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

June 6, 2016

Maryanne Schretzman, Executive Director
Center for Innovation through Data Intelligence
W. Cyrus Garrett, Executive Director
NYC Young Men's Initiative
City Hall
New York, NY 10007

Re: Methodological Problems in Center for Innovation through Data Intelligence's 2016 Disparity Report

Dear Executive Directors Schretzman and Garrett:

The purpose of this letter is to bring to the attention of the Center for Innovation through Data Intelligence (CIDI) and the NYC Young Men's Initiative methodological problems in the CIDI's 2016 *Disparity Report*[1] arising from the report's failure to recognize patterns by which measures tend to be affected by frequency of an outcome. In particular, the report fails to recognize the pattern whereby the rarer an outcome, the greater tends to be the relative difference between rates of experiencing the outcome and the smaller tends to be the relative difference between rates of avoiding the outcome. Thus, for example, the report fails to recognize that general improvements (increases) in academic proficiency or graduation rates will tend to reduce relative racial differences in proficiency and graduation rates while increasing relative racial differences in rates of failing to achieve proficiency and failing to graduate, or that general improvements (decreases) in dropout rates will tend to increase relative differences in dropout rates while reducing relative differences in rates of avoiding dropping out. Consequently, the report, which measures disparities in terms of relative differences in favorable outcomes for some matters and relative differences in adverse outcomes for other matters, provides nothing that is statistically sound regarding whether differences in the circumstances of racial groups have increased or decreased over time.

This letter is similar in certain respects to letters to a variety of other institutions or organizations whose activities involve, or are affected by, the interpretation of data on group differences. Among the more comprehensive of such letters is that of October 8, 2015, to the American Statistical Association (ASA) urging the organization, among other things, to explain

[1] To facilitate consideration of issues raised in letters such as this by their recipients and others 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.

to the federal government that reducing the frequency of adverse outcomes like suspension from school tends to increase, not decrease, relative differences in rates of experiencing such outcomes. That aspect of the ASA letter is of special pertinence to the subject of Section B of this letter. But it, and the other letters to institutions and organizations,[2] also illustrate how universal is the misunderstanding of the statistical issues addressed here. That misunderstanding is also reflected in the RAND Corporation report *Reparable Harm: Assessing and Addressing Disparities Faced by Boys and Men of Color in California* that the CIDI report indicates served to inform and inspire it. Thus, while I would remiss in failing to stress that none of the CIDI report's conclusions about the changes in the size of disparities has a sound statistical basis, I also note that the analytical problems identified here and in the ASA letter apply to the work of essentially all institutions and organizations analyzing demographic differences in outcome rates.

In addition to the ASA letter, extended treatments of statistical issues pertinent to the subject of this letter may be found in my "Race and Mortality Revisited," *Society* (July/Aug. 2014), "Measuring Health and Healthcare Disparities," Federal Committee on Statistical Methodology 2013 Research Conference (March 2014), and "The Mismeasure of Discrimination," Faculty Workshop, University of Kansas School of Law (Sept. 2013).[3] The substantial discussion in "Race and Mortality Revisited" and "Measuring Health and Healthcare Disparities" of the recognition by the National Center for Health Statistics 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, and the agency's belief that it could satisfactorily address the matter simply by measuring all health and healthcare disparities in terms of relative differences adverse outcomes,

