

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

August 24, 2015

The Honorable Sylvia M. Burwell,
Secretary of Health and Human Services
United States Department of Health and Human Services
200 Independence Avenue, SW
Washington, DC 20201

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

Re: Joint HHS and DOE “Policy Statement on Expulsion and Suspension Policies in Early Childhood Settings” and the Failure of the Agencies to Understand That Generally Reducing Discipline Rates Tends to Increase Racial and Other Types of Disproportionality in Discipline Rates

Dear Secretaries Burwell and Duncan:

The principal purpose of this letter, in addition to that of explaining to the Department of Health and Human Services and the Department of Education ways in which their missions are generally undermined as a result of the failure to understand certain fundamental statistical principles, is to urge the agencies to withdraw the document, jointly issued by the agencies in December 2014, titled “Policy Statement on Expulsion and Suspension Policies in Early Childhood Settings.” For the document will lead the public and school administrators erroneously to believe that reducing the frequency of preschool expulsions and suspensions will tend to reduce racial and other disproportionality in preschool expulsions and suspensions. In fact, reducing the frequency of those outcomes will tend to increase disproportionality. A secondary purpose of the letter is to alert the agencies that most members of the public would regard the document’s representation that preschool expulsion and suspension rates are “high” to be materially misleading, given that the best evidence indicates that nationally such rates are under 1% yearly.

On occasion I write to institutions or organizations whose activities involve the interpretation of data on demographic differences in the law or the social or medical sciences

alerting them to ways in which their activities are undermined by the failure to understand patterns by which standard measures of differences between favorable or adverse outcome rates of advantaged and disadvantaged groups – or differences between the proportion a group comprises of persons potentially experiencing an outcome and the proportion it comprises of persons actually experiencing the outcome – tend to be systematically affected by the overall frequency of an outcome. Recipients of other letters involving measurement issues of the type discussed in this letter include [Robert Wood Johnson Foundation](#) (Apr. 8, 2009),¹ [National Quality Forum](#) (Oct. 22, 2009), [Institute of Medicine](#) (June 1, 2010), [The Commonwealth Fund](#) (June 1, 2010), [United States Department of Education](#) (Apr. 18, 2012), [United States Department of Justice](#) (Apr. 23, 2012), [Board of Governors of the Federal Reserve System](#) (Mar. 4, 2013), [Harvard University](#) (Oct. 9, 2012), [Harvard Medical School, Massachusetts General Hospital, et al.](#) (Oct. 26, 2012), [Senate Committee on Health, Education, Labor and Pensions](#) (Apr. 1, 2013), [Mailman School of Public Health of Columbia University](#) (May 24, 2013), [Investigations and Oversight Subcommittee of House Finance Committee](#) (Dec. 4, 2013), [Education Trust](#) (April 30, 2014), [Annie E. Casey Foundation](#) (May 13, 2014), [Institute of Medicine II](#) (May 28, 2014), [IDEA Data Center](#) (Aug. 11, 2014), [Education Law Center](#) (Aug. 14, 2014), [Financial Markets and Community Investment Program, Government Accountability Office](#) (Sept. 9, 2014), [Wisconsin Council on Families and Children’s Race to Equity Project](#) (Dec. 23, 2014), [Portland, Oregon Board of Education](#) (Feb. 25, 2015), [Vermont Senate Committee on Education](#) (Feb. 26, 2015), [United States Department of Justice and City of Ferguson, Missouri](#) (Mar. 9, 2015), [Senate Committee on Health, Education, Labor and Pensions II](#) (Mar. 20, 2015), [Texas Appleseed](#) (Apr. 7, 2015), [City of Minneapolis, Minnesota](#) (June 8, 2015), and [Agency for Healthcare Research and Quality](#) (July 1, 2015). [Comments](#) I posted on May 16, 2014, regarding Federal Drug Administration proposed subgroup regulations may be deemed a similar communication to that agency, and an *amicus curiae* [brief](#) I filed on November 17, 2014, in *Texas Department of Housing and Community Development, et al. v. The Inclusive Communities Project, Inc.*, Sup. Ct. No. 13-1371 (TDHCD brief), might be deemed a similar communication to the United States Supreme Court.

