse outcomes rates that fails to make such an attempt. See pages of 3-4 of the attached July 17, 2017 letter to the Departments of Education, Justice, and Health and Human Services and pages 46-47 of the November 14, 2016 Comments for the Commission on Evidence-Based Policy Making.

Please forward this letter to Matthew Henry, author of the 2015 MEP Policy Brief, and Erin Janulis, co-author of the 2018 MEP Data Brief, as well as other members of the staff of MEP or MSEP involved in analyses of (or teaching about) demographic differences.

Sincerely,

/s/ James P. Scanlan

James P. Scanlan

Attachments

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 groups 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.

James P. Scanlan
Attorney at Law
1529 Wisconsin Avenue, NW
Washington, D.C. 20007
(202) 338-9224
jps@jpscanlan.com

July 17, 2017

The Honorable Betsy DeVos
Secretary of Education
United States Department of Education
400 Maryland Avenue, SW
Washington, DC 20202

The Honorable Thomas E. Price, M.D.
Secretary of Health and Human Services
United States Department of Health and Human Services
200 Independence Avenue, SW
Washington, DC 20201

The Honorable Jeff Sessions
Attorney General
United States Department of Justice
950 Pennsylvania Avenue, NW
Washington, DC 20530-0001

Re: Obligations of the Departments of Education, Health and Human Services, and Justice to Correct Their Erroneous Guidance Suggesting That Relaxing Discipline Standards Tends to Reduce, Rather Than Increase, (a) Relative Demographic Differences in Discipline Rates and (b) the Proportions Groups More Susceptible to Adverse Discipline Outcomes Make Up of Persons Experiencing the Outcomes

Dear Secretary DeVos, Secretary Price, and Attorney General Sessions:

The purpose of this letter to advise the Departments of Education (DOE), Health and Human Services (HHS), and Justice (DOJ) of an obligation to correct erroneous guidance the three agencies have been providing the public, policymakers, and school administrators regarding the relationship between the stringency of school discipline standards and racial and other demographic differences in discipline outcomes. At least since the early years of this decade DOE and DOJ have been promoting the belief that relaxing standards and otherwise reducing rates of suspension and other adverse discipline outcomes will tend to reduce (a) relative (percentage) racial and other demographic differences in rates of experiencing the outcomes and (b) the proportions more susceptible groups make up of persons experiencing the outcomes. In

December 2014, the Secretary of HHS, in a document titled “[Policy Statement on Expulsion and Suspension Policies in Early Childhood Settings](#)” (Policy Statement) and an associated [Dear Colleague Letter](#), joined the Secretary of Education in promoting the belief that generally reducing adverse discipline outcomes would tend to reduce (a) and (b).

In fact, generally reducing any outcome tends to increase both (a) and (b) as to the outcome. Thus, the agencies have been leading a wide range of persons and entities to believe something about an important matter that is the opposite of reality. In any situation where government agencies have provided misleading guidance to the public the agencies have an obligation to correct the misleading guidance. The obligation is heightened where, as here, the agencies represent themselves to have, or are assumed by the public to have, expertise in the matter.

I briefly explain below the pertinent statistical point, which I have recently also explained in an April 13, 2017 [letter](#)¹ to Attorney General Jeff Sessions and Acting Assistant Attorney General T. E. Wheeler, III (Sessions letter) and in other communications to DOJ attorneys. Before doing so, however, I make certain preliminary points regarding the relationship of the principal subject of this letter to larger subjects the agencies must address if they are to fulfill their missions in a responsible manner.

Preliminary points regarding the instant subject and the larger subjects the agencies must address

This letter focuses on a discrete matter that agency officials, once having focused on a statistical pattern recognized more than a decade ago by the National Center for Health Statistics, should understand both to be undebatable and to involve agency actions that are the antithesis of responsible government. Further, the matter is something the three agencies can immediately begin to address at least by a Dear Colleague Letter explaining that express or implied guidance in prior such letters was incorrect. The matter also is quite pressing because thousands of school administrators across the country are continually endeavoring to implement policies promoted by the government (or incorporated into agreements with the government) while relying on the government’s mistaken guidance as to the effects of those policies on the measures of demographic differences that the government employs.² Numerous state and local governmental authorities have already taken actions based on the government’s erroneous guidance and others are considering like actions.

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.

² The matters is particularly pressing in the case of the school districts acting pursuant to agreements with DOE where the agency’s failure of understanding has created situations in which the more the school districts (or parts thereof) endeavor to comply with the agreement the more likely it is that DOE will regard them to have violated the agreement. See my September 20, 2016 [letter](#) to Oklahoma City School District.

But the agencies should recognize that the failure of understanding of elementary statistics that has led the agencies to provide the aforementioned mistaken guidance is part of a larger failure of understanding on the part of the agencies regarding the ways measures commonly employed in the analyses of demographic differences tend to be affected by the prevalence of an outcome. As a result of the larger failure of understanding, virtually nothing the agencies have themselves done, or that has been done pursuant to grants and contracts awarded by the agencies, regarding the analyses of demographic differences involving outcome rates has been statistically sound. See, *e.g.*, my [Comments for Commission on Evidence-Based Policymaking](#) (Nov. 14, 2016) (first CEBP comments), “[The Mismeasure of Health Disparities](#),” *Journal of Public Health Management and Practice* (July/Aug. 2016), “[Race and Mortality Revisited](#),” *Society* (July/Aug. 2014), and “[Measuring Health and Healthcare Disparities](#),” Proceedings of Federal Committee on Statistical Methodology 2013 Research Conference (March 2014). See also my “[Will Trump Have the First Numerate Administration?](#)” Federalist Society Blog (Jan. 4, 2017), regarding prospects that the current administration will be able understand things about analyses of demographic differences that other administrations have failed to understand.

