

Using Student Data to Address Inequities

Claus von Zastrow, Maxine T. Roberts and John Squires

State education data systems help policymakers use data to evaluate the impact of their efforts to improve education. By disaggregating the data — that is, breaking it out by different student subgroups — policymakers can ensure that their efforts address the needs of students who have been traditionally underserved in educational settings. Fortunately, states' data systems have become much more sophisticated in recent decades, making it easier to collect and report data on a growing number of student subgroups. Recent decades have also witnessed a growing commitment to using data as a tool for addressing inequities in education.

Yet barriers to collecting and reporting on disaggregated data persist: It can be expensive, and efforts to report on small groups of students can violate state and federal privacy laws. Nonetheless, states that disaggregate data as fully as possible can best support the increasingly diverse student populations they serve in their schools and higher education institutions.

A Note on Terminology: This brief uses terms established by the federal government to describe racial and ethnic categories. Researchers have raised questions about some of these terms — like Hispanic or Latino — that may not accurately reflect the ethnic or gender composition of the groups they aim to describe.

The brief adheres to the federal language when it traces the evolution of that language over the past half century. In doing so, however, it does not endorse that language.



Part 2 of 3

The brief is the second installment in a three-part series on how the use of an equity framework can support success for students who are underserved in educational settings.

The series can be helpful for policymakers who are committed to identifying the barriers that impede success for students who are underserved in their state and working with their constituents to advance improved educational outcomes for these learners.

Past Data Practices: From Repression to Opportunity

Since the United States government began collecting data on people living within its borders, it has been disaggregating those data. Over the past two centuries, disaggregation has evolved from a mechanism for limiting opportunity to a means of expanding it.

For Years, Disaggregation Supported Discrimination. The history of data collection in the United States offers a stark reminder that disaggregation may do little good — and may even do harm — without a strong framework for addressing inequities. The 1790 census counted each enslaved person as [three-fifths of a free person](#), a strategy that strengthened the political power of slave-holding states without benefiting their enslaved populations.

Subsequent census collections in the 19th and early 20th centuries added new racial, ethnic and national groups, often to help leaders track and restrict immigration. For example, [finer distinctions](#) among East Asian populations in the 1890 census helped [government officials administer](#) the [Chinese Exclusion Act of 1882](#). Similarly, a new race category in the 1930 census — Mexican — supported [aspirations](#) to restrict immigration from Mexico. Just over a decade later, government leaders infamously used census data to [identify Japanese Americans](#) for internment during World War II.

The Goals of Disaggregation Shifted in the 20th Century. The second half of the 20th century saw a marked shift away from discriminatory data policies. In the 1970s, the federal government began to require federal agencies to disaggregate data to support the goals of legislation like the Civil Rights Act of 1964 and the Equal Employment Opportunity Act of 1972. In 1977, the Office of Management and Budget (OMB) issued [Directive No. 15](#), which presented new requirements for how federal agencies should collect and report on racial and ethnic information.

The directive, which went into effect on Jan. 1, 1980, required agencies to collect and report on data on four racial categories (American Indian/Alaska Native, Asian/Pacific Islander, Black and white) and two ethnic categories (Hispanic origin, not of Hispanic origin). Each person would be classified under one ethnic category and at least one race — for example, Hispanic and American Indian/Alaska Native. The directive specified that agencies should use those categories to show compliance with civil rights and equal employment opportunity law.

More Recently, Federal Agencies Have Expanded Racial and Ethnic Categories. Early in this century, the OMB expanded and refined its list of racial and ethnic categories, and the U.S. Department of Education embraced those categories. In 2000, [two revisions](#) to Directive No. 15 went into effect:

- The Asian/Pacific Islander category split into two separate categories — Asian and Native Hawaiian/Pacific Islander.
- The term Hispanic became Hispanic or Latino.

In 2007, the U.S. Department of Education issued [guidance](#) requiring education institutions and other recipients of agency funds to adopt those racial/ethnic categories. It also required them to provide students the option of selecting more than one race. Recipients of federal funds, including state education agencies, were required to implement the new guidelines by fall 2010. The No Child Left Behind Act of 2001 had already required states to report education data disaggregated by race, ethnicity and gender.