This letter is immediately prompted by my learning that the Department of Education (DOE) recently launched a “[Rethinking Discipline Campaign](#)” and my review of the DOE and Department of Health and Human Services (HHS) December 10, 2014 document titled “[Policy Statement on Expulsion and Suspension Policies in Early Childhood Settings](#)” (Policy Statement) and the associated [Dear Colleague Letter](#) by the Secretaries of the two agencies. This letter may be deemed a follow-up to the above-mentioned April 12, 2012 letter to DOE (addressed to Secretary Duncan and Assistant Secretary for Civil Rights Russlynn Ali), which explained that, contrary to the understanding of the agency and things it was generally leading the public and school administrators to believe, relaxing public school discipline standards tends to increase (not decrease) relative racial/ethnic differences in discipline rates.² A number of the

¹ To facilitate consideration of issues raised in letters such as this I include links to referenced materials in electronic copies of the letters. Such copies may be found by means of the Institutional Correspondence subpage of the Measuring Health Disparities page of [jpscanlan.com](#).

The Honorable Sylvia M. Burwell
The Honorable Arne Duncan
August 24, 2015
Page 3

other letters mentioned above, including the April 23, 2012 letter to the Department of Justice (DOJ) and the April 1, 2013, and March 20, 2015 letters to the Senate Committee on Health, Education, Labor and Pensions, also addressed that specific issue. But the above-mentioned letter most pertinent to the subject of this letter is the August 11, 2014 letter to the [IDEA Data Center](#), which involves the DOE-funded guide titled “[Methods for Assessing Racial/Ethnic Disproportionality in Special Education](#).” The IDEA Data Center letter and the guide will be discussed further below.

In a limited sense, this letter may also be deemed a follow-up to the above-mentioned July 1, 2015 letter to the Agency for Healthcare Research and Quality (AHRQ). That letter addressed the failure of AHRQ to understand the ways that measures it employed to appraise health and healthcare disparities tend to be systematically affected by the frequency of an outcome, as well as an anomaly whereby the National Healthcare Disparities Report (NHDR) finds substantial decreases in disparities for a healthcare outcome between two points in time even though it would also find a substantially larger disparity for the outcome at the end of the period than the beginning of the period. That letter also discusses failures of understanding of measurement issues at all HHS agencies involved in health and healthcare disparities research. But, since this letter principally involves school discipline issues, I will give the health and healthcare disparities issues only limited attention below.

As with the earlier DOE letter, much of the discussion below involves the pattern 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 and the fact that, consistent with that pattern, relaxing standards and otherwise generally reducing discipline rates will tend to increase relative differences in discipline rates while reducing relative differences in rates of avoiding discipline. That discussion will include explaining that seemingly stark racial disproportionality in preschool suspensions is a function of the fact that suspensions are very rare in preschool. In light of such focus, it is necessary initially to point out a substantial problem with the statement about frequency of preschool expulsions and suspensions in the first paragraph of the Policy Statement. While the statement seems to provide the rationale for making preschool discipline a priority issue that may cause the expenditure of many millions of dollars nationally by preschool authorities attempting to follow guidance provided by HHS and DOE, most people aware of the actual rates would consider the statement to be materially misleading.

Specifically, the second sentence of the Policy Statement states (at 1, footnotes 2, 3, 4 omitted): “Recent data indicate that expulsions and suspensions occur at high rates in preschool settings.”³ The document, however, never states what these purportedly high rates are.

² The Department replied by [letter](#) dated May 4, 2012, from the Customer Service Team of the Office of Civil Rights. The reply attempted to summarize the content of my letter. But it does not appear that the drafter of the reply in fact understood the statistical issues addressed in my letter.

³ The Dear Colleague Letter similarly states in its second sentence that “[r]ecent data indicate that expulsions and suspensions occur with regularity in preschool settings.”