In the case of DOE, the larger failure of understanding has prevented the agency from conducting any useful analyses of whether racial differences in educational outcomes like retention in grade, graduation, proficiency, assignment to disabled status, and various other matters have increased or decreased over time. See the [Educational Disparities](#) page of [jpscanlan.com](#) and its subpages, my August 24, 2015 [letter](#) to the HHS Secretary Sylvia M. Burwell and DOE Secretary Arne Duncan (at 9-11), and my April 18, 2012 [letter](#) to DOE Secretary Arne Duncan and Assistant Secretary of Education for Civil Rights Russlyn Ali (at 4). For example, as proficiency rates generally improve, relative demographic differences in rates of achieving proficiency tend to decrease while relative differences in rates of failing to achieve proficiency tend to increase; as proficiency rates generally improve, absolute demographic differences between rates of achieving basic proficiency (where rates are often well above 50 percent) tend to decrease, while absolute differences between rates of achieving advanced proficiency (where rates usually are well below 50 percent) tend to increase.³ To my knowledge, nothing DOE or any entity assisting it has done regarding analyses of demographic differences involving outcome rates has reflected an awareness of these patterns. Thus, DOE should undertake a complete review of the soundness of the methods by which it has analyzed demographic differences and of the soundness of the guidance it has provided on this subject. The agency should also institute a moratorium on grants and contracts (and activities pursuant to grants and contracts already awarded) to which these measurement issues pertain.⁴

³ Examples of these patterns may be found in the [Education Trust Glass Ceiling Study](#) subpage of the [Educational Disparities](#) page of [jpscanlan.com](#).

⁴ A minimum requirement of federally-funded research on demographic differences in outcome rates should be a commitment of the researchers to attempt to address the implications of the effects of the frequency of an outcome on the measures employed in the research. See fourth recommendation of the [first CEBP comments](#) (at 47). But the measurement issues addressed in those comment are pertinent both to activities involving analyses of demographic differences and activities that, while not necessarily involving analyses of such differences, are based on mistaken understandings regarding effects of policies on measures of demographic differences. The latter include, for example, activities that are based on the mistaken belief that positive behavioral intervention and support programs will tend to reduce relative racial differences in discipline rates, as in the case of the \$1 million grant discussed in

In the case of HHS, as discussed in the references at the top of page 3, the larger failure of understanding has led to the expenditure of many billions of dollars in research into demographic differences in health and healthcare outcome that has yielded very little of value even when it has not been patently misleading. One of the many situations exemplary of the failures of understanding on the part of HHS and its arms is the following. The National Center for Health Statistics (NCHS) more than a decade ago recognized that, as health and healthcare improve relative differences in favorable health and healthcare outcomes and relative differences in the corresponding adverse outcomes tend to change systematically in opposite directions as the prevalence of an outcome changes; yet, so far as the published record reveals, no other arm of HHS has recognized that it is even possible for relative differences in a favorable health and healthcare outcome and relative differences in the corresponding adverse outcome to change in opposite directions as the prevalence of an outcome changes. To my knowledge, no health or healthcare disparities research conducted or funded by arms of HHS has considered whether an observed pattern of changes in a measure employed in the research was anything other than a function of the change in the prevalence of the outcome. See the first four references at the top of page 3 and my [“The Mismeasure of Health Disparities in Massachusetts and Less Affluent Places,”](#) Quantitative Methods Seminar, Department of Quantitative Health Sciences, University of Massachusetts Medical School (Nov. 18, 2015). The points in the last two sentences of the prior paragraph regarding DOE apply equally to HHS.

In the case of DOJ, the consequences of the larger failure of understanding are summarized to a degree in the Sessions letter and include many situations where the more an entity complies with DOJ guidance (or obligations imposed by decrees in suits brought by the DOJ) the more likely the entity is to be sued by DOJ (or found not to comply with decree-imposed obligations). See my [“Compliance Nightmare Looms for Baltimore Police Department,”](#) Federalist Society Blog (Feb. 8, 2017), [“Things DoJ doesn’t know about racial disparities in Ferguson,”](#) *The Hill* (Feb. 22, 2016), [“Things government doesn’t know about racial disparities,”](#) *The Hill* (Jan. 28, 2014), [“Misunderstanding of Statistics Leads to Misguided Law Enforcement Policies,”](#) *Amstat News* (Dec. 2012). See also my [Comments on the Selection of Monitor of the Baltimore Police Consent Decree](#) (June 26, 2017) regarding the unlikelihood that the experts identified in the monitor proposals for the consent decree covering Baltimore Police practices understand the effects of reducing adverse criminal justice on measures of demographic differences any better than the government does.

Thus, each of the agencies has a responsibility to examine the problems in the analyses of demographic differences that it conducts or funds with an aim toward ensuring that future analyses are sound and that no further research, even on existing grants and contracts, continues to employ unsound methods. I may contact the agencies again regarding such matters. But there is no need for the agencies to await such contacts before examining the extent to which their failures to understand the ways measures tend to be affected by the prevalence of an outcome have undermined their activities.