The federal [Every Student Succeeds Act \(ESSA\) of 2015](#) further expanded the list of required subgroups to include such categories as students experiencing homelessness, students in foster care, students with disabilities, English language learners and students with a military parent. [Thirty-two states](#) reported on all required subgroups in their 2020-21 school report cards.

These changes to federal policy over the past four decades have helped make data a tool for exposing barriers to opportunity that disproportionately affect some groups of students.

Current Data Practices: Continued Challenges

For all their value, federal requirements set a floor, not a ceiling, for disaggregating data. Subgroups of students still disappear into larger groups, leaving the barriers those subgroups face undetected and unaddressed. States, local governments and institutions can avoid this problem by exceeding federal mandates for disaggregating data.

Some Groups of Students Remain Invisible. Current practices for collecting data can obscure as much information about students as they reveal. Some categories defined in the OMB directives can become catchalls for very diverse groups of people. The Asian classification describes a group of people with the [most diverse](#) economic and educational outcomes of any racial or ethnic group in the United States. For example, [a report](#) by the National Forum on Education Statistics notes that 67% of Asian Americans ages 18 to 24 are enrolled in college. After disaggregating those data, however, the report reveals that enrollment rates among South Asian students alone range from 70% (Pakistani) to 20% (Bhutanese). Similarly, the American Indian/Alaska Native classification conflates [hundreds](#) of tribes whose children and youth have distinct histories, cultures and needs. Few state education data systems capture these distinctions.

American Indian/Alaska Native students are also [much more likely](#) than students of other races to disappear into the Hispanic or Latino or two-or-more-races categories. The U.S. Department of Education's 2007 guidance counsels federal and state agencies to classify students who identify as Hispanic or Latino in that category alone, regardless of what race they select. The guidance also states that students who select more than one race and who do not identify as Hispanic or Latino should appear only in the two-or-more-races category. A [2017 study in Oregon](#) found that only 13% of students who identified themselves as American Indian did not also identify as Hispanic/Latino or multiracial. State reports would not typically include the remaining 87% among American Indian students. Such incomplete information can hobble plans to understand and meet American Indian students' unique needs.

Technical, Budgetary and Legal Challenges Can Hamper Disaggregation.

Surveys that rely on representative samples of students limit possibilities for fine disaggregation because sample sizes of some student populations can be too small to generate reliable results. The National Council of American Indians has described American Indians and Alaska Natives as “the [Asterisk Nation](#),” because research reports often represent them with an asterisk, indicating a lack of reliable data.

Even if data systems aim to capture all students in an institution or school system, disaggregating data by finer subcategories, like Black females or Latino students experiencing homelessness, can increase the complexity — and the cost — of analyzing and visualizing data.

In addition, federal and state privacy laws can prohibit public reporting on small subgroups of students, even if the data about those students are available and reliable. As the number of students in any subgroup grows very small, individual students in the group become easier to identify, putting their privacy at risk. Federal and state laws specify the numbers, or the [minimum n-size](#), below which any public reporting system must suppress student data. Reports that focus on large districts, or the state as a whole, are less likely to encounter those minimums.

Public Reporting Often Presents Incomplete Information. Even in cases where states collect enough data to support more detailed disaggregation, state agencies or institutions might not fully disaggregate the data they report. The Data Quality Campaign's [review of state report cards](#) has found that some states fall short of [ESSA requirements](#) for disaggregating data in their K-12 school report cards. Higher education institutions often do [not fully disaggregate](#) data in their reporting, even to accreditation bodies.

An informal scan of recent system reports in higher education shows that some authors present outcomes in the aggregate, even when disaggregated data are available. The tendency to disaggregate data only partially — that is, to omit some student groups or to compare outcomes for specific student groups with the

outcomes for all students — can deprive education leaders and affected students of the information they need to assess the effectiveness of reform efforts.

This table has been adapted from an actual report from a higher education institution and displays the failings of partially disaggregated data.

Retention and Graduation by Institution

Institution	All Students		Pell-Eligible Recipients		Black	
	Retention	Graduation	Retention	Graduation	Retention	Graduation
College 1	71%	42%	65%	38%	70%	40%
College 2	65%	34%	59%	31%	51%	27%
College 3	59%	28%	49%	18%	45%	14%

First, the report falls far short of OMB guidance on disaggregating data by racial and ethnic group. Instead, it provides data on Black students alone. Second, as presented, the data undermine the ability to make full comparisons: How, for example, did students who qualified for Pell Grants perform relative to those who didn't? State leaders and other stakeholders can only compare Pell-Eligible Recipients' outcomes with the outcomes of All Students.