The first reference cited for the quoted sentence (note 2 of the Policy Statement) is a 2005 study,⁴ which was limited to expulsions. The study found, based on a national sample that seems to involve either the 2001-02 or 2002-03 school year, a preschool expulsion rate of 0.67%. The second reference (note 3 of the Policy Statement) is DOE's March 2014 report on preschool discipline,⁵ which provided information only on suspensions. The report found a national preschool suspension rate of 0.75%.⁶ The third reference (note 4 of the Policy Statement) is a 2006 study⁷ that found that during the 1999-2000 school year, Massachusetts had a preschool expulsion rate of 2.7% and a preschool suspension rate of 1.3%.

Few would regard the under 1% national expulsion and suspension rates found in the 2005 study and the DOE report as high, and many readers of the Policy Statement would be quite surprised to learn that the rates the Policy Statement described as high were in fact this low. But even if the agencies were to maintain that rates of under 1%, or the somewhat higher rates found in Massachusetts in one school year, should properly be deemed high,⁸ the Policy Statement ought to have made clear to the readers the exact nature of the rates that the agencies were regarding as high.

Rather than doing so, however, when the Policy Statement returns to the subject of high rates (at 3-4), it cites the 2005 study's finding that "over 10% of preschool teachers in state-

⁴ Gilliam, W. S. (2005). [Prekindergarteners left behind: Expulsion rates in state prekindergarten systems](#). New York, NY: Foundation for Child Development.

⁵ U.S Department of Education Office for Civil Rights (2014). [Data Snapshot: Early Childhood Education](#).

⁶ The 0.75% rate is based on the note to the second chart on page 3 of the report, which indicates that there were approximately one million preschool students and that nearly 5,000 students were suspended once and over 2,500 students were suspended more than once. Inasmuch as the report indicated (at 3) that nearly 10,000 school districts had preschool programs, presumably a substantial proportion of districts, and a very substantial proportion of the individual preschool programs within districts, had no preschool suspensions whatever.

⁷ Gilliam, W.S., & Shahar, G. (2006). [Preschool and child care expulsion and suspension: Rates and predictors in one state](#). *Infants & Young Children*, 19, 228-245.

⁸ I note, however, that in other discipline contexts, DOE has regarded such rates to be low rather than high. In a January 8, 2014 press release titled "[Rethinking School Discipline](#)," the DOE makes the following statement:

And we know that discipline policy and practices matter tremendously—there is nothing inevitable about high rates of suspension and expulsion. We can, and must, do much better.

According to CRDC data, schools in South Carolina suspended 12.7 percent of students—about one in eight students during the 2009-10 school year. By contrast, schools in North Dakota suspended 2.2 percent of students—about one out of every 50 students.

Thus, the agency cited a 2.2% suspension rate as an example of a low suspension rate.

funded prekindergarten programs reported expelling at least one preschooler in the past year.”⁹ Few readers would immediately recognize that a situation where 10% of teachers expelled at least one student would be consistent with expulsion rates of under 1%. Indeed, the reference to 10% might well lead many readers to think that expulsion rates were in the 10% range.

Thus, particularly in a context where the rareness of the outcome is so pertinent to understanding seemingly high racial disproportionality (as discussed below), this misleading aspect of the Policy Statement provides additional reason for its withdrawal. The key reason for withdrawal, however, rests in the fact that the Policy Statement, like other documents issued by DOE, leads the public and school administrator erroneously to believe that reducing the frequency of expulsions and suspensions will tend to reduce disproportionality.

Since the earlier letter to DOE, I have treated the agency’s misunderstanding of this issue in a number of published articles. These include: “[Racial Differences in School Discipline Rates](#),” *The Recorder* (June 22, 2012); “[Misunderstanding of Statistics Leads to Misguided Law Enforcement Policies](#),” *Amstat News* (Dec. 2012); “[The Paradox of Lowering Standards](#),” *Baltimore Sun* (Aug. 5, 2013); “[Things government doesn’t know about racial disparities](#),” *The Hill* (Jan. 28, 2014); and “[Race and Mortality Revisited](#),” *Society* (July/Aug. 2014). I have also treated the DOE’s misunderstanding in workshops given at arms of various universities, including [American University](#) (2012), [Harvard University](#) (2012), [University of Kansas School of Law](#) (2013), [University of Minnesota](#) (2014), [University of Maryland](#) (2014),¹⁰ [George Mason University](#) (2014), and [University of California, Irvine](#) (2015). These workshops provide many graphical and tabular illustrations of the statistical patterns mentioned in the earlier letter and only touched upon here.