Further, Section 5 of the Evidence-Policymaking Commission Act of 2016 imposes on each of the heads of DOE, HHS, and DOJ a responsibility to advise and consult with the Commission on Evidence-Based Policymaking regarding matters within the agency heads' areas of responsibility. Thus, the aforementioned reviews by DOE and HHS (and like actions suggested in the Sessions letter) should be conducted in a sufficiently timely fashion for the agencies to fulfill their responsibility to the Commission on Evidence-Based Policymaking before the Commission issues its report to Congress and the President this fall. I suggest that my comments for the Commission dated [November 14, 2016](#), and [November 28, 2016](#), provide the agencies a useful guide for advising the Commission as to the ways the agencies' missions have so far been undermined by the failure to understand the statistical patterns described in the comments.

Attention to these larger subjects, however, should not interfere with the agencies' fulfilling their responsibilities to immediately correct their guidance regarding the effects of relaxing discipline standards on measures of difference in school discipline outcomes.

Patterns by which restricting adverse outcomes to those most susceptible to them tends to increase measures of demographic differences as to the outcomes

For reasons related to the shapes of underlying distributions of factors associated with experiencing an outcome or its opposite, all standard measures of differences between outcome rates (*i.e.*, the proportions of demographic groups experiencing a binary outcome) tend to be affected by the frequency of an outcome. The pattern most pertinent here is that whereby the rarer an outcome, the greater tends to be the relative difference in experiencing it and the smaller tends to be the relative difference in avoiding it (*i.e.*, experiencing the opposite outcome). A corollary to this pattern is a pattern whereby the rarer an outcome, the greater tend to be the proportions groups most susceptible to the outcome make up of both persons who experience the outcome and persons who avoid the outcome.

The patterns can be easily illustrated with normally distributed test score data. Table 1 below, which is also Table 1 of the Sessions letter, shows the pass and fail rates of an advantaged group (AG) and a disadvantaged group (DG) at two cutoff points in a situation where the groups have normally distributed test scores with means that differ by half a standard deviation (a situation where approximately 31 percent of DG's scores are above the AG mean) and both distributions have the same standard deviation. The table also shows (in columns 5 through 8) measures that might be used to appraise differences in test outcomes of AG and DG.

Column 5, which presents the ratio of AG's pass rate to DG's pass rate,⁵ shows that at the higher cutoff, where pass rates are 80 percent for AG and 63 percent for DG, AG's pass rate is 1.27

⁵ While I commonly refer to patterns of relative differences in this letter, 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

times (27 percent greater than) DG’s pass rate. If the cutoff is lowered to the point where AG’s pass rate is 95 percent, DG’s pass rate would be about 87 percent. At the lower cutoff, AG’s pass rate is only 1.09 times (9 percent greater than) DG’s pass rate.

Table 1. Illustration of effects of lowering a test cutoff on measures of differences in test outcomes

Row	(1) AG Pass Rate	(2) DG Pass Rate	(3) AG Fail Rate	(4) DG Fail Rate	(5) AG/DG Pass Ratio	(6) DG/AG Fail Ratio	(7) DG Prop of Pass	(8) DG Prop of Fail
1	80%	63%	20%	37%	1.27	1.85	44%	65%
2	95%	87%	5%	13%	1.09	2.60	48%	72%

That lowering a cutoff tends to reduce relative differences in pass rates is well understood and underlies the widespread view that lowering a cutoff tends to reduce the disparate impact of tests on which some groups outperform others.

But, whereas lowering a cutoff tends to reduce relative differences in pass rates, it tends to increase relative differences in failure rates. As shown in column 6, initially DG’s failure rate was 1.85 times (85 percent greater than) AG’s failure rate. With the lower cutoff, DG’s failure rate is 2.6 times (160 percent greater than) AG’s failure rate.

Columns 7 and 8 show the proportions DG makes up of persons who pass and fail the test at each cutoff in a situation where DG makes up 50 percent of persons taking the test. Column 7 shows that lowering the cutoff increases the proportion DG makes up of persons who pass from 44 percent to 48 percent (hence, *reducing* all measures of difference between the proportions DG makes up of persons who took the test and persons who passed the test). Column 8 shows that lowering the cutoff increases the proportion DG makes up persons who fail the test from 65 percent to 72 percent (hence, *increasing* all measures of difference between the proportions DG makes up of persons who took the test and persons who failed the test).

The patterns reflected in Table 1 are not peculiar to test score data or the numbers I used to illustrate them. Rather, the patterns can be found in virtually any setting where two groups have different, more or less normal, distributions of factors associated with experiencing some outcome. Income and credit score data, for example, show how lowering an income or credit score requirement, while tending to reduce relative racial differences in meeting the requirement, will tend to increase relative racial differences in failing to meet the requirement. See Tables 2 and 3 of the Sessions letter. The information in the tables necessarily also means that lowering the requirements increases the proportions African Americans make up of persons who meet the

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). 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 ratio, the larger the relative difference. Choice of numerator in the rate ratio, however, has no bearing on the patterns by which as the frequency of an outcome changes, the two relative differences tend to change in opposite directions.

requirement and persons who fail to meet the requirement. Many other examples may be found in the longer references listed at the top of page 3, the scores of web pages on jpscanlan.com devoted to measurement issues, and the university methods workshops and conference presentations listed under the [Conference Presentations](#) subpage of the Publications page of jpscanlan.com.

The patterns are also evident in many types of data on school discipline outcomes, including data in DOE publications. Tables 2 through 5 below are based on data from a March 2014 DOE publication titled “[Data Snapshot: School Discipline](#).” 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) more than one time. Tables 2 and 3 present that information for black and male K-12 students and Tables 4 and 5 present the information for black and male preschool students.⁶

The tables illustrate the effects of relaxing standards in a way that would cause all students to receive a reprimand rather than what would otherwise be their first suspension. Such a modification would cause the proportion the indicated groups makes up of students with one or more suspensions to change from that in the first row to that in the second row. Thus, for example, as shown in Table 2, relaxing the standard in the manner indicated would cause the proportion African American students make up of K-12 students suspended one or more times to increase from 37 percent to 42 percent.