---

[2] Recipients of such letters include: City of Madison, Wisconsin (Mar. 12, 2016), Stanford Center on Poverty and Inequality (Mar. 8, 2016), City of Boulder, Colorado (Mar. 5, 2016), Houston Independent School District (Jan. 5, 2016), Boston Lawyers' Committee for Civil Rights and Economic Justice (Nov. 12, 2015), House Judiciary Committee (Oct. 19, 2015), Chief Data Scientist of White House OSTP (Sept. 8, 2015), McKinney, Texas Independent School District (Aug. 31, 2015), Department of Health and Human Services and Department of Education (Aug. 24, 2015), Agency for Healthcare Research and Quality (July 1, 2015), City of Minneapolis, Minnesota (June 8, 2015), Texas Appleseed (Apr. 7, 2015), Senate Committee on Health, Education, Labor and Pensions (Mar. 20, 2015), United States Department of Justice and City of Ferguson, Missouri (Mar. 9, 2015), Vermont Senate Committee on Education (Feb. 26, 2015), Portland, Oregon Board of Education (Feb. 25, 2015), Wisconsin Council on Families and Children's Race to Equity Project (Dec. 23, 2014), Financial Markets and Community Investment Program, Government Accountability Office (Sept. 9, 2014), Education Law Center (Aug. 14, 2014), IDEA Data Center (Aug. 11, 2014), Institute of Medicine II (May 28, 2014), Annie E. Casey Foundation (May 13, 2014), Education Trust (April 30, 2014), Investigations and Oversight Subcommittee of House Finance Committee (Dec. 4, 2013), Mailman School of Public Health of Columbia University (May 24, 2013), Senate Committee on Health, Education, Labor and Pensions (Apr. 1, 2013), Federal Reserve Board (March 4, 2013), Harvard University et al. (Oct. 26, 2012), Harvard University (Oct. 9, 2012), United States Department of Justice (Apr. 23, 2012), United States Department of Education (Apr. 18, 2012), The Commonwealth Fund (June 1, 2010), Institute of Medicine (June 1, 2010), National Quality Forum (Oct. 22, 2009), Robert Wood Johnson Foundation (Apr. 8, 2009).

[3] Older extended treatments may be found in "Can We Actually Measure Health Disparities?," *Chance* (Spring 2006), "Race and Mortality," *Society* (Jan./Feb. 2000), and "The Perils of Provocative Statistics," *Public Interest* (Winter 1991).

should be considered in light of the agency's recent reversal of that position with regard to healthcare disparities, as discussed in my "The Mismeasure of Health Disparities," *Journal of Public Health Management & Practice* (July/Aug. 2016).

More succinct treatments of the principal statistical issue addressed here may be found in my "Things government doesn't know about racial disparities," *The Hill* (Jan. 28, 2014), "The Paradox of Lowering Standards," *Baltimore Sun* (Aug. 5, 2013), and "Misunderstanding of Statistics Leads to Misguided Law Enforcement Policies," *Amstat News* (Dec. 2012). Extensive graphical and tabular illustrations of the pertinent patterns may be found in methods workshops given at arms of American universities between 2012 and 2015.[4]

Section A of this letter addresses the general implications of the failure of the CIDI report to recognize the ways that the measures it employs tend to be affected by the frequency of an outcome. Section B addresses certain issues pertinent to the report's treatment of disparities in public school suspensions.

A. The Failure of the CIDI Disparity Report to Recognize the Pattern by Which the Measures it Employs Tend to Be Affected by the Frequency of an Outcome.

For reasons related to the shapes of underlying risk distributions, all standard measures of differences between outcome rates tend to be systematically affected by the frequency of an outcome. The pattern of such effects most pertinent to the subject of this letter 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.

The pattern can be illustrated with virtually any data showing points on a continuum of quantifiable factors associated with the likelihood of experiencing or avoiding some outcome. A simple illustration of the pattern is set out in Table 1. The table 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. The table presents pass and fail rates for each group at two cutoff points, along with

[4] See "The Mismeasure of Health Disparities in Massachusetts and Less Affluent Places," Department of Quantitative Health Sciences, University of Massachusetts Medical School (Nov. 18, 2015); "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); "Rethinking the Measurement of Demographic Differences in Outcome Rates," Maryland Population Research Center of the University of Maryland (Oct. 10, 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).

the ratio of AG's pass rate to DG's pass rate and the ratio of DG's fail rate to AG's fail rate at each cutoff.

Table 1. Illustration of effects on relative differences in pass rates and relative differences in failure rates of lowering a cutoff from a point where 80% of AG passes to a point where 95% of AG passes (when mean scores differ by approximately half a standard deviation), with the higher rate used as the numerator in both rate ratios

Cutoff	AG Pass	DG Pass	AG Fail	DG Fail	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

At the higher cutoff, AG's pass rate (80%) is 1.27 times (27% greater than) [5] DG's pass rate (63%). At the lower cutoff, AG's pass rate (95%) is only 1.09 times (9% greater than) DG's pass rate (87%). Thus, lowering the cutoff, and thereby generally increasing pass rates and generally reducing failure rates, reduced the relative difference in pass rates.