Each of the referenced articles uses the same example as that in earlier DOE letter to illustrate that lowering a test cutoff, while tending to reduce relative differences in pass rates, will tend to increase relative differences in failure rates. The example is based on a situation where the means of normal test score distributions of an advantaged group (AG) and a disadvantaged group (DG) differ by half a standard deviation and both distributions have the same standard deviation. In such circumstances, at the cutoff where 80% of AG passes the test, 63 percent of DG would pass the test (with corresponding failure rates of 20% for AG and 37%

⁹ I exclude from the quotation the following language: “a rate more than three times higher than estimates for teachers of K-12 public school students.” I note for clarity that the study actually found that the rate of expulsions in preschool was 3.2 times as high as in K-12, not that the proportion of teachers reporting one or more expulsions was 3.2 times as high in preschool as in K-12. See page 3 of the study referenced in note 4 *supra*. (I use the phrase “times as high” rather than “times higher” for reasons discussed on the [Times Higher](#) subpage of the [Vignettes](#) page of [jpscanlan.com](#) (*i.e.*, the latter usage is incorrect and often misleads). The distinction, however, is not pertinent to the error in the Policy Statement.)

¹⁰ Subsequent references to materials in these workshops will mainly be to the University of Maryland workshop (though similar materials will be found in the other workshops). The Maryland workshop was attended by representative of two agencies of HHS (the Agency for Healthcare Research and Quality and the Food and Drug Administration). I invited employees of the DOE to the workshop, but none attended.

for DG); the ratio of AG's pass rate to DG's pass rate would thus be 1.27 while the ratio of DG's fail rate to AG's fail rate would be 1.85.¹¹ When the cutoff is lowered to the point where the pass rate for AG is 95%, the pass rate for DG would be 87% (with corresponding failure rates of 5% for AG and 13% for DG); the ratio of AG's pass rate to DG's pass rate would thus decrease to 1.09, while the ratio of DG's fail rate to AG's fail rate would increase to 2.60.

These figures are set out in Table 1 below, which is the same as Table 1 of the above-mentioned April 1, 2013 letter to the Senate Committee on Health, Education, Labor and Pensions and similar to Table 1 of "Race and Mortality Revisited," an article to which I will refer further below with respect to the size of racial differences in preschool suspensions and other matters.

Table 1. Illustration of effects on relative differences in pass and fail rates of lowering a cutoff from a point where 80% of the advantaged (higher-scoring) group passes to a point where 95% of the advantaged group passes (when mean scores differ by approximately half a standard deviation)

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

As with the earlier DOE letter, the referenced articles commonly explain that discipline standards operate like test cutoffs. Relaxing the standards, while tending to reduce relative difference in rate of avoiding discipline (the equivalent of passing the test), tends to increase relative differences in discipline rates (the equivalent of failing the test).

The articles also commonly highlight the following law enforcement anomaly. The Departments of Education and Justice have for some time been encouraging public schools to relax standards in order to reduce relative racial/ethnic differences in discipline rates. Unaware that relaxing standards in fact tends to increase relative differences in discipline rates, the agencies continue to monitor the fairness of practices on the basis of relative differences in discipline rates. Thus, by complying with DOE and DOJ encouragements to relax standards, schools districts increase the chances that the agencies will find them to have violated civil rights laws.¹²

¹¹ While I commonly refer to patterns of relative differences in this letter, the table actually presents rate ratios. 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. One should be careful not to mistakenly refer to the rate ratio as the relative difference. But the distinction between the two terms is not pertinent to the discussion here of patterns by which relative differences tend to be affected by the frequency of an outcome.

¹² As discussed in several of the articles, a similar anomaly has existed in the enforcement of fair lending laws for over two decades. Since the early 1990s, DOJ and other federal agencies have been encouraging lenders to relax lending standards, while mistakenly believing that doing so will tend to reduce relative differences in adverse borrower outcomes like rejection of mortgage applications. See also the above-mentioned TDHCD brief and the

There have been several other developments regarding this subject since the earlier letter. In November 2012 the DOE's Office of Civil Rights issued a document titled "[Helping to Ensure Equal Access to Education: Report to the President and Secretary](#)." As discussed on the [DOE Equity Report](#) subpage of the [Discipline Disparities](#) page of [jpscanlan.com](#), data in the report show 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.