Table 2. Illustration of effect of giving all persons a reprimand instead of their first suspension on proportion black students make up of K-12 students suspended one or more times

Outcome	Black Proportion of K-12 Students Experiencing the Outcome
One or more suspensions	37%
Two or more suspensions	42%

Tables 3 shows a like pattern for male K-12 students, and Tables 4 and 5 shows like patterns for black and male preschool students.

⁶ Demographic differences in rates of experiencing things like single suspensions cannot be effectively analyzed, just as differences in rates of receiving grades of C or experiencing fair health cannot be effectively analyzed. See the [Intermediate Outcomes](#) subpage of the Scanlan’s Rule page of jpscanlan.com. It is possible that DOE has come to appreciate aspects of this issue. In DOE’s 2016 publication on school discipline titled “[2013-2014 Civil Rights Data Collection – A First Look](#),” the agency no longer presented data on single suspensions but included information on single suspensions within the category of “one or more suspensions.”

Table 3. Illustration of effect of giving all persons a reprimand instead of their first suspension on proportion male students make up of K-12 students suspended one or more times

Outcome	Male Proportion of K-12 Students Experiencing the Outcome
One or more suspensions	70%
Two or more suspensions	72%

Table 4. Illustration of effect of giving all persons a reprimand instead of their first suspension on proportion black preschool students make up of preschool students suspended one or more times

Outcome	Black Proportion of Preschool Students Experiencing the Outcome
One or more suspensions	44%
Two or more suspensions	48%

Table 5. Illustration of effect of giving all persons a reprimand instead of their first suspension on proportion male preschool students make up of preschool students suspended one or more times

Outcome	Black Proportion of Preschool Students Experiencing the Outcome
One or more suspensions	80%
Two or more suspensions	82%

If standards were further relaxed such that all persons were given reprimands for what would otherwise be their first two suspensions, the figures for the proportion black and male students make up of persons experiencing one or more suspensions would almost certainly rise still further. Rarely will one fail to observe such a pattern in circumstances where there are large numbers of observations.

In the school discipline context, in point of fact, one observes that all across the country recent reductions in discipline rates have been accompanied by increased relative racial/ethnic differences in discipline rates. See the following web pages discussing such patterns with respect to the jurisdictions indicated in the page titles: [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](#), [Los Angeles SWPBS](#), [Minneapolis Disparities](#), [Montgomery County, MD Disparities](#), [Portland, OR Disparities](#), [St. Paul Disparities](#), [South Bend Disparities](#).⁷ These patterns are occurring notwithstanding that

⁷ These situations usually caught my attention as a result of press reportage of the fact that discipline rates had generally declined but racial disparities had increased, often while reflecting the mistaken belief that the general declines in discipline rates should have resulted in reductions in the racial disparity. Reportage that general declines in discipline rates were accompanied by decreased racial differences in discipline generally involves situations where the observers are measuring discipline disparities in terms of absolute differences between rates.

school districts may well be doing many things beyond relaxing standards in attempting to reduce racial/ethnic differences in discipline rates.

See also (a) the [DOE Equity Report](#) subpage of the [Discipline Disparities](#) page of [jpscanlan.com](#) (regarding data in a November 2012 DOE Office of Civil Rights document titled "[Helping to Ensure Equal Access to Education: Report to the President and Secretary](#)") showing that, contrary to the agency's attribution of large relative differences in adverse discipline outcomes to zero tolerance policies, relative racial differences in expulsions are smaller in districts with zero tolerance policies than in districts without such policies) and (b) Table 8 of "Race and Mortality Revisited" (showing that relative differences in multiple suspensions are larger, though relative differences in avoiding multiple suspensions are smaller, in the setting where multiple suspensions are less common (preschool) than in the setting where multiple suspensions are more common (K-12)).

These patterns, of course, will not be observed in every case, since other factors will be at work. But that does not alter the fact that general reductions in discipline rates will tend to affect measures of demographic difference in ways that are the exact opposite of what the government has been leading school administrators and others to believe. Further, the effects of the misunderstanding promoted by the government are substantial, as teachers and administrators must struggle to explain to supervisors, oversight authorities, and the public (and, in the case of agreements with the DOE, to the DOE itself) why relaxing of standards are accompanied by effects on measures of disparity in adverse discipline outcome that are the opposite of what DOE and other government agencies have led them to expect.

In these circumstances, the obligation of the agencies to correct the misunderstandings it has promoted, and to do so as soon as possible, should be evident.⁸

Sincerely,

/s/ James P. Scanlan

James P. Scanlan

⁸ One closely related matter that also requires early attention from DOE involves the agency's perceptions about the implications of the fact that students with disabilities make up a high proportion of persons subject to physical restraints. See the [Restraint Disparities](#) subpage of the [Discipline Disparities](#) page of [jpscanlan.com](#) regarding the agency's singling out of states based on the proportion students with disabilities make up of students physically restrained where the states the agency singles out favorably are those least likely to adhere to DOE guidance to employ physical restraints as a last resort, while the states the agency singles out unfavorably are those most likely to adhere to DOE guidance on the matter.

