



The Education Trust

Closing the gaps in opportunity and achievement, pre-k through college.

August 20, 2013

Seth Galanter
U.S. Department of Education
400 Maryland Avenue, SW
Washington, DC 20202

Re: Mandatory Civil Rights Data Collection

Dear Mr. Galanter,

We strongly support the Office of Civil Rights' (OCR) efforts to collect critical information on students' opportunity to learn through the Civil Rights Data Collection (CRDC). Thanks to the CRDC, the public now has access to a set of key, critically important indicators about students' school experiences. The continuation of this important survey is absolutely essential to our country's ongoing efforts to tackle inequities in educational opportunity. In the comments that follow, we suggest several clarifications and refinements to the collection that are aimed at one thing: improving the quality and usability of these important data.

The CRDC has the potential to be an invaluable resource to advocates and civil rights groups fighting for equity by providing a wealth of information not available from other sources about students' school experiences and educational opportunities. These data can help shed light on inequities in opportunities to learn that have plagued students of color, students with disabilities, and English language learners for far too long. Unfortunately, the latest publicly available CRDC data — results of the 2009-10 collection — contain a number of data inaccuracies and inconsistencies that undermine the usability of this information. In these comments, we suggest modifications to how the data are gathered and reported that we think will greatly improve the quality and utility of the CRDC.

Over the past six months, we have worked extensively with the 2009-10 CRDC data to understand the types of analyses these data support. Our work focused on four sets of indicators:

- 1) School funding,
- 2) Course access,
- 3) School discipline, and
- 4) General school information.

We focus our comments on these four sets of indicators not because we think other data in the CRDC aren't important, but because we believe we have the most insight to offer regarding data we've worked with directly.

A note on Public vs. Restricted-Use Data

The Education Trust has obtained both the public and restricted-use versions of the CRDC databases. Because of the challenges associated with getting access to and reporting findings from the restricted-use files, however, we focus our comments on ensuring the data available in the public-use file can be meaningfully analyzed and understood as guides for

School funding

Importance of Collecting School Finance Indicators

The funding inequities among states and among districts within the same state are well documented. Less well known, but equally devastating, are the funding inequities that exist among schools in the same district. Yet these, for the most part, have remained largely invisible to parents and the public.

These within-district inequities are driven in large part by differences in teacher salaries. In many districts, the lowest paid teachers are concentrated in the highest poverty schools. But because school district accounting practices often allocate staff positions, rather than dollars, more affluent schools are able to employ a disproportionate share of the highest paid teachers without regard to budgetary implications. Consequently, the district spends more per teacher in more affluent schools than in high-poverty schools.

But salaries don't tell the whole story. You might expect that those schools with the least experienced and thus least expensive teachers are made whole by receiving additional money due to their "savings" in teacher salaries — money that could be used to help with professional development or acquisition of other resources to support student learning. Available research suggests, however, that in at least some districts and states, these schools are not made whole; indeed they often receive less in non-personnel resources, too.

Understanding how widespread these inequities are is critical. Inequitable allocations of state and local funds undermine the very intent of federal Title I dollars, which are meant to provide supplementary funds for schools serving the least advantaged students. More important, they deprive our most vulnerable students of educational opportunity and all the rewards academic success offers.

Making school-level state and local expenditure data public and transparent is the first step toward righting these wrongs. **For this reason, we strongly support s OCR's efforts to collect data on school-level expenditures from state and local funds.**

The Challenge

In the 2009-10 CRDC, there was a great deal of variability in the quality of school-level expenditure data that districts reported for their schools. Some districts, for example, reported state and local expenditures for their schools that look unreasonably low given the district-level funding data available from other sources. Consider, for example, a district in Arizona that serves over 30,000 students. In 2009-10, the school-level expenditures the district reported to CRDC indicate that it spent about \$2,500 per student in state and local funds. But according to the U.S. Census Public Elementary–Secondary Education Finance Data (F-33 survey), this district actually received nearly \$8,500 in state and local revenues per student that year.

This example is extreme, but it highlights a problem in many districts' data. Part of the nearly \$6,000 discrepancy in this district's figures is likely due to the difference in the data collected (Census collects revenue figures, while CRDC collects expenditures, and CRDC explicitly excludes certain types of spending — such as employee benefits — from its collection that Census includes). However, the difference also likely has to do with how the district tracks its expenditures — which costs it assigns to individual schools and which costs it does not. For example, some of the teachers in the district may be working in more than one school, and the district may not know how to attribute their salaries to different schools.

The problem is that it's hard to draw any conclusions about equity based on data that represent only a small portion of the resources that schools actually receive. For example, two schools might be spending the same amount per student on all of the staff explicitly assigned to that building, but one school might receive far more support from the district in the form of coaching, professional development, etc. — expenses the district may not assign to the school.