Further, since the earlier letter, information has become available (or I have discovered data that were already available) consistently showing that recent general reductions in discipline rates in schools across the country have been accompanied by increased relative racial/ethnic differences in discipline rates. These data are discussed on the following subpages of the Discipline Disparities page (with jurisdiction indicated in the title of the subpage): [Los Angeles SWPBS](#), [Denver Disparities](#), [Florida Disparities](#), [Maryland Disparities](#),¹³ [California Disparities](#), [Connecticut Disparities](#), [Maryland Disparities](#), [Minnesota Disparities](#), [Rhode Island Disparities](#), [St. Paul Disparities](#), [Minneapolis Disparities](#), [Beaverton \(OR\) Disparities](#), [Portland \(OR\) Disparities](#), [Montgomery County \(MD\) Disparities](#), and [Henrico County \(VA\) Disparities](#).

Racial differences in preschool suspension rates first received substantial publicity upon DOE's issuance of its March 2014 report on preschool discipline. The report included information such as that African Americans make up 18% of preschoolers but 48% of preschoolers suspended more than once. Many observers regarded such facts as reflecting huge disparities.¹⁴ But observers failed to recognize that the size of relative differences in suspensions or multiple suspensions (of which the proportion disadvantaged groups comprise of persons suspended is a function, as discussed below) in preschool was a consequence of the fact that suspensions and multiple suspensions are very rare in preschool.

I treated this subject in Table 8 (at 342) of "Race and Mortality Revisited," which table is reproduced in part as Table 2 below. The table is based on the multiple suspension rates for African Americans and whites that can be derived from information in the DOE March 2014 report. The table shows the common pattern whereby in a setting where the outcome is less common than in another setting (*i.e.*, preschool versus K-12), relative differences in experiencing

letters to the Board of Governors of the Federal Reserve System (Mar. 4, 2013), the Investigations and Oversight Subcommittee of House Finance Committee (Dec. 4, 2013), the Financial Markets and Community Investment Program, Government Accountability Office (Sept. 9, 2014), as well as the articles collected on the [Lending Disparities](#) page of [jpscanlan.com](#), including "[The Perverse Enforcement of Fair Lending Laws](#)," Mortgage Banking (May 2014)

¹³ The Maryland Disparities subpage is based on the March 2014 Institute for Education Sciences study [Disproportionality in School Discipline: An Assessment of Trends in Maryland, 2009–12](#) highlighted on the DOE's [School Climate and Discipline: Know the Data](#) page.

¹⁴ The Policy Statement highlights these figures at page 4 but mistakenly states that the figures are for African American boys. Page 3 of the DOE report makes clear that the figures are for all African Americans.

the outcome are larger, while relative differences in avoiding the outcome are smaller, than in the setting where the outcome is more common.

Table 2. White and black multiple suspension rates in preschool and K-12, with measures of difference

Level	White rate multiple suspensions	Black rate multiple suspensions	B/W ratio multiple suspensions	W/B ratio no multiple suspension
Preschool	0.15%	0.67%	4.41	1.01
K12	2.23%	6.72%	3.01	1.05

The table in the article also shows that, to the extent that we can effectively measure the strength of the forces causing black and white multiple suspension rates to differ, such strength is essentially the same in preschool as in K-12. I have not included that information here because it involves a subject beyond the scope of this letter. But the contrasting patterns of the comparative size of the relative difference in the two outcomes in preschool and K-12 shown in the table are simply what one with a sound understanding of statistics would expect in the circumstances.