Materials for Meeting of Department of Education Staff with James Scanlan
(Mar. 22, 2018)

Issues (pages 1-2)

Illustrative Tables and Figures (pages 3-7)

Extended References (pages 8-11)

ISSUES

Issue 1:

Guidance by the Department of Education (DOE), as well as Departments of Justice (DOJ) and Health and Human Services (HHS), regarding school discipline policies has been premised on the belief that relaxing standards and otherwise generally reducing suspension rates will tend to reduce (a) the ratio of the African American suspension to the white suspension rate and (b) the proportion African Americans make up of suspended students. In fact, exactly the opposite is the case.

Recommendations for DOE action:

a. Communicate (ideally in conjunction with DOJ and HHS) to school administrators, the public, and Congress (by Dear Colleague letters and otherwise) that prior guidance as to the effects of policies on measures of racial disparity was incorrect.

b. Advise Congress of the ways statutes involving education and youth justice issues are premised on the mistaken belief that generally reducing adverse outcomes will tend to reduce the measures of disproportionality typically used by the government.

c. Review all agreements with school districts to determine whether the agreements require modifications to practices that tend to increase (a) and (b) while contemplating measuring compliance in terms of reductions in (a) and (b).

Issue 2:

There exists a general failure of persons and entities analyzing demographic differences regarding rates at which advantaged and disadvantaged groups experience favorable or adverse outcomes to recognize the ways measures employed in such analyses tend to be affected by the prevalence (frequency) of the outcomes. Analyses of such differences and guides thereon have almost invariably been unsound and misleading because they have not addressed (a) the extent to which observed patterns of changes in a measures are functions of the change in the prevalence of the outcome and (b) the extent to which such patterns reflect something significant about underlying processes, including the effects of policies aimed at mitigating the comparative disadvantage of certain groups.

Recommendations for DOE action:

a. Withdraw (or withdraw DOE association with) all research involving analyses of demographic differences that has attempted to quantify such differences, and all materials providing guidance on quantifying those differences, that have failed to consider the effects of the prevalence of an outcome on measures employed or discussed.

b. Review all DOE research and research grants to determine whether they fail to address the implications of the effects of the prevalence of an outcome on the measures employed or discussed; halt all funding that cannot be shown to address those implications in a useful manner.

c. In conjunction with other agencies, form a committee to reform the analyses of demographic differences.

Key references (available on web by means of title search or on Measurement Letters page of jpscanlan.com):

Statement of James P. Scanlan 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)

“Innumeracy at the Department of Education and the Congressional Committees Overseeing It,” Federalist Society Blog (Aug. 24, 2017)

Letter to United States Departments of Education, Health and Human Services, and Justice (July 17, 2017)

Comments of James P. Scanlan for Commission on Evidence-Based Policymaking (Nov. 14, 2016)

ILLUSTRATIVE TABLES AND FIGURES

Table 1. Illustration of effects of lowering a test cutoff on measures of differences in test outcomes of advantaged group (AG) and disadvantaged group (DG) (based on situation where groups are of equal size) (Table 1 of July 17, 2017 letter to DOE, HHS, DOJ)

Row	(1) AG Pass Rate	(2) DG Pass Rate	(3) AG Fail Rate	(4) DG Fail Rate	(5) AG/DG Pass Ratio	(6) DG/AG Fail Ratio	(7) DG Prop of Pass	(8) DG Prop of Fail
1	80%	63%	20%	37%	1.27	1.85	44%	65%
2	95%	87%	5%	13%	1.09	2.60	48%	72%

Table 1 illustrates that lowering a test cutoff – and thereby generally increasing pass rates and generally reducing failure rates – tends to *reduce* relative differences in pass rates (Column 5) and *increase* relative difference in failure rates (Column 6). Table also shows that lowering cutoffs tends to *increase* both the proportion DG makes up persons who pass (Column 7) and the proportion DG makes up of persons who fail (Column 8).

Considerations:

- Improving education in way that enables everyone scoring between the two cutoffs to reach the higher cutoff will have the same effect as lowering the cutoff.
- In circumstances where favorable and adverse outcome rates in the two rows result from actions of decisionmakers, there is no rational basis for distinguishing between the two rows with respect to the likelihood of decisionmaker bias.
- Other things being equal, decisionmaker who employs more relaxed standards or are more cautious about imposing adverse outcomes will tend show results more like those in Row 2 than Row 1.
- Patterns in the two rows are akin to those one would find where Row 1 involves more serious (often deemed objectively-identified) offenses while Row 2 involves less serious (often deemed subjectively-identified) offenses. See Offense Type Issues subpage of Discipline Disparities page of jpscanlan.com.
- Regarding Columns 4 and 8, a pattern that it is crucial to know, though virtually no one in fact knows, is that generally reducing an adverse outcome tends to (a) *reduce* the proportion of a disadvantaged group that experiences the outcome but (b) *increase* the proportion the disadvantaged group makes up of persons who experience the outcome.
- Lowering the cutoff decreased the absolute (percentage point) difference between pass (or fail) rates from 17 to 8. Usually when observers say that general reductions in suspensions decreased a disparity (mainly Daniel Losen and colleagues), they are referring to the percentage point difference. That does not mean that the absolute difference is a useful measure of association. See "Race and Mortality Revisited," *Society* (July/Aug. 2014) and Figures 1 and 2 and Table 6 *infra*.

Table 2. 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 (Table 4 of testimony to Commission on Civil Rights)

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 2 illustrates that a policy of giving reprimands instead of what would otherwise be first suspensions will tend to increase proportion African Americans make up of persons with one or more suspensions.