The second problem can lead policymakers and others to underestimate outcome differences among groups. In College 3, the difference in graduation outcomes between All Students and students who qualified for Pell Grants is roughly 10 percentage points — but the All Students category includes students who qualified for Pell. If Pell enrollment rates in College 3 mirror the national average of 34%, then the graduation rate among students who did not qualify for Pell Grants is 33% — 15 percentage points higher than the rate for those who did. This calculation is, of course, entirely speculative. The table offers too little information to illuminate outcome disparities among subgroups.

In this example, it is unclear whether the education system is in fact serving each student group or how serious the disparities among groups might be.

A Rising Tide Might Not Lift All Boats Equally

In 1963, John F. Kennedy popularized the saying *a rising tide lifts all boats* in a speech arguing that all citizens benefit from a better economy. In education policy, the aphorism can mean that a reform meant to benefit all students will benefit each student subgroup as well. Reformers can test that proposition by disaggregating data. The reform may improve student outcomes in the aggregate, yet data on different student groups can reveal persistent or even widening disparities, suggesting that the reform has not succeeded.

In this example, a reform has lifted all boats, so to speak, but differences in outcomes have widened.

College Completion Rates

Group	College Completion Before Intervention	College Completion After Intervention
Pell-Eligible	42%	48%
Pell-Ineligible	52%	63%
Difference in Outcomes	10%	15%

Given the growing importance of completing college in a competitive job market, the reform in this scenario may worsen economic disparities between the groups. A truly successful reform might not close outcome differences entirely, but it should benefit every student group while narrowing these differences considerably. Leaders who design a reform will have to determine what sort of disparities are reasonable in the short or medium term (e.g., < 5%). Fully disaggregating outcomes data can determine whether the reform has met that standard for success.

Strategies for Improving Data Collection and Presentation

Some challenges of collecting and reporting on disaggregated data are unavoidable. For example, states and districts cannot, and should not, circumvent minimum n-sizes, which protect students' privacy. Still, states can adopt policies

or practices that go well beyond federal requirements for disaggregating data. By doing so, they equip decision-makers with the information they need to address educational inequities, as detailed in the framework presented in the first report in this series.

Expanding the Number of Student Subgroups

Some states have enacted laws or established policies to collect and report data on more student subgroups.

California [Government Code Section 8310.5](#) requires state agencies, boards or commissions to collect data on each major Asian and Pacific Islander group in the state: Asian Indian, Cambodian, Chinese, Filipino, Guamanian, Hawaiian, Hmong, Japanese, Korean, Laotian, Samoan, Tahitian and Vietnamese Californians.

Florida's annual [Fact Book](#), which reports on the state's college system, publishes data on subgroups of disabled students, including students with visual, hearing, physical and speech disabilities, students with brain injuries and autistic students.

Minnesota's 2016 All Kids Count Act (sections [120B.31, subdivision 3a](#) and [120B.35, subdivision 3](#) of the [Minnesota Statutes](#)) requires school districts to disaggregate information about enrollments, K-12 graduation rates and assessment results by more detailed racial and ethnic categories. For example, the state [now publishes](#) state-level data on outcomes for Asian American students in the aggregate as well as for 12 Asian subgroups — including Hmong, Filipino and Korean students, who together comprise more than half of Asian American enrollments in the state. In the same year, **Washington's** governor signed a similar bill [into law](#).

Rhode Island has enacted [similar legislation](#) focused on Asian Americans and Pacific Islanders.

Counting All American Indian and Alaska Native Students

Agencies in some states have taken steps to recognize American Indian or Alaska Native tribal identities and to ensure that American Indian/Alaska Native students do not vanish into other reporting categories.

In 2014 and 2017, the **Oregon** Department of Education invited the nine federally recognized tribes in the state to compare their data with state data and to add detail on students' tribal affiliations. While [the public reports](#) do not include information on individual tribes, each tribe received briefings and reports on the performance of its students. The 2014 report informed the state's [American Indian/Alaska Native education plan](#) and improved the state's efforts to engage Native

education leaders in state decision-making. The reports also published total enrollments of American Indian/Alaska Native students in the state — including those who identified as Hispanic or multiracial.