While the discussion above is largely cast in terms of relative differences between outcome rates, the Policy Statement is largely cast in terms of comparisons of the proportion certain groups make up of students and the proportion such group's make up of persons experiencing the adverse outcome (though without discussing how one might measure the difference between those two proportions). But a corollary to the pattern whereby reducing the frequency of an outcome tends to increase the relative difference in rates of experiencing the outcome while reducing relative differences in rates of avoiding the outcome is a pattern whereby reducing the frequency of an outcome tends to increase the proportion groups most susceptible to the outcome make up of both (a) persons experiencing the outcome and (b) persons failing to experience the outcome. Thus, reducing the frequency of an outcome will tend to increase disproportionality regarding the outcome, regardless of whether the disproportionality is measured in terms of the relative difference between those proportions or the absolute difference between those proportions.¹⁵

I discuss and illustrate these patterns, with reference to measurement approaches recommended in the above-mentioned DOE-funded "Methods for Assessing Racial/Ethnic Disproportionality in Special Education," in the above-mentioned May 28, 2014 letter to the IDEA Data Center, and in the [IDEA Data Center Disproportionality Guide](#) subpage of the Discipline Disparities page, as well in slides 98 to 117 of the University of Maryland workshop.¹⁶ These references also address additional problems with efforts to appraise the

¹⁵ Reducing the frequency of the outcome will, however, tend to reduce relative and absolute differences between the proportion groups most susceptible to the outcome make up of persons potentially experiencing the outcome (or its opposite) and the proportion such groups make up of persons failing to experience the outcome.

¹⁶ Simpler illustrations of the way that reducing the frequency of an outcome tends to cause groups most susceptible to the outcome to make up higher proportions of both (a) persons experiencing the outcome and (b) persons failing

strength of an association based on the proportion a group comprises of persons potentially experiencing an outcome and the proportion it comprises of persons experiencing the outcome. But, given that, as with relative differences in either the favorable or the adverse outcome, neither the relative difference nor the absolute difference between the two proportions just referenced is a sound measure of association, it is unnecessary to address those problems here.

I add one final point pertinent solely to discipline issues. While this letter focuses on measurement issues, I would be remiss in failing to note that the part of the body of research with which I am familiar that addresses the consequences of stringent discipline policies is of extremely dubious validity. I discuss this issue with regard to the American Psychological Association (APA) Zero Tolerance study referenced in note 7 of the Policy Statement on the [APA Zero Tolerance Study](#) subpage of the Discipline Disparities page. The criticisms of the APA study on that subpage would apply to essentially all statements in the Policy Statement regarding the correlations of suspensions with adverse outcomes. That is, the conclusions fail to consider that students who engage in conduct that leads to suspensions tend to have characteristics associated with adverse educational outcomes irrespective of suspensions. Hence, such students will tend to have higher rates of adverse outcomes than other students regardless of the effects of suspensions or other disciplinary measures. That it may be difficult or impossible to determine whether disciplinary measures caused the adverse outcome rates to be higher than they otherwise would be is not a reason to treat the higher adverse outcome rates of disciplined students as meaningful evidence of the deleterious consequences of discipline practices. Similar issues exist with respect to interpretations of data on correlations of frequency of suspensions with school environmental factors. Thus, I suggest that the agencies give careful consideration to the points made on the referenced subpage in formulating policies relating to school discipline.

The concluding paragraphs of the earlier DOE letter (at 4) called the agency's attention to the [Educational Disparities](#) page of [jpscanlan.com](#). In doing so, the letter noted that the statistical issues it addressed regarding relative differences in favorable and adverse outcomes would apply to various matters within the purview of DOE, including those involving demographic differences in proficiency rates. The letter pointed out, for example, that improvements in overall performance on proficiency exams would tend to reduce relative differences in proficiency rates while increasing relative differences in non-proficiency rates.¹⁷

to experience the outcome may be found in Table 1 of "[Divining Difference](#)," *Chance* (Fall 1994) and Table 1 of "[Can We Actually Measure Health Disparities?](#)," *Chance* (Spring 2006). Assuming that the advantaged group and the disadvantaged group each made up 50% of test takers, the figures in Table 1 of this letter would mean that lowering the cutoff would cause the proportion the disadvantaged group makes up of persons failing the test to increase from 65% to 72% and the proportion the disadvantaged group makes up of persons passing the test to increase from 44% to 48%.