On the other hand, at the higher cutoff, DG's failure rate was 1.85 times (85% greater than) AG's failure rate (20%). At the lower cutoff, DG's failure rate (13%) is 2.60 times (160% greater than) AG's failure rate (5%). Thus, while lowering the cutoff reduced the relative difference in pass rates, it increased the relative difference in failure rates.

One would observe the same pattern if, instead of a cutoff's being lowered, there occurred a general improvement in test performance such as to enable everyone with scores falling between the two cutoffs points to reach the higher cutoff. Raising cutoffs, or the general worsening of test performance, would tend to show opposite patterns of changes in the two relative differences, as I will address further below with regard to the revisions of the New York State standardized tests discussed at page 13 of the CIDI report.

At this juncture I note that while I discuss changes in relative differences, I illustrate the changes by means of rate ratios. The relative difference is the rate ratio minus 1 where the rate ratio is above 1 (in which case, the larger the rate ratio, the larger the relative difference) and 1 minus the rate ratio where the rate ratio is below 1 (in which case, the smaller the rate ratio, the

[5] I note that in the cases where the CIDI report discusses rate ratios with the larger rate in the numerator of the ratio it consistently describes a rate ratio of x as indicating that one rate is x "time higher" than another, rather than x "times as high" as another The former usage technically means the same thing as x plus 1 times as high (that is, treats a ratio of 2.0 as indicating that one figure is 200% higher than another rather than 100% higher). In addition to being technically incorrect, this usage in fact sometimes causes observers to regard an x percent difference to be an x plus 100% difference. See the Times Higher subpage of the Vignettes page of jpscanlan.com. See also the Risk Ratios and Rate Ratios (Relative Risk) subpage of the Boston University Medical Center Measures of Association page. While this is a minor matter compared with the main issues addressed in this letter, an organization striving to provide sound statistical analyses should give careful attention to all precision issues.

larger the relative difference). While I have sometimes used the disadvantaged group's rate as the numerator in both rate ratios, and hence with the lower rate in the numerator for the favorable outcome, I believe that presenting the larger figure in the numerator for both rate ratios makes it easier for the observer to recognize that changes in the frequency of an outcome tend to cause the two relative differences to change in opposite directions. I note, however, that for all subjects where the CIDI report measures disparities in terms of relative differences in favorable outcomes, it uses the non-white rate as the numerator of the rate ratio. Thus, in order to demonstrate that choice of numerator is irrelevant to the subject of this letter, in Table 2 I present the same information as in Table 1, except that I use the DG rate as the numerator in both rate ratios.

Table 2. Illustration of effects on relative differences in pass rates and relative differences in failure rates of lowering a cutoff from a point where 80% of AG passes to a point where 95% of AG passes (when mean scores differ by approximately half a standard deviation), with the DG rate used as the numerator in both rate ratios

Cutoff	AG Pass	DG Pass	AG Fail	DG Fail	DG/DG Pass Ratio	DG/AG Fail Ratio
High	80%	63%	20%	37%	.79	1.85
Low	95%	87%	5%	13%	.92	2.60

When the DG rate is used as the numerator in the pass ratio, lowering the cutoff increases the rate ratio from .79 to .92 (which means that the relative difference decreases from 21% to 8%). Thus, as with Table 1, one observes that lowering the cutoff caused the two relative differences to change in opposite directions, though the matter is presented somewhat differently.[6]

Because I believe that patterns of directions of changes in the two relative differences are easier to recognize when the larger outcome rate is used as the numerator in both rate ratios (and because that is the way I present it in most of the recent references), I will use that approach in the subsequent tables in this letter even though it is not the approach of the CIDI report.