But this is by no means the only quality problem. The level of variability in the per-pupil expenditures in the CRDC also indicates widely varying interpretations of reporting instructions and policies, as well as some outright data errors.

In order to improve the quality and utility of the critically important school finance indicators in the CRDC, we recommend that OCR make the following changes to the data collection. While these focus mostly on the reporting of expenditures from state and local sources, we also want to be clear: In order to give parents, advocates, and educators as complete a picture as possible of their schools' resources, we recommend that OCR collect both total expenditures (including state, local, and federal dollars), and expenditures from state and local funds only.

Recommendations for improving the school finance data collection

We recommend that OCR:

- 1) **Focus the collection on a limited number of key indicators. These should include the following, for each school:**
 - a. ***Number of FTE teachers, and the amount spent on their actual salaries (not district averages), in total and separately, from state and local funds combined.***
 - b. ***Amount spent on salaries of instructional support (e.g., nurses, guidance counselors, instructional coaches, etc.) and administrative staff, in total and separately, from state and local funds combined.*** This amount should include actual salaries of **both** the staff who work only at the given school, and itinerant staff who split their time between multiple schools. We strongly recommend that OCR add the following clarifications regarding itinerant staff:
 - i. Clarify that costs should be attributed to individual schools based on the percentage of FTE hours that the employee spends at the school. For example, if a school nurse spends 25 percent of her time at School A, 25 percent of her total salary should be attributed to this school.
 - ii. Clarify that salaries of staff who provide district-wide services (e.g., accounting, district administration, district-wide professional development, etc.) **should not be included.**
 - c. ***Non-personnel expenditures related to instruction (in total and separately, from state and local funds combined).*** We are aware that some non-personnel expenditures are difficult to report and may not paint an accurate picture of a school's resources. As such, we recommend that OCR further clarify the types of expenditures that districts should include and exclude in this category. Specifically, we recommend that this category of expenditures be clearly defined for districts and **only include:**
 - i. Instructional materials and other supplies, such as textbooks, curricular materials, classroom, office, and laboratory supplies, etc.
 - ii. Computers, software and other technology, including SMART Boards, projectors, etc.
 - iii. Library books and other media center materials, such as subscriptions to research databases or journals
 - iv. Contracted services related to instruction, such as comprehensive school reform programs like Success for All
 - v. Additional funds, not captured elsewhere in this reporting, allocated to schools for the express purpose of professional development that is developed and/or purchased at the school level. This can include stipends for staff to work additional hours, contracted professional development trainings, etc.
- 2) **Clarify that districts should report both total expenditures (including federal dollars) and expenditures from state and local funds only.** Currently, the fact that districts are asked to report expenditures from state and local funds appears only in the definitions portion of the data collection form. We strongly recommend including clear instructions next to the tables where districts are entering the data.
- 3) **Build automated data checking mechanisms into the data collection form.** We recommend that the CRDC data entry form automatically calculate the total per-pupil spending for each school based on the school finance data entered. Seeing the per-pupil amount may alert the individual entering that data to potential errors.

Course Access Data

Research shows that the biggest predictor of a student's success in college is the rigor of the high school curriculum.¹ Research also shows that low-income students and students of color are less likely to complete a rigorous course of study in high school.² To us, and to other equity-minded educators and advocates, these findings indicate that students of color and low-income students often do not get equal access to classes that will help prepare them for postsecondary success. When we cite low numbers on course enrollments, however, the response we often get is, "Well, clearly these students are not interested in taking these challenging courses." The power of the course access data in the CRDC is that in addition to course enrollments, it provides information on *course offerings*. Moreover, the CRDC contains information not only on whether a course is offered at all, but how many classes in that course a school has. Many schools today may claim to offer calculus, for example. Not surprisingly, however, wealthier and predominantly white schools often offer far more calculus classes — and thus greater access to them — than schools serving predominantly black and Latino students. CRDC data have the potential to help advocates make the critical case that too many students attend schools where enrolling in a rigorous course of study isn't even an option.

The Challenge

The course access data in the CRDC contain a variety of inconsistencies that make us question the accuracy of these data for some schools. For example, some schools report that they offer a particular class, but then report that no students are enrolled in that course. Other schools report student enrollment in a course, but no classes offered. Others still may be reporting enrollment for a course under a different variable. For example, the CRDC asks schools to report the number of classes offered and students enrolled in calculus, and, separately, the number of students enrolled in AP math courses. It's possible that some schools reported their calculus students under the AP math variable, leaving calculus enrollment at zero. Many of these errors and inconsistencies can be prevented by making some simple changes to the CRDC data entry form. These changes would increase confidence in the data and minimize the number of schools that have to be eliminated from an analysis. We summarize our key recommended changes below.