¹⁷ The penultimate paragraph of the earlier letter (at 4) discussed issues addressed on the [Relative Versus Absolute](#) subpage of the [Measuring Health Disparities](#) page of [jpscanlan.com](#). Those issue have more recently been addressed in "Race and Mortality Revisited" at 335-336.

Since the earlier letter, I have given a good deal more attention to proficiency disparities issues, especially with regard to the common practice of analyzing racial/ethnic differences in proficiency rates in terms of absolute (percentage point) differences between rates. The Educational Disparities page now has seven subpages addressing problematic appraisals of changes in the size of disparities in reaching or failing to reach certain proficiency levels without consideration of the way the measure employed is systematically affected by changes in the frequency of the outcome. Such is also the subject of the above-mentioned April 30, 2014 letter to the Education Trust and the May 13, 2014 letter to the Annie E. Casey Foundation.

In addition, the issues discussed above play importantly into efforts to determine whether observed differences in outcome rates, whether with respect to discipline or anything else, are the result of biased decision making. I have discussed this subject in the above-mentioned TDHCD brief, in certain portions of the "Race and Mortality Revisited," and in the University of California, Irvine and University of Kansas School of Law workshops, as well as various other places.

The following is one example of the implications of the failure to understand these issues that is pertinent to the school discipline context. Statements to the effect that white students tend to be punished more often for objective offenses while black students tend to be punished more often for subjective offense play importantly into contentions that racial bias is responsible for a substantial part of observed racial differences in discipline rates. But implications of the pattern underlying such (commonly ambiguous) statements – which pattern is presumably that relative differences in discipline rates are greater for subjective offenses than for objective offenses – cannot be analyzed without an understanding of role of the frequency of the outcomes. See the [Offense Type Issues](#) subpage of the Discipline Disparities page. More generally, as discussed especially at page 339-341 of "Race and Mortality Revisited," one can never draw a sound inference about processes based on either the comparative size of two relative differences in a favorable outcome or the comparative size of two relative differences in the corresponding adverse outcome without a firm understanding of the patterns described in that article.

I therefore encourage the DOE to have its staff carefully review the Educational Disparities and Discipline Disparities pages and all their subpages, as well as the articles and workshop materials mentioned above. I especially encourage that review with regard to the agency's ongoing consideration of the Government Accountability Office's recommendation that the agency provide states more specific guidance on determinations of significant disproportionality in special education, the subject on which the DOE issued a [Request for Information](#) in June 2014. The IDEA Data Center guide mentioned above, though not leading observers erroneously to believe that reducing the frequency of a putative adverse outcome will tend to reduce relative differences in rates of experiencing the outcome, nevertheless shows no awareness of the way the frequency of the outcome tends to affect each of the measures the

guide recommends.¹⁸ As with other guides that fail to reflect such awareness, the guide cannot provide useful instruction on the appraisal of the strength of the forces causing outcome rates of advantaged and disadvantaged groups to differ, and necessarily will commonly lead users to believe things that are not true. See the above-mentioned October 26, 2012 letter to Harvard Medical School and Massachusetts General Hospital (and others) regarding their jointly produced *Commissioned Paper: Health Care Disparities Measurement* and the discussion of that document in "Race and Mortality Revisited" at 344-345.

In considering the value of such a review to the agency, please keep in mind that so far in the DOE's history, the agency has not merely failed to show any understanding that the measures it has employed for appraising demographic differences tend to be systematically affected by the frequency of an outcome or even to show an awareness that lowering a test cutoff (or improving test performance) will tend to increase relative differences in failure rates while reducing relative differences in pass rates. As a result of the agency's failure to understand the effects of relaxing discipline standards on relative differences in discipline rates, the agency has also affirmatively misled public school authorities (as well as local and state governments) all across the country to believe that relaxing standards will tend to reduce relative racial/ethnic differences in discipline rates. In such circumstances, I submit, the agency should employ every available avenue to educate its staff on these issues and to reform its practices in accordance with an improved understanding of the issues.

An important first step in such reform is the withdrawal of the Policy Statement and all other documents that suggest that relaxing standards or otherwise reducing the frequency of adverse discipline outcomes will tend to reduce relative differences between rates of experiencing those outcomes or reduce the disproportionality that is the function of those relative differences. But there are many aspects of agency policies that warrant reconsideration in light of the statistical issues discussed above and addressed more fully in referenced materials.

