August 31, 2015

Dr. Rick McDaniel, Superintendent
McKinney Independent School District
One Duvall Street
McKinney, Texas 75069

Re: Misunderstanding of Statistics in August 11, 2015 Letter from Texas Appleseed to McKinney Independent School District

Dear Superintendent McDaniel:

The purpose of this letter is to explain that an August 11, 2015 letter1 to McKinney Independent School District (MISD) from the organization Texas Appleseed regarding MISD discipline practices reflects a fundamental misunderstanding of statistics. The Texas Appleseed letter reflects the mistaken view that relaxing standards and otherwise reducing public school discipline rates will tend to reduce the proportion African Americans make up of students disciplined. In fact, the opposite is the case. Generally reducing discipline rates will tend to increase the proportion African Americans make up of students disciplined, as is being demonstrated all across the country. The Texas Appleseed letter also mistakenly interprets data to indicate that only white discipline rates have decreased.

I have addressed the relevant statistical principles in a great many places. These include a large number of letters to institutions and organizations whose interpretations of data on demographic differences 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 makes up of persons potentially experiencing an outcome and the proportion it makes up of persons actually experiencing the outcome – tend to be systematically affected by the overall frequency of an outcome.

Recent letters of particular pertinence to the subject of this letter include an April 7, 2015 letter to Texas Appleseed regarding its March 2015 report titled “Class, Not Court”.

---

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 to the Measuring Health Disparities page of jpscanlan.com.
Reconsidering Texas’ Criminalization of Truancy” and an August 24, 2015 letter to the U.S. Department of Health and Human Services (HHS) and U.S. Department of Education (DOE) regarding a December 2014 document jointly issued by the agencies titled “Policy Statement on Expulsion and Suspension Policies in Early Childhood Settings.” The letter to Texas Appleseed, which bore the subject line “The Inverse Relationship Between the Frequency of Truancy Charges and the Proportion Disadvantaged Groups Comprise of Persons Subject to Truancy Charges,” explained that, contrary to the belief expressed in the organization’s truancy report, reducing the frequency of truancy actions in Texas would tend to increase the proportion disadvantaged groups make up of persons subject to those actions. The letter to HHS and DOE explained, that contrary to the premise of the agencies’ Policy Statement, reducing preschool expulsions and suspensions would tend to increase the proportion groups most susceptible to expulsion or suspension make up of persons expelled or suspended.

While not involving school discipline, another recent letter of particular pertinence to issues addressed here is a March 9, 2015 letter to the U.S. Department Justice (DOJ) and the City of Ferguson, Missouri. The letter explained that, contrary to the premise of the DOJ’s March 4, 2015 report on the disparate impact of Ferguson’s police and court practices, reducing the frequency of adverse interactions between the police and courts and the city’s population will tend to increase the proportion African Americans make up of persons experiencing those interactions. That is, for example, increasing the number of missed court appearances necessary to trigger issuance of an arrest warrant will tend to increase, not reduce, the proportion African Americans make up of persons against whom such warrants are issued.²

Another recent item of particular pertinence to the subject of this letter is an amicus curiae brief I filed on November 17, 2014, in the case of Texas Department of Housing and Community Development, et al. v. The Inclusive Communities Project, Inc., Sup. Ct. No. 13-1371 (TDHCD brief), which involves the same statistical issues addressed here, but in the fair housing context. See especially Section I.B of the brief (at 23-27). The content of the brief is familiar to attorneys in the office of the Texas Attorney General involved with briefing and arguing the case in the U.S. Supreme Court.

***

In quite a few published articles since 1987, I have explained the statistical pattern whereby the rarer an outcome, the greater tends to be the relative (percentage) difference in experiencing it and the smaller tends to be the relative difference in avoiding it. In the school discipline context, this means that the belief so often expressed by the Departments of Justice and Education that relaxing standards and otherwise reducing the frequency of adverse school discipline outcomes will tend to reduce relative racial/ethnic differences in discipline rates is the exact opposite of reality. While reducing the frequency of adverse discipline outcomes tends to reduce relative differences in rates of avoiding those outcomes, it tends to increase relative differences in discipline rates.

I have explained this issue in school discipline context in the following recent articles: “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). These articles all illustrate the relevant statistical pattern by showing that lowering a test cutoff, and thereby generally reducing test failure, will tend to increase relative differences between the failure rates of a higher-scoring and a lower-scoring group, while reducing relative differences between the two groups’ pass rates.