¹ C. Adelman, "The Toolbox Revisited: Paths to Degree Completion From High School Through College," U.S. Department of Education (2006), <http://www2.ed.gov/rschstat/research/pubs/toolboxrevisit/index.html>.

² C. Nord, S. Roey, R. Perkins, M. Lyons, N. Lemanski, J. Brown, and J. Schuknecht. "The Nation's Report Card: America's High School Graduates," (NCES 2011-462), U.S. Department of Education, National Center for Education Statistics (2011).

Recommendations for improving course access data collection and reporting

Recommended changes to the data collection form

- 1) **Collect data on the number of classes offered and the number of students enrolled in the same table.** Currently, the CRDC data entry form collects data on the number of classes offered and the number of students enrolled in separate tables, with slightly different data entry instructions. This increases potential for data entry error (for example, district staff could enter enrollment for calculus in the geometry table). We recommend that the CRDC create data entry tables for each course (Algebra I, Geometry, Algebra II, Advanced Math, Calculus, Biology, Chemistry and Physics, respectively) requesting the number of classes offered and students enrolled at the same time.
- 2) **Build automated data checking mechanisms into the data collection form.** We recommend that OCR build additional data checking functions into the data collection form, including, at minimum, the following:
 - a. If a district reports a number of classes offered, but no enrollment in those courses, the district should have to verify (via a checkbox) that the school is on a block schedule and the course will be offered in a different semester. Otherwise, the form should not allow individuals to enter a number of classes without also specifying enrollment in those classes.
 - b. If a district reports enrollment in a course for a school, but no classes offered in that course, the district should have to verify (via a checkbox) that the class is offered off-site.
 - c. The form should also automatically calculate an average class size based on the data reported, and flag any unusually low (e.g., fewer than five students) or high (e.g., more than 40 students) class sizes for the individual entering the data to signal potential data errors.

Recommended changes to the public-use data set

- 3) **Differentiate between true zeroes and suppressed values in the public use dataset.** In the public dataset, all student counts are rounded to the nearest five, or, for very small numbers (such as one or two), to zero. As a result, it is impossible to differentiate between small positive values and true zero values. While we understand the necessity to protect student privacy, we recommend that OCR differentiate zero values and positive values in the public dataset. For example, OCR could code any positive numbers that are currently getting rounded to zero as “less than five.”

School Discipline Data

In recent years, educators, parents, and civil rights advocates have become increasingly concerned about excessive use of out-of-school suspensions and expulsions in our public schools, and the disproportionate use of these practices with students of color. Data from the National Center for Education Statistics show that nearly 50 percent of African American high school students have been suspended at least once during their academic career, compared with less than 20 percent of white high school students.³ And while there aren't many studies on the impact of suspensions and expulsions on student outcomes, the data that are available indicate significant detrimental effects. For example, Fabelo, et al. (2011) found that Texas students who were suspended or expelled were far more likely than other students to repeat a grade. Moreover, students who were suspended or expelled for minor (discretionary) offenses

³ “Parent and Family Involvement in Education Survey of the National Household Education Surveys Program,” (PFI-NHES), U.S. Department of Education, National Center for Education Statistics (1999, 2003, and 2007), http://nces.ed.gov/pubs2012/2012026/tables/table_14.asp.

such as school code violations were almost three times more likely to become involved with the juvenile justice system.⁴

The data on school discipline practices in the CRDC, including in-school and out-of-school suspensions, and expulsions with and without educational services, could allow advocates to examine inequities in discipline practices in their districts and in their schools. We strongly support OCR's efforts to collect and publish these data, and recommend a few changes to the data collection and reporting to increase the usability of these indicators.

The Challenge

Throughout the CRDC dataset, all student counts are disaggregated by both race and gender. To obtain a total count of students enrolled in calculus, for example, one has to add the total number of male and female students enrolled in the class.

In addition to race and gender, discipline data are further disaggregated by disability status. In order to find out the total number of students expelled *without* education services, for example, one has to add:

- The number of female students with disabilities expelled without education services,
- The number of female students without disabilities expelled without education services,
- The number of male students with disabilities expelled without education services, and
- The number of male students without disabilities expelled without education services.

To get the total number of expulsions, one would need to add in four additional categories of students (male and female with and without disabilities) expelled *with* education services — a total of eight groups of students.