The **Washington state statute** that created the Office of Native Education requires that it annually report to the governor, legislator and the Governor’s Office of Indian Affairs on the state of Native education. The **resulting report** on Native students’ graduation and dropout rates explicitly includes data on Native students who also identify as Hispanic or multiracial. The report even notes that the Office of Native Education is working with “Student Information staff at [the Washington State Office of Superintendent of Public Instruction] to explore ways to establish a system to retrieve accurate data, including graduation rates, dropout rates, academic success, and accurate Native student enrollment counts for all American Indian/Alaska Native students in Washington state’s schools.”

Reporting on the Intersection of Student Groups

States are also disaggregating their education data more finely, reporting on intersections between race and gender, for example, or disability and race.

The **District of Columbia’s** Office of the State Superintendent of Public Instruction publishes **annual assessment results** disaggregated by both race and gender. The results reveal gender gaps within racial groups that can rival the gaps between racial groups.

When the **Illinois** Board of Higher Education reports on **college enrollment** and **credentials**, it breaks out the data by both race and gender. This level of detail allows decision-makers to explore how race and gender interact to advance or limit opportunity. In 2020, for example, the board released an **analysis** of the state’s enrollment patterns revealing dramatic gender differences within specific racial categories. It also found that women of color typically earn less than men in their first jobs, despite enrolling in college at higher rates. The analysis quotes board Executive Director Ginger Ostro’s assessment of why the data matter: “This new data, disaggregated by gender, surprised us and gives us more information as we focus on how to pursue equity.”

In 2021, the **Maryland** State Board of Education’s **Task Force on Achieving Academic Equity and Excellence for Black Boys** recommended that the state education agency show gender breakdowns for each racial group in the state. The task force’s **final report** examines gender disparities in Black students’ school performance, graduation rates and suspension rates. These data inform the report’s recommendations for strategies to support Black male students, including structured mentoring programs, teacher professional development and tailored academic support. The task force encourages the Maryland State Department of Education to publish a gender

breakdown for each racial group in the state's [school report card](#) or through other statewide reporting mechanisms.

Final Thoughts

The past four decades have witnessed [major improvements](#) in data collection and reporting. As technologies for collecting and visualizing data have improved, so have federal and state agencies' commitments to using data to rectify inequities. Yet it remains all too easy for groups of students to become buried in aggregate data, where they may remain out of sight and out of mind.

Decision-makers who examine disaggregated data can discover barriers to opportunity that stifle students' long-term prospects. One important step in designing any reform is to divide student groups into the smallest categories that are both meaningful and feasible to explore. Only then can reformers properly evaluate whether their reforms are meeting the needs of students whose challenges would otherwise remain invisible.

Fortunately, states are going well beyond federal guidelines to make such data available to policymakers, community leaders, educators and families.



This brief is the second installment in a three-part series. Don't miss [Part 1](#) and [Part 3](#).

About the Authors



Claus von Zastrow



Claus von Zastrow oversees efforts to improve statewide longitudinal data systems and provide state-by-state data on STEM education. He has held senior positions in education policy and research for 20 years and has spent much of that time helping diverse stakeholders find consensus on important education issues. Claus is dedicated to ensuring that state leaders have the information and guidance they need to make the best possible decisions affecting young people. Contact Claus at cvonzastrow@ecs.org.

Maxine T. Roberts



As director of Strong Start to Finish, Maxine works with state leaders, researchers and technical assistance partners to address inequities in developmental education for students who are racially minoritized, from lower-income communities and returning adults. Prior to joining Education Commission of the States, Maxine directed after-school and college-preparatory programs, worked with community college faculty to improve their course outcomes in developmental education, and studied the factors that foster success for students of color in community college developmental math courses. Contact Maxine at mroberts@strongstart.org.

John Squires



John Squires is the executive director of advancement and planning at Cleveland State Community College in Cleveland, Tennessee. Previously, he served as the director of college and career readiness at the Southern Regional Education Board, working with states to increase student preparedness in literacy and mathematics. John has coauthored papers for New Directions in Institutional Research and IES What Works Clearinghouse, and his work has been recognized by the American Association of Community Colleges.