Table 3. African American and white rates of multiple suspensions in preschool and K-12, with measures of difference (Table 8 of Commission on Civil Rights testimony and Table 8 or “Race and Mortality Revisited,” Society (July/Aug. 2014))

Level	(1) AA Multiple Susp Rate	(2) Wh Multiple Susp Rate	(3) AA/Wh Ratio Mult Susp	(4) Wh/AA Ratio No Mult Susp	(5) EES
Preschool	0.67%	0.15%	4.41	1.01	.49
K12	6.72%	2.23%	3.01	1.05	.51

Table 3 illustrates that relative differences in receiving multiple suspensions are larger (Column 3), but relative differences in avoiding multiple suspensions are smaller (Column 4), in preschool (where multiple suspensions are comparatively rare) than in K-12 (where multiple suspensions are more common). Column 5 shows that, to the extent that the forces causing black and white rates to differ can be measured, they are about the same in both settings. Illustration is based on data from March 21, 2014 DOE report titled “Data Snapshot: Early Childhood Education” underlying the fact highlighted in the document, and much-cited in discussions of it, that African American children, who make up 18% of preschool students, make up 48% of preschool students with multiple suspensions.

Table 4. States regarded favorably and unfavorably in March 21, 2014 DOE document. 1) titled “Data Snapshot: School Discipline.”

State	Proportion of restrained students who were students with disabilities	Way state was regarded by DOE	Likely degree to which states follows DOE guidance on restraints
Nevada	96%	Unfavorably	High
Florida	95%	Unfavorably	High
Wyoming	93%	Unfavorably	High
Arkansas	43%	Favorably	Low
Louisiana	41%	Favorably	Low
Mississippi	40%	Favorably	Low

See Restraint Disparities subpage of the Discipline Disparities page of jpscanlan.com regarding reasons why following DOE guidance to restrict the use of physical restraints to the most extreme cases tends to increase, not reduce, the proportion students with disabilities make up of restrained students.

Table 5: Proportions African Americans make up of expelled students overall and in schools with zero tolerance policies, with ratio of the African American expulsion rate to the white expulsion rate (based on 2012 DOE report titled “Helping to Ensure Equal Access to Education: Report to the President and Secretary”) (Table 9 of Commission on Civil Rights testimony)

Setting	(1) AA Proportion of Students	(2) AA Proportion of Expulsions	(3) AA/Non-AA Expulsion Ratio
Overall	18%	39%	2.91
Zero Tolerance Schools	19%	33%	2.10

Table 5 illustrates that the African American/white expulsion ratio is greater in schools without zero tolerance policies than in schools with zero tolerance policies.¹

¹ One can derive the rate ratio in Column 3 from the figures in Columns 1 and 2 even though one does not have the actual rates. One needs the actual rates to attempt to determine whether forces causing rates to differ are greater in schools with or without zero tolerance policies. This is one of the reasons, but not the only reason, one can never analyze a demographic difference in the basis of a comparison between the proportion a group makes up of students and the proportion it makes up of students experiencing an outcome. See Section C the Kansas Law paper “The Mismeasure of Discrimination,” Section I.B of the Texas Department of Housing brief, and Section C of the November 14, 2016 Comments to the Commission on Evidence-Based Policy Making (listed in Section B of Extended References); see also the IDEA Data Center Disproportionality Guide subpage of the Discipline Disparities page of jpscanlan.com.

Figure 1. Absolute differences between rates of AG and DG pass (or fail) rates at various cutoff points defined by AG fail rate (Figure 2 CEBP Comments)

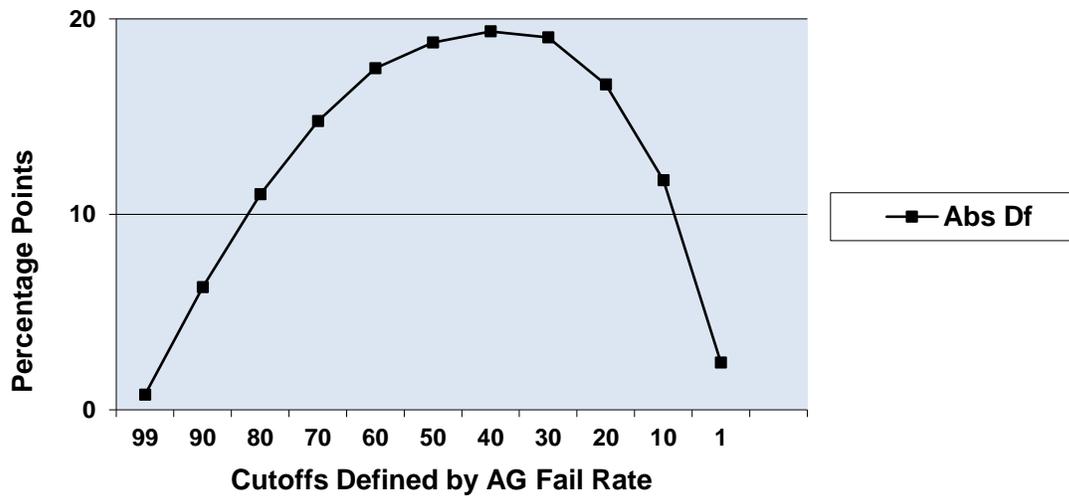
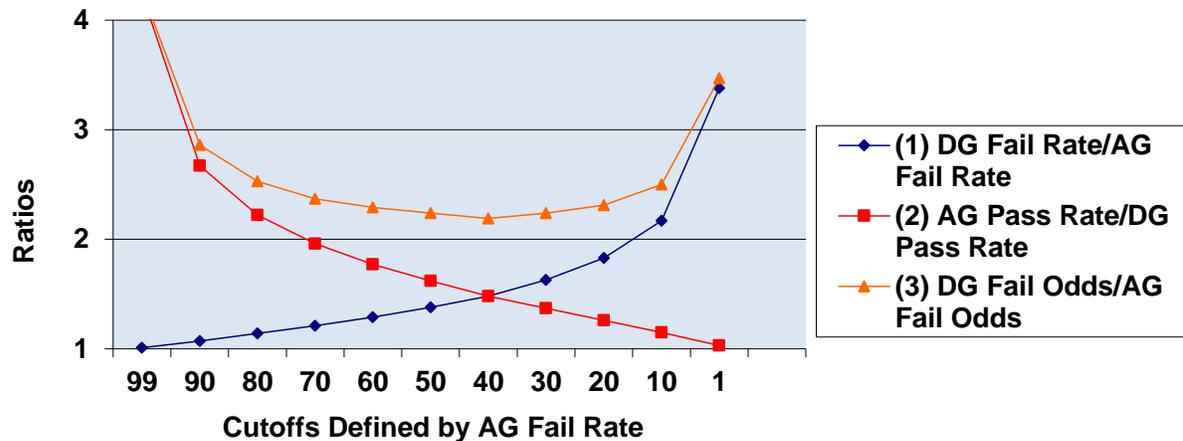