In each of these eight groups, the reported numbers of affected students are rounded to the nearest five to protect student privacy. Very small counts (counts of one or two students) appear to be rounded to zero. Added across eight different categories, this rounding could lead to substantial under or over-estimates of disciplinary actions, particularly when the actual numbers are small. For example, consider a school that expelled 2 students in each of the eight groups (male and female students with and without disabilities, expelled with and without services, respectively). In total, this school expelled 16 students. The public dataset, however, could report zero expulsions since each of the eight reporting categories gets rounded to zero.

The effects of rounding are largest when the underlying numbers are small. For example, the practical difference between rounding three expulsions up to five is bigger (a 67 percent difference) than that of rounding 23 expulsions up to 25 (a 9 percent difference). By adding up the total rounded counts of female and male students enrolled in a class, we are probably not introducing a whole lot of error in most schools. Adding up the numbers of students expelled or

⁴ T. Fabelo, M.D. Thompson, M. Plotkin, D. Carmichael, M.P. Marchbanks, and E.A. Booth. "Breaking Schools' Rules: A Statewide Study of How School Discipline Relates to Students' Success and Juvenile Justice Involvement," Council of State Governments Justice Center (2011).

suspended across eight reporting categories, especially for students of a particular race/ethnicity, is far more problematic.

While OCR could certainly build additional data checks into the data collection form (see below for specific recommendations), the changes that would make the most difference in improving the quality of the suspension and expulsion data actually need to happen on the reporting side. While we commend OCR for the level of disaggregation (by race, sex, and disability status) that the data provide, the need to protect student privacy makes it imperative that OCR also publish suspension and expulsion data at a more aggregated level. More specifically, we recommend the following changes to the data collection and reporting:

Recommendations to OCR for improving discipline data collection and reporting

Recommended changes to the public-use data set

- 1) Publish total numbers of suspensions/expulsions for students overall and by race:** The level of disaggregation in the CRDC could, hypothetically, allow for very detailed analyses of discipline practices. However, the combination of this level of disaggregation (by gender, race, and disability status) with the rounding rules make it nearly impossible to answer any questions about discipline practices with confidence. We recommend that in addition to reporting disaggregated numbers, CRDC also report some commonly requested totals. These should include, but need not be limited to:
 - a. The total number of students expelled, by race
 - b. The total number of students suspended out of school at least once, by race
 - c. The total number of students expelled, by race and gender
 - d. The total number of students suspended out of school at least once, by race and gender
 - e. The total number of students expelled, by race and disability status
 - f. The total number of students suspended out of school at least once, by race and disability status.
- 2) Differentiate between true zeroes and suppressed values in the public-use data set.** To enable data users to differentiate between schools that suspended or expelled no students in a particular group, and those that expelled or suspended a small number of students, we recommend that OCR code any positive numbers that are currently getting rounded to zero as “less than five.”

Recommended changes to the data collection form

- 3) Build additional data checks into the data collection system:** Similar to our recommendations for collecting course access data, we recommend that OCR modify the data collection form to provide some data checks that automatically flag questionable numbers. For example, when district staff enter suspension or expulsion information, the form should:
 - a. Calculate the percent of students suspended or expelled and display that number for the individual entering the data. Seeing the actual suspension or expulsion rates may help district staff recognize cases where they are reporting disciplinary actions instead of the students affected by those actions.
 - b. Flag unusually high or low values and require the individual entering the data to certify that the numbers are accurate in order to proceed with the submission.

General School Information

The CRDC collects a substantial amount of general information about schools from districts. These indicators include school type (i.e., whether the school serves primarily special education students, or is an alternative school), grades offered, and student demographics. This information is critical in providing context for any analysis of inequities in educational opportunity.

Many of these indicators, however, are also collected via other Department of Education surveys. For example, the National Center for Education Statistics collects data on school type, grade level, and student enrollment by race/ethnicity via its Common Core of Data Public School Universe survey. Rather than asking districts to enter these data “from scratch,” we recommend that OCR pre-populate any data fields for which information is available from the CCD or EdFacts, and ask districts to verify that this data is correct. Doing so would reduce reporting burden on districts and improve the quality of the data.

Recommendations for improving general school information data collection and reporting

To reduce reporting burden, and improve the quality of the data in the CRDC, we strongly recommend that OCR identify general school indicators for which data are already being collected via other Department of Education collections, and pre-populate this information for each school. Districts should then be asked to either certify that the pre-populated information is accurate, or edit it as necessary.

We believe that with the changes above, the 2013-14 Civil Rights Data Collection can go a long way toward empowering educators, policymakers, and advocates with the information they need to expose — and then fix — inequities in the opportunities to learn that students are given. Please let us know if we can provide additional information on our experiences with the data set or on these recommendations for improving this critically important public accountability tool.

Cordially,



Kati Haycock
The Education Trust