Such illustration is also set out in Table 1 below, which is modified version of Table 1 of "Race and Mortality Revisited." 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. In such circumstances, at the cutoff where 80% of AG passes the test, approximately 63 percent of DG would pass the test (with corresponding failure rates of 20% for AG and 37% 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. When the cutoff is lowered to the point where the pass rate for AG is 95%, the pass rate for DG would be approximately 87% (with corresponding failure rates of 5% for AG and 13%

3 See the Bibliography subpage of the Scanlan’s Rule page of jpscanlan.com.

4 A good example of the two agency’s promotion of this belief may be found in the “Dear Colleague” letter the agencies jointly issued in January 2014, which is cited in note 2 of the August 11, 2015 letter from Texas Appleseed to MISD.

5 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.
for DG). The ratio of AG’s pass rate to DG’s pass rate would thus decrease to 1.09 (from 1.27), while the ratio of DG’s fail rate to AG’s fail rate would increase to 2.60 (from 1.85).6

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)

<table>
<thead>
<tr>
<th>Cutoff</th>
<th>AG Pass Rate</th>
<th>DG Pass Rate</th>
<th>AG Fail Rate</th>
<th>DG Fail Rate</th>
<th>AG/DG Pass Ratio</th>
<th>DG/AG Fail Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>80%</td>
<td>63%</td>
<td>20%</td>
<td>37%</td>
<td>1.27</td>
<td>1.85</td>
</tr>
<tr>
<td>Low</td>
<td>95%</td>
<td>87%</td>
<td>5%</td>
<td>13%</td>
<td>1.09</td>
<td>2.60</td>
</tr>
</tbody>
</table>

Like patterns exist in virtually every situation where two groups differ in their susceptibility to an outcome and regardless of the nature of the forces causing the susceptibilities to differ.7 In the school discipline context, the implication is that relaxing standards or otherwise generally reducing discipline rates, while tending to reduce relative differences in rates of avoiding discipline (the equivalent of passing the test), tends to increase relative differences in rates of being disciplined (the equivalent of failing the test).

Further, in point of fact, all across the country, recent reductions in discipline rate have been accompanied by increased relative racial/ethnic differences in discipline rates. Evidence to that effect is discussed on the following subpages to the Discipline Disparities page of jpscanlan.com (with jurisdiction indicated in the title of each subpage): Los Angeles SWPBS.8

---

6 One may note that the failure rate decreased by a larger proportionate amount for AG than DG (75% for AG versus 65% for DG). But 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 reducing the frequency of an outcome tends to cause a larger proportionate decrease in the outcome for the group with the lower baseline rate while causing a larger proportionate increase in the opposite outcome for the other group. Thus, with the lowering of the cutoff, DG experienced a larger proportionate increase in its pass rate than AG (38% for DG versus 19% for AG). There is no basis, however, for maintaining that the comparative size of each group’s change, either in the failure rate or in the pass rate, indicates that one group particularly benefitted from the lowering of the cutoff. See discussion in the section titled “Illogical Expectations and Unfounded Inferences” in “Race and Mortality Revisited” at 339-341.

7 Many other tabular and graphical illustrations of this pattern (and related patterns) with various types of data may be found, among other places, in the above-mentioned “Race and Mortality Revisited” and in my guest editorial titled “Can We Actually Measure Health Disparities?” in the Spring 2006 issue of the American Statistical Association publication Chance, as well as in methods workshops given in recent years at U.S. universities, including American University (2012), Harvard University (2012), University of Kansas School of Law (2013), University of Minnesota (2014), University of Maryland (2014), George Mason University (2014), and University of California, Irvine (2015). See also the Collected Illustrations subpage of the Scanlan’s Rule page of jpscanlan.com.

8 The principles discussed here apply to any type of discipline (and, indeed, any type of outcome). But since ticketing is the focus of the letter from Texas Appleseed to MISD, it may be worthwhile to mention findings in an October 2013 report concerning Los Angeles school discipline by the Community Rights Campaign of the
Denver Disparities, Florida Disparities, Maryland Disparities, 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. McKinney, Texas may be added to this list, based on information provided in the Texas Appleseed letter. But, as with the other jurisdictions, there should be no cause for puzzlement over the fact that relative differences in discipline rates increased during times of general decreases in discipline rates. Rather, the pattern is precisely what observers with a sound understanding of statistics would expect.