[6] One may notice that while whether the larger or smaller rate is used as the numerator of the rate ratio does not affect the direction of the change in the relative difference, it does affect the size of the relative difference. One consequence of using the smaller rate in the numerator of the rate ratio is that the relative difference can never exceed 100%. Thus, it is hardly remarkable that in describing largest and smallest disparities (as in the Executive Summary), the CIDI report will always refer to subjects where the disparity is measured in terms of relative differences in the adverse outcome (and where the larger rate is used as the numerator of the rate ratio) as showing the largest disparities and subjects where the disparity is measured in terms of relative differences in the favorable outcome (and where the smaller rate is used as the numerator of the rate ratio) as the smallest disparities. The potential for choice of numerator to affect the size of a relative difference is among a number of general problems with rate ratios and relative differences as measures of association. See "Race and Mortality Revisited" (at 339) and the ASA letter (at 12-13) regarding reasons why the rate ratio and relative difference are illogical, as well as unsound, measures of association. But issues concerning which rate is used as the numerator of the rate ratio are distinct from those relating to patterns of directions of changes in relative differences regardless of which rate is used as the numerator.

Since the CIDI report discusses percentage changes in rates at which different racial/ethnic groups experienced various favorable or adverse outcomes, I also note here that a corollary to (a) 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 is (b) a pattern whereby when the frequency of outcome rate changes, the group with the lower baseline rate tends to experience the larger proportionate change in its rate for the outcome while the other group tends to experience a larger proportionate change in the opposite outcome rate. In the case of the hypothetical lowering of test cutoffs shown in Tables 1 and 2, AG (the group with the lower baseline rate for test failure) experienced a 75% reduction in its failure rate compared with a 65% reduction for DG, while DG experienced a 38% increase in its pass rate compared with a 19% increase for AG.[7] Thus, issues raised here about the CIDI report's interpretations regarding directions of changes in relative differences apply equally to the report's interpretations regarding the comparative size of the changes in the outcome rates of different racial/ethnic groups.

The above-described patterns of change in relative differences, or of comparative proportionate changes in rates of groups with lower and higher baseline rates for an outcome, will not be observed in every case. But they are an essential part of the picture and, without understanding them, it is not possible to effectively determine whether the differences in the circumstances of advantaged and disadvantaged groups reflected by their favorable or adverse outcome rates have increased or decreased over time. See discussion in "Race and Mortality Revisited" (at 329, 343) regarding its Table 2 and the pointlessness of studying, with regard to a period of general declines in poverty, either why relative differences in black and white poverty rates increased, or why relative differences in black and white rates of avoiding poverty decreased, without consideration of the pattern shown in the table.

Tables 3 though 7 use data from the early charts in the CIDI report to show how, in fact, when there occurred changes in the frequency of an outcome, the relative difference in the outcome that was generally increasing in frequency decreased while the relative difference in the corresponding opposite (decreasing) outcome increased. The tables include a column titled "EES," for "estimated effect size," containing a measure that is theoretically unaffected by the frequency of an outcome. The measure, which can be derived from any pair of favorable or adverse outcome rates, shows the difference, in terms of percentage of a standard deviation, between the means of hypothesized underlying distributions of factors associated with experiencing an outcome. The measure, along with it imperfections, is discussed and illustrated the ASA letter and the longer papers identified in the text on page 2, as well as the methods workshops listed in note 2 on that page.

For consistency with Tables 1 and 2, the tables all present the favorable outcome rates before the adverse outcome rates, and the rate ratio for the favorable outcome before the rate

[7] The matter could be viewed from the opposite perspective as well. The group with the lower baseline rate for test passage (DG) experienced the larger proportionate change in its rate for that outcome while the other group (AG) experienced the larger proportionate change in the opposite outcome.

ratios for the adverse outcome, even in situation where (as in the subject of Table 7) the CIDI report analyzes the subject in terms of the relative difference in the adverse outcome.

Table 3 is based on the portions of the charts on pages 14 and 16 of the report pertaining to the proportions of males in different racial/ethnic groups in grades 3 to 8 meeting or exceeding English and math standards. The table shows rates of meeting and failing to meet the standards for white and black males in the years 2006 and 2012, years between which there were substantial increases in rates of meeting the standards (and before the imposition of more rigorous standards in 2013). The rates shown in Table 3 and subsequent tables are estimates based on visual examination of the heights of the bars in the underlying CIDI report charts and may be inexact by a percentage point or two. Such inexactness may affect the rate ratio and EES figures slightly. But any inexactness will not affect the pattern by which the relative difference in the favorable outcome and the relative difference in the corresponding adverse outcome change in opposite directions as there occur substantial overall changes in outcome rates.

