

Understanding Minimum N-Size and Student Data Privacy: A Guide for Advocates

When used effectively, data empowers educators, students, and families to make decisions that help all learners succeed. Access to high-quality education data is essential to improving students' achievement in school and preparing them for success in life. To that end, states are required to:

- **publicly report aggregate information on student academic performance** each year and provide that information broken down by subgroup (e.g., gender, economically disadvantaged status, major racial and ethnic groups)
- **protect students' privacy and comply with data sharing guidelines** established under the Family Educational Rights and Privacy Act (FERPA) and state privacy laws.

WHAT IS MINIMUM N-SIZE?

The minimum number (n) of students necessary to create a student subgroup without jeopardizing privacy.

These responsibilities can seem at odds with one another, as fears of disclosure increase as more types of information are shared. This tension becomes especially evident in conversations about minimum n-size, the minimum number of students necessary to create a student subgroup without jeopardizing student privacy.

As states design and implement their new accountability systems and re-think how they publicly report aggregate information about student academic performance, n-size becomes a central topic of discussion and debate. Some groups will advocate for a smaller n-size, wanting to ensure that all students are represented in public reporting and school accountability, and other groups will hesitate due to student privacy concerns.

Why Consider N-Size?

Minimum n-size is the lowest number of students needed to create a student subgroup without inadvertently revealing personally identifiable information about any given student. Because No Child Left Behind required that states establish a minimum n-size, it is a focal point of many privacy conversations. Too often