Figure 2. Ratios of (1) DG fail rate to AG fail rate, (2) AG pass rate to DG pass rate, (3) DG failure odds to AG failure odds (Figure 2 from the CEBP Comments)



Figures 1 and 2, which are based on the same specifications as Table 1, illustrate the effect of lowering a cutoff from a point where almost everyone fails to the point where almost everyone passes. Notice that direction of change in the absolute difference tends to track direction of change of the smaller of the two relative differences (initially (1)/diamond marker, later (2)/rectangle marker). Because observers who rely on relative differences to measure disparities commonly rely on the larger of the two relative differences (school discipline, mortgage outcomes, poverty, unemployment), such observers tend to reach opposite conclusions about directions of changes in disparities from observers who rely on absolute differences.

Table 6. Favorable outcome rates of advantaged group (AG) and disadvantaged group (DG) at four levels of prevalence with different favorable outcome frequencies, with measures of difference

	(1) AG Fav Rate	(2) DG Fav Rate	(3) AG/DG Fav Ratio	(4) DG/AG Adv Ratio	(5) Absolute Diff (Perc Points)	(6) Odds Ratio
A	20.0%	9.0%	2.22 (1)	1.14 (4)	11.0	2.53
B	40.0%	22.6%	1.77 (2)	1.29 (3)	17.4	2.28
C	70.0%	51.0%	1.37 (3)	1.63 (2)	19.0	2.24
D	80.0%	63.4%	1.26 (4)	1.83 (1)	16.6	2.31

Table 6 Illustrates that across all prevalence ranges general increases in favorable outcomes tend to reduce relative differences in those outcomes (Column 3) while increasing relative differences in the corresponding adverse outcomes (Column 4). The highlighted absolute difference column (5) shows that generally increasing an uncommon outcome (e.g., rates of advanced proficiency) tends to increase absolute (percentage point) differences between rates, as reflected by movement from row A to Row B; but generally increasing a common outcome (e.g., rates of achieving basic proficiency) tends to reduce absolute differences between rates, as reflected by movement from Row C to Row D.

See Educational Disparities page of jpscanlan.com and its subpages. See discussion of Table 5 in "Race and Mortality Revisited.," *Society* (July/Aug. 2014) and discussion (at 337-339) regarding the implications of failure to understand the pattern by which absolute differences tend to be affected by the prevalence of an outcome with respect to disparities reduction elements in pay-for-performance programs, especially in Massachusetts.

EXTENDED REFERENCES

All items listed below are available online and most can be accessed by web searches for their titles. Items that may not be found by web searches should be available on the Measurement Letters page of jpscanlan.com.

A. Short items explaining the mistaken understanding of effects of relaxing standards on measures of demographic difference involving school discipline or criminal justice outcomes (essentially primers on Issue 1)

“Things Do doesn’t know about racial disparities in Ferguson,” *The Hill* (Feb. 22, 2016)

“Things government doesn’t know about racial disparities,” *The Hill* (Jan. 28, 2014).

“The Paradox of Lowering Standards,” *Baltimore Sun* (Aug. 5, 2013)

“Misunderstanding of Statistics Leads to Misguided Law Enforcement Policies,” *Amstat News* (Dec. 2012)

“An Issue of Numbers,” *National Law Journal* (Mar. 5, 1990)²

B. More extensive treatments of Issue 1 or Issue 2 with respect to the full range of matters to which the issues pertain

Statement of James P. Scanlan 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)

Comments of James P. Scanlan for Commission on Evidence-Based Policymaking (Nov. 14, 2016)

“The Mismeasure of Health Disparities,” *Journal of Public Health Management and Practice* (July/Aug. 2016)

“Race and Mortality Revisited,” *Society* (July/Aug. 2014)

Amicus curiae brief of James P. Scanlan 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)

“Measuring Health and Healthcare Disparities,” Federal Committee on Statistical Methodology 2013 Research Conference (Nov. 2013)

² Explains that lowering National Collegiate Athletic Association academic standards for participation in intercollegiate athletics will tend to increase the proportion African Americans make up of athletes disqualified from participation.