I mentioned at the outset that in a limited sense, this letter may be deemed a follow-up to the July 1, 2015 letter to AHRQ. As reflected in the AHRQ letter and "Race and Mortality Revisited" – and addressed at greater length my Federal Committee on Statistical Methodology 2013 Research Conference paper "[Measuring Health and Healthcare Disparities](#)" – the same failures of understanding that have undermined DOE efforts to appraise demographic differences in discipline rates have undermined virtually all efforts of federal agencies to appraise health and healthcare disparities. I will be communicating further with HHS or various of its agencies regarding such matters (as well as the subgroup effect issues that are the subject of the above-mentioned May 16, 2014 comments on Federal Drug Administration proposed subgroup regulations). But for the present I will leave the addressing of those matters to the items just mentioned and other materials previously created, save for the following two points.

¹⁸ The guide also shows no awareness of other problems with certain of those measures, which problems were alluded to several paragraphs above.

First, each of the three items mentioned in the preceding paragraph discusses, sometimes at great length, the disarray among federal agencies involved with health and healthcare research. Such disarray includes a failure of various agencies to realize that other agencies are using different measures of disparities and that such different measures tend to yield (or have in fact yielded) opposite conclusion about such things as direction of changes in disparities over time from those yielded by the measures employed by the first agencies. I add to such discussion the point that in 2015, the Center for Medicare and Medicaid Services issued the [2015 National Impact Assessment of the Centers for Medicare & Medicaid Services \(CMS\) Quality Measures Report](#). The document read the 2013 National Healthcare Disparities Report of AHRQ as identifying disparities on the basis of a 10 percentage point difference between rates (and adopted that approach, though using a 5 percentage point standard). In fact, the NHDR identifies a disparity on the basis of a 10% (i.e., relative) difference between adverse outcome rates (although, as discussed in the AHRQ letter, in 2010 the NHDR began to measure changes in disparities in terms of the comparative size of absolute changes in rates without AHRQ's recognizing the inconsistency of that approach with its reliance on relative differences in adverse outcomes).

Second, all three of the items mentioned two paragraphs above give attention to the fact that the National Center for Health Statistics (NCHS), while recognizing since 2004 that relative differences in favorable health and healthcare outcomes and relative differences in the corresponding adverse outcomes tend to change in opposite directions as the frequency of an outcome changes, did not appreciate that such pattern called into question the utility of either relative difference for appraising the strength of the forces causing outcome rates of advantaged groups and disadvantaged groups to differ. Rather, NCHS simply sought to address the matter by measuring all health and healthcare disparities in terms of relative differences in adverse outcomes. That course caused many healthcare disparities that previously were deemed to be decreasing now to be deemed to be increasing, but it did not address the fundamental measurement problem. It is my understanding that NCHS is now intending, for purposes of Healthy People 2020, to measure healthcare disparities in terms of relative differences in favorable outcomes. Such intent is already reflected in an online document titled "[Health Disparities Tool: A User's Guide](#)" (and associated online materials). Such course will tend to reverse the pattern that has existed since 2004 whereby improvements in healthcare tended to be associated with increasing healthcare disparities to a pattern whereby improvements in healthcare will tend to be associated with decreasing disparities. But, while creating additional disarray within the government, the contemplated NCHS action will by no means address the underlying measurement issues.

For all the reasons stated above, I urge DOE and HHS to withdraw the Policy Statement and urge DOE to withdraw all documents that would lead readers to believe that reducing the frequency of disciplinary actions will tend to (a) reduce relative differences in rates of

The Honorable Sylvia M. Burwell
The Honorable Arne Duncan
August 24, 2015
Page 13

experiencing those outcomes or (b) reduce the proportion disadvantaged groups make up of persons experiencing those outcomes.

Sincerely,

/s/ James P. Scanlan

James P. Scanlan

[This version of the letter contains a correction of the letter originally sent, specifically, elimination of a line in the address for Secretary of Education Arne Duncan.]