Table 3. White and black male rates of meeting standards in English and math in grades 3 through 8 (favorable outcome) and failing to meet those standards (adverse outcome) in 2006 and 2012, with ratios of white to black favorable outcome rates and black to white adverse outcome rates and EES

Subject	Year	White Fav Rt	Black Fav Rt	White Adv Rt	Black Adv Rt	Wh/Bl Fav Ratio	Bl/Wh Adv Ratio	EES
English	2006	52%	20%	48%	80%	2.60	1.67	0.89
English	2012	64%	32%	36%	68%	2.00	1.89	0.82
Math	2006	53%	18%	47%	82%	2.94	1.74	0.99
Math	2012	77%	44%	23%	56%	1.75	2.43	0.89

Consistent with the frequency-related pattern described above, the table shows that, for both English and math, substantial increases in rates of meeting the standards were accompanied by reductions in relative differences in meeting the standards but increases in relative differences in failing to meet the standards. The EES figures indicate that, to the extent that the racial differences in academic performance reflected by the rates can be measured, the difference decreased somewhat for both English and math. It should be noted, however, that inasmuch as the EES approach is aimed at inferentially identifying the differences between means of the underlying distributions, comparisons of mean scores on the proficiency tests would likely be a better way to quantify disparities.

Table 4 presents similar information for the years 2012 and 2013, the year before and during which implementation of revised New York State standardized tests resulted in substantial decreases in overall rates of meeting standards. I have elsewhere treated interpretations of effects on racial disparities of the implementation of the revised New York State tests with regard to circumstances where disparities were measured in terms of absolute

differences between rates. See the New York Proficiency Rate Disparities subpage of the Educational Disparities page of jpscanlan.com. [8]

Table 4. White and black male rates of meeting standards in English and math in grades 3 through 8 (favorable outcome) and failing to meet those standards (adverse outcome) in 2012 and 2013, with ratios of white to black favorable outcome rates and black to white adverse outcome rates and EES

Sub	Year	White Fav Rt	Black Fav Rt	White Adv Rt	Black Adv Rt	Wh/Bl Fav Ratio	Bl/Wh Adv Ratio	EES
English	2012	64%	32%	36%	68%	2.00	1.89	0.82
English	2013	42%	13%	58%	87%	3.23	1.50	0.92
Math	2012	77%	44%	23%	56%	1.75	2.43	0.89
Math	2013	49%	14%	51%	86%	3.50	1.69	1.06

With regard to relative differences in meeting and failing to meet standards, the table shows the common pattern whereby imposition of more rigorous standards increases relative differences in meeting the standards while reducing relative differences in failing to meet the standards. The EES figures increased for both English and math. Again, however, comparison of the actual mean scores would likely be more informative.

Table 5 is based on the portion of the chart on page 18 of the report pertaining to the proportions of white and black male students in grade 9 accumulating more than 10 credits. It presents the same information regarding achieving and failing to achieve that outcome as shown for the outcomes addressed in Tables 3 and 4.

[8] See also the Harvard CRP NCLB Study subpage of the Educational Disparities page regarding analyses of differences in disparities in meeting standards of different rigor without consideration of the patterns of relative differences described here. See generally the Education Disparities page itself and its Disparities by Subject, Education Trust High Achiever Study, Education Trust Glass Ceiling Study, McKinsey Achievement Gap Study, and Annie E. Casey 2014 Proficiency Disparities Study subpages regarding the near universal failure of analyses of disparities in educational outcomes to consider the effects of the frequency of the outcome on the measure employed. But see the discussion of the work of Andrew Ho and Sean Reardon at pages 3-4 of the March 8, 2016 letter to Stanford Center on Poverty and Inequality.