C. Recent articles or blog posts discussing, with respect to certain current issues, government policies or actions based on an understanding of the effects of generally reducing school discipline or criminal justice outcomes on measures of racial disparity that is the opposite of reality

“The misunderstood effects of the Baltimore police consent decree,” *The Daily Record* (Feb. 15, 2018)

“The Misunderstood Relationship Between Racial Differences in Conduct and Racial Differences in School Discipline and Criminal Justice Outcomes,” Federalist Society Blog (Dec. 20, 2017).³

“United States Exports Its Most Profound Ignorance About Racial Disparities to the United Kingdom,” Federalist Society Blog (Nov. 2, 2017)

“The Pernicious Misunderstanding of Effects or Policies on Racial Differences in Criminal Justice Outcomes,” Federalist Society Blog (Oct. 12, 2017).

“Innumeracy at the Department of Education and the Congressional Committees Overseeing It,” Federalist Society Blog (Aug. 24, 2017) *

“The Government’s Uncertain Path to Numeracy,” Federalist Society Blog (July 21, 2017)

D. Web pages on jpscanlan.com

Discipline Disparities page and 41 subpages

Subpages address various issues. About 25 pertain to situations where general reductions in discipline rates were in fact associated with increased relative racial/ethnic differences in discipline rates or where the settings with comparatively low discipline rates had comparatively high relative demographic differences in discipline rates.

Education Disparities page and its 7 subpages

The subpages mainly pertain to research examining demographic differences in educational outcomes in terms of relative differences in the favorable or the adverse outcome, or absolute differences between rates, without consideration of the ways the measures employed tend to be affected by the prevalence of the outcome. That is, researchers failed to understand that general improvements in educational outcomes tend to reduce relative differences in favorable outcomes while increasing relative differences in the corresponding adverse outcomes, or that such improvements tend to increase absolute differences for uncommon outcomes like advanced proficiency but reduce absolute differences for common outcomes like basic proficiency.

³ This item also discusses some complex issues regarding inferences related to likelihood that bias plays a role in racial differences akin to those addressed on the Offense Type Issues subpage of the Discipline Disparities page of jpscanlan.com.

E. Letters to DOE, DOJ, or HHS Regarding School Discipline Issues

Departments of Education, Health and Human Services, and Justice (July 17, 2017)
Department of Justice (Apr. 13, 2017)
Departments of Education and Health and Human Services of Education (Aug. 24, 2015)
Department of Justice (Apr. 23, 2012)
Department of Education (Apr. 18, 2012)

F. Letters to DOE contractors and grantees and other entities that conduct research or provide guidance on research regarding demographic differences in discipline or education outcomes (known DOE contractors/grantees denoted with asterisk)

American Institutes for Research (Aug. 25, 2017) *
Pyramid Equity Project (Nov. 28, 2016) *
University of Oregon Institute on Violence and Destructive Behavior and University of Oregon Law School Center for Dispute Resolution (July 5, 2016) *
University of Oregon Institute on Violence and Destructive Behavior and University of Oregon Law School Center for Dispute Resolution (July 3, 2016) *
New York City Center for Innovation through Data Intelligence (June 6, 2016)
Texas Appleseed (Apr. 7, 2015)
Wisconsin Council on Families and Children's Race to Equity Project (Dec. 23, 2014)
Education Law Center (Aug. 14, 2014)
IDEA Data Center (Aug. 11, 2014) *
Annie E. Casey Foundation (May 13, 2014)
Education Trust (April 30, 2014)

G. Letters to school districts regarding difficulties in their particular situations arising from their own mistaken beliefs, or the mistaken beliefs of others, that generally reducing discipline rates will tend to reduce (a) relative differences in discipline rates or (b) the proportion disadvantaged groups make up of persons disciplined

Metro Nashville Public Schools (Feb. 14, 2018)
Loudoun County Public Schools (Sept. 5, 2017)
Duval County Public Schools (Aug. 2, 2017)
Oklahoma City School District (Sept. 20, 2016)
Antioch Unified School District (Sept. 9, 2016)
Houston Independent School District (Jan. 5, 2016)
McKinney, Texas Independent School District (Aug. 31, 2015)

H. DOE-sponsored documents warranting withdrawal

As suggested in the Recommendations regarding Issue 2, all DOE-sponsored documents measuring or providing guidance on measuring demographic differences in educational outcomes should probably be withdrawn. Those listed below are merely some notable examples.

IDEA Data Center Technical Assistance Guide titled “Methods for Assessing Disproportionality in Special Education (revised March 2014).”⁴

Institute of Education Sciences study titled “Disproportionality in school discipline: An assessment of trends in Maryland, 2009-12” (March 2014).⁵

Institute of Education Sciences/Regional Educational Laboratory guide titled “School discipline data indicators: A guide for districts and schools” (April 2017).⁶

DOE Regulation 24 CFR Part 300 – Assistance to States for the Education of Children with Disabilities; Preschool Grants for Children with Disabilities.⁷

⁴ See the IDEA Data Center Disproportionality Guide subpage of the Discipline Disparities page of jpscanlan.com. See also pages 8-9 of the August 24, 2015 letter to the Secretaries of DOE and HHS.

⁵ This item, which is made available on the DOE “School Climate and Discipline: Know the Data” page and treated on the Maryland Disparities subpage of the Discipline Disparities page of jpscanlan.com, is problematic both because it measures suspension disparities in relative terms and because it reflects the mistaken belief that generally reducing discipline suspension rates would be expected to reduce relative racial differences in suspension rates.

⁶ This item has problems similar to those of the IDEA Data Center Technical Assistance Guide.

⁷ On February 28, 2018, DOE postponed implementation of this regulation until 2020. By then the agency should recognize that one cannot usefully measure demographic based on relative differences in outcome rates (or other measures that tend to change solely because the prevalence of an outcome changes).