In addition, in November 2012 the U.S. Department of Education’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 DOE’s attribution of large relative racial 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.

***

Most of the materials just mentioned are cast in terms of relative differences. On the other hand, the letter from Texas Appleseed to MISD is cast in terms of comparisons of the proportion certain groups make up of all students and the proportions such group’s make up of persons experiencing adverse outcomes (which is also the principal way disparities were discussed in the Texas Appleseed March 2015 truancy report). The Texas Appleseed letter to MISD both attributes high proportions African Americans students made up of persons disciplined to putatively unwarrantedly high discipline rates and finds something amiss in the fact that such proportions increased during periods of general declines in discipline rates.

But the following is a corollary to the pattern whereby reducing the frequency of an outcome tends to increase relative differences in rates of experiencing the outcome while reducing relative differences in rates of avoiding the outcome: reducing the frequency of an outcome tends to increase the proportions groups most susceptible to the outcome make up of both (a) persons experiencing the outcome and (b) persons failing to experience the outcome.

Labor/Community Strategy Center titled “Black, Brown and Over-Policed in L.A. Schools. Structural Proposals to End the School-to-Prison Pipeline in the Los Angeles Unified School District and to Build a National Movement to Stop the Mass Incarceration of Black and Latino Communities.” The study, which (currently) receives only brief mention in the introductory material to the Los Angeles SWPBS subpage, showed that during a period of very dramatic reductions in ticketing for truancy in Los Angeles schools, the ratio of the black ticketing rate to the white ticketing rate increased from 3.8 to 5.8. The larger the reduction in the adverse outcome, the larger will tend to be the increase in the relative difference between rates at which advantaged and disadvantaged groups experience the outcome.
Thus, reducing the frequency of an outcome will tend to increase disproportionality regarding the outcome, though reducing disproportionality regarding the opposite outcome.\(^9\)

I illustrate this pattern in Table 2 below, which is based on Table 1 of my April 7, 2015 letter to Texas Appleseed and employs the same hypothetical test score information underlying Table 1 above.\(^10\) Table 1 of the letter to Texas Appleseed illustrated the pattern of relative differences described above and addressed in the articles on school discipline. But, based on a situation where the two groups each made up half of the test takers, the table in the letter also showed the proportion the disadvantaged (lower-scoring group) made up of persons who pass and persons who fail at each cutoff (the first pair of shaded columns in Table 2 below, with tan shading in the electronic copy of this letter). I have added to the table below two further columns showing the proportion the advantaged group makes up of persons who pass and persons who fail at each cutoff (the second pair of shaded columns, with red shading in the electronic copy). Ordinarily, I do not include such information in tables like this, since the information is so easily inferred from the proportions presented and usually is not discussed. But the information is pertinent to an interpretive error in the letter from Texas Appleseed to MISD.

**Table 2. Illustration of effects on relative differences in pass and fail rates of lowering a cutoff from a point where 80% of AG passes to a point where 95% of AG passes, with proportions DG and AG make up of persons who pass and of persons who fail (when mean scores differ by approximately half a standard deviation and DG comprises 50% of test takers)**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>80%</td>
<td>63%</td>
<td>20%</td>
<td>37%</td>
<td>1.27</td>
<td>1.85</td>
<td>44%</td>
<td>65%</td>
<td>56%</td>
<td>35%</td>
</tr>
<tr>
<td>Low</td>
<td>95%</td>
<td>87%</td>
<td>5%</td>
<td>13%</td>
<td>1.09</td>
<td>2.60</td>
<td>48%</td>
<td>72%</td>
<td>52%</td>
<td>28%</td>
</tr>
</tbody>
</table>

The second of the first two shaded columns (DG Prop of Fail) contains the information that is key to understanding the relationship between disproportionality in experiencing an outcome an outcome and the frequency of the outcome in settings where observers focus on the

\(^9\) As with the discussion of a corollary in note 6 supra, determinations of which phenomenon is designated the pattern itself, and which is designated a corollary to it, is arbitrary. In early treatments of this topic I focused on misinterpretations of patterns of disproportionality in adverse outcomes, while treating relative difference issues as a secondary matter. See *The ‘Feminization of Poverty’ is Misunderstood,* *Plain Dealer* (Nov 11, 1987); “An Issue of Numbers,” *National Law Journal* (Mar. 5, 1990); and “The Perils of Provocative Statistics,” *Public Interest* (Winter 1991).