Table 5 White and black male rates of accumulating more than 10 credits (favorable outcome) and failing to accumulate more than 10 credits (adverse outcome) in grade 9 in the 2005-06 and 2013-14 school years, with ratios of white to black favorable outcome rates and black to white adverse outcome rates and EES

Year	White Fav Rt	Black Fav Rt	White Adv Rt	Black Adv Rt	Wh/Bl Fav Ratio	Bl/Wh Adv Ratio	EES
2005-06	58%	35%	42%	65%	1.66	1.55	0.59
2013-14	76%	55%	24%	45%	1.38	1.88	0.58

As in the case of the outcomes addressed in Table 3, the general increase in favorable outcome rates resulted in a reduction in the relative difference in the favorable outcome and an increase in the relative difference in the corresponding adverse outcome. The EES shows a negligible change in disparity.[9]

Tables 6 and 7 provide similar information on favorable and adverse outcome rates of blacks and whites, and measures of disparity, with respect to two matters related to one another – graduating in four years (from the chart on page 20) and dropping out of school (from the chart on page 22) – but where the CIDI report measured the former matter in terms of relative differences in the favorable outcome and the latter matter in terms of relative differences in the adverse outcome, with the consequence that general overall improvements showed contrasting directions of change for the two relative differences.

Table 6. White and black male rates of graduating in four years (favorable outcome) and not graduating in four years (adverse outcome) in 2005 and 2014, with ratios of white to black favorable outcome rates and black to white adverse outcome rates and EES

Year	White Fav Rt	Black Fav Rt	White Adv Rt	Black Adv Rt	Wh/Bl Fav Ratio	Bl/Wh Adv Ratio	EES
2005	58%	31%	42%	69%	1.87	1.64	0.70
2014	77%	57%	23%	43%	1.35	1.87	0.56

Table 6 shows the standard pattern whereby over a period of general increase in the favorable outcome (with corresponding decrease in the adverse outcome), the relative difference in the increasing (favorable) outcome decreased while the relative differences in the decreasing (adverse) outcome increased. The EES shows a modest decline in the disparity.

[9] Here, too, comparisons of means might be more useful than the EES for quantifying differences in the circumstances of an advantaged and a disadvantaged group.

Table 7. White and black male rates of not dropping out of school (favorable outcome) and dropping out of school (adverse outcome) in 2005 and 2014, with ratio of white to black favorable outcome rates and black to white adverse outcome rates and EES

Year	White Fav Rt	Black Fav Rt	White Adv Rt	Black Adv Rt	Wh/Bl Fav Ratio	Bl/Wh Adv Ratio	EES
2005	80%	73%	20%	27%	1.10	1.35	0.23
2014	93%	89%	7%	11%	1.04	1.57	0.25

Table 7 shows the standard pattern whereby over a period of general increase in the favorable outcome (with corresponding decrease in the adverse outcome), the relative difference in the favorable outcome decreased while the relative differences in adverse outcome increased. The EES shows a negligible change in disparity.

Similar illustrations of contrasting interpretations of directions of changes in disparities depending on whether one examines the relative difference in the favorable outcome or the relative difference in the corresponding adverse outcome may be found with regard to various other subjects of the report. In many cases, the measure on which the report relies would identify a change in direction that is consistent with that identified by the EES. But such consistency would merely be a matter of chance. And, even when the direction of change is correctly identified, the report would be misleading by suggesting that it is soundly quantifying the disparity and the change in the disparity. The last point also holds in any circumstances where the relative difference in the favorable outcome and the relative difference in the adverse outcome changed in the same direction and where one may therefore properly infer that the differences in the circumstances of the two groups in fact changed in that direction.

B. Issue Relating to the Interpretation of Data on Discipline Disparities

New York City is one of many jurisdictions that have recently relaxed discipline standards under the belief that doing so would tend to reduce relative racial/ethnic differences in suspensions and expulsions. As explained in many of the references cited in the introduction, the exact opposite is the case.

In fact, all across the country jurisdictions have been finding that recent reductions in discipline rates have been attended by increased relative racial differences in discipline rates, notwithstanding that efforts to generally reduce discipline rates are often accompanied by other actions aimed at reducing racial/ethnic differences. Jurisdictions where such patterns have been observed are discussed on the following subpages of the Discipline Disparities page of jpscanlan.com (with jurisdiction indicated in title of the subpage): California Disparities, Colorado Disparities, Connecticut Disparities, Maryland Disparities, Minnesota 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.[10]

Further, school discipline disparities are often cast in terms of the difference between the proportion a group makes up of students and the proportion the group makes up of students suspended, as, according to a February 14, 2015 *New York Post* article, was done by groups arguing for modification of New York City public school discipline standards. But another corollary to the (a) 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 is (b) a pattern whereby the rarer an outcome the greater tends to be the proportions groups most susceptible to the outcome make up of persons experiencing the outcome and persons avoiding the outcome. In the case of the hypothetical reduction in test cutoffs shown in Tables 1 and 2, assuming DG makes up 50% of test takers, lowering the cutoff would increase the proportion DG comprises of persons who pass the test from 44% to 48% and the proportion DG comprises of persons failing the test from 65% to 72%. See Table 1 (at 11) of the ASA letter.[11]

I have not examined the extent to which the data on suspension in the CIDI report are consistent with these patterns. But I note that further appraisals of discipline disparities by CIDI must be undertaken with an understanding of the issues addressed above.

I also note the following technical issue. Most analyses of group differences (including those reflecting no understanding of the issues addressed in this letter) appear to recognize that sound analyses of demographic differences in outcome rates must examine rates at which groups fall above or below certain points, not rates at which the groups fall between two points. That is, for example, in analyses of self-rated health, where categories commonly include health rated as (a) Excellent, (b) Very Good, (c) Good, (d) Fair, or (b) Poor, researchers typically analyze rates at which the groups being compared rate their health as less than Very Good or less than Good (or the opposites). Few would think to analyze differences between rates of falling into (or out of) the Very Good category, the Good category, or the Fair category, though some might analyze rates of falling into (or out of) the Excellent category or the Poor category (that is, in or out of the extreme categories). Similarly, few would think to analyze rates at which students of different demographic groups received grades of B, or grades of C, or grades of D. While one may have an interest in examining distributions across the categories, information about differing rates of falling into an intermediate category cannot even identify which is the disadvantaged

[10] These jurisdictions caught my attention as a result or news coverage of discipline disparity issues in the jurisdictions. There have no doubt also been cases where a general reduction in discipline rates was accompanied by a reduction in relative differences in discipline rates. But I have not seen news coverage of such situations. Reportage of declines in disparities during periods of general reductions in discipline rates has involved situations where disparities were measured in terms of absolute differences between rates.

[11] In addition to urging ASA to explain to the federal government that reducing the frequency of an outcome tends to increase, not decrease, relative differences in rates of experiencing it, the letter urges ASA to explain to the federal government that reducing the frequency of an outcome tends to increase, not decrease, the proportions groups make most susceptible to the outcome make up of persons experiencing it.

group (as, for example, where 30% of students from Group X and 40% of students from Group Y receive C grades).

There are a few situations, however, when analyses of demographic differences in fact examine rates at which groups fall into intermediate categories. One such case is found in the CIDI report in its analyses of rates at which different racial/ethnic groups receive one out of school suspension. One can soundly analyze rates at which groups experience (a) one or more suspension, (b) two or more suspension, or (c) any other number or more suspensions, just as one might analyze rates of receiving grades C or below, D or below, or F. But one cannot soundly analyze rates at which group experience only one suspension. I discuss this matter more fully on the Intermediate Outcomes subpage the Scanlan's Rule page of jpscanlan.com and in note 25 (at 15-16) of the ASA letter.

While this technical issue is a minor matter compared with the broader measurement issues addressed in this letter and the materials it references, the issue should be kept in mind by those attempting to monitor discipline disparities.

I hope you will carefully consider the issues raised in this letter and the materials it references, both with regard to further attention to the subjects of the *Disparities Report* and with regard to all other matters involving the City of New York's monitoring of demographic differences in outcome rates.

Sincerely,

/s/ James P. Scanlan

James P. Scanlan