\(^10\) Illustrations of the pattern using income data, and showing, for example, that reducing poverty will tend to increase the proportion African Americans make up of the poor, and of the non-poor, may be found in Table 1 of my Chance editorial “Can We Actually Measure Health Disparities?” mentioned in note 7. But, given that the pattern is a corollary to the previously-described pattern by which the two relative differences are affected by the frequency of an outcome, all illustrations of the pattern of relative differences could be adapted to illustrate the pattern whereby reducing the frequency of an outcome causes an increase in the proportions groups most susceptible to the outcome make up of persons experiencing the outcome and an increase in the proportion such groups make up of persons failing to experience the outcome.
proportion disadvantaged groups made up of persons experiencing adverse outcomes. The column shows that lowering the cutoff caused the proportion DG made up of persons who failed to increase from 65% to 72%. But rather than suggesting anything of consequence, such increase, like the increase in the proportion DG made up of persons experiencing the favorable outcome (from 44% to 48%, as shown in first shaded column), is simply the standard result of reducing the frequency of the adverse outcome. And certainly the pattern gives one no basis to believe that reducing the frequency of an adverse outcome will tend to reduce disproportionality in experiencing that outcome, which is the general theme of the letter from Texas Appleseed to MISD, as well as of the March 2015 Texas Appleseed truancy report and the December 2014 federal Policy Statement on preschool discipline that is the subject my August 24, 2015 letter to HHS and DOE.

In addition to emphasizing that the proportions African Americans and other minorities made up of persons disciplined increased as various types of discipline rates declined, the letter from Texas Appleseed to MISD (at 3) makes a particular point of the fact that during a general decline in ticketing, “the decrease in ticketing can be attributed almost entirely to the reduction of tickets issued to white students, for whom citations went from 28% to 15%, from 2012 to 2015.” The 28% and 15% figures are described as is they are rates at which white students were ticketed. But, based on the fact that the letter generally cites proportions groups make up of persons experiencing an outcome rather than rates at which groups experience the outcome, as well as the fact that the cited figures seem improbably high for white citation rates, I assume that these figures are the proportions of citations that went to white students.

If the proportion of citations going to white students decreased, the proportion going to other groups necessarily increased. But this does not mean that only (or almost only) whites experienced reductions in citation rates. As shown in the last column (AG Prop of Fail) and third last column (AG Prop of Fail) of Table 2, even though lowering the cutoff reduced failure rates for both DG and AG – and even though there is no basis to say that anything occurred that was other than a function of the general decrease in adverse outcomes effected by lowering the test cutoff – the proportion AG made up of persons who experienced the adverse outcome decreased while the proportion DG made up of those who experienced the adverse outcome increased. Thus, a decrease in the proportion of citations going to whites and an increase in the proportion going to other groups (i.e., African Americans and other groups with higher discipline rates than whites), as well as an increase in the proportion other groups made up of persons not ticketed and a decrease in the proportion whites make up of persons not ticketed, are simply the usual consequences of reducing the number of citations.11

11 While it is necessary to clarify what the numbers in the Texas Appleseed letter to MISD mean, as explained in note 8 supra, whites would typically have experienced a larger proportion decrease in their citation rate than African Americans (and would necessarily have experienced a larger proportion decrease than all non-white groups combined if the white proportion of persons ticketed declined). But the larger proportionate reduction in the adverse outcome rate for the group with the lower baseline rate is simply what one generally observes when the adverse outcome generally declines, along with a larger proportionate increase in the favorable outcome rate for other groups. So it would not materially affect the discussion here if the 28% and 15% figures were in fact white citation rates.
Another apparently faulty characterization of data in Texas Appleseed’s letter to MISD may be found in the transition between the second and third paragraphs on page 3. At the end of the second paragraph, the letter notes an increase in the proportion of disorderly conduct tickets that went to African American students (from 47% to 61%) and then refers to “this increase in disorderly conduct ticketing.” But the increase in the proportion African Americans made up of persons ticketed is a very different thing from an increase in ticketing, and, as discussed, typically an increase in the proportion African Americans make up of persons receiving tickets occurs in the context of decreases in the number of tickets. See note 8 supra regarding the way dramatic reductions in truancy ticketing in Los Angeles were accompanied by a substantial increase in the relative difference between black and white ticketing rates, which presumably also involved a substantial increase in the proportion African Americans made up of persons ticketed.

The above points do not necessarily mean there is nothing amiss in the ways that various groups benefited from general reductions in discipline rates. But in order to identify anything meaningful in observed patterns of changes one must have a sound understanding of statistics, including the patterns described in "Race and Mortality Revisited" and, with respect to disproportionality (as distinguished from relative differences) the discussion above, as well as the discussion in Section I.B (at 23-27) of the TDHCD brief, Section C (at 23 to 26) of my September 2013 University of Kansas School of Law faculty workshop paper “The Mismeasure of Discrimination,” and the IDEA Data Center Disproportionality Guide subpage of the Discipline Disparities page of jpscanlan.com.12

Those references explain that, using the appropriate tools, one may cautiously draw sound inferences about underlying processes on the basis of adverse (or favorable) outcome rates of advantaged and disadvantaged groups. But one can never draw sound inferences about processes on the basis of the proportion a group makes up of persons potentially experiencing an outcome and the proportion it makes up of persons actually experiencing the outcome.

***

I add two final points about discipline and discipline disparities. The first is related to the statistical issue discussed above while the second concerns a different matter.

First, the Texas Appleseed letter states (at page 3, citation omitted):

The disproportionate ticketing and arrests of these African-American students is especially stark for the offenses that are most subject to the discretion of teachers, school administrators, and school police officers. For example, in this period beginning in 2012, African-American students received approximately 46% of tickets for “disruption of

12 See also the January 2015 methods workshop at the University of California, Irvine titled 31. “The Mismeasure of Discrimination.”
class”—a vague, subjective offense that is similar to what research has shown to lead to the disproportionate punishment of students of color.

Statements based on the fact that relative differences in discipline rates are greater for subjective offenses than for objective offenses (which would translate into greater disproportionality for the former than the latter) play importantly into arguments that racial bias is responsible for a substantial part of observed racial differences in discipline rates. But arguments about the implications of the comparative size of relative differences (or disproportionality) for different types of outcomes cannot be appraised without an understanding of the effects 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.

Second, the letter from Texas Appleseed to MISD, and other general calls for relaxing discipline standards, all reflect a common perception that research generally shows that stringent discipline standards do not improve student outcomes or schooling environments. The part of that the body of research with which I am familiar, however, is of extremely dubious validity, as I discuss on the APA Zero Tolerance Study subpage of the Discipline Disparities page. See also the August 24, 2014 letter to HHS and DOE at 9.

***

While this letter is prompted by the letter from Texas Appleseed to MISD regarding the district’s discipline practices, I note that statistical issues addressed above (or in the references above) bear importantly on a range of matters where a school district attempts to appraise demographic differences in outcome rates. For example, as should be evident from Table 1, lowering a proficiency standard, or general improvements in performance on proficiency tests, will tend to reduce relative differences in proficiency rates while increasing relative differences in non-proficiency rates. On the other hand, raising standards, or general declines in performance, will tend to do the opposite – that is, to increase relative differences in proficiency rates while reducing relative differences in non-proficiency rates.

Increasingly, researchers and school administrators are appraising demographic differences in reaching and failing to reach various proficiency levels in terms of absolute (percentage point) differences between rates. Appraisal of demographic differences in outcome rates in terms of absolute differences are unaffected by whether one examines the adverse outcome or the corresponding favorable outcome. But like the two relative differences, the absolute difference also tends to be systematically affected by the frequency of an outcome, though in a more complicated way than the two relative differences. Roughly, as uncommon outcomes become more common, absolute differences between rates of advantaged and
disadvantaged groups tend to increase; when common outcomes become even more common absolute differences tend to decrease.

This pattern will have a variety of implications regarding appraisals of demographic differences regarding proficiency issues. These can include that for difficult subjects (where proficiency rates are low) general improvements in performance may increase absolute difference, while improvement in easier subjects (where proficiency rates are high) may reduce absolute differences. Similarly, and depending on the situation at particular schools or school districts, general improvements in performance will often tend to reduce absolute differences in rates of achieving basic proficiency (where rates can be fairly high) while increasing absolute differences between rates of reaching advanced proficiency levels (where rates often are quite low). Declines in performance will tend to show patterns of changes in absolute differences that are the opposite of those just described. These issues are covered in the subpages to the Educational Disparities page of jpscanlan.com and in the April 30, 2014 letter to the Education Trust and the May 13, 2014 letter to the Annie E. Casey Foundation listed in note 2.

Sincerely,

/s/ James P. Scanlan

James P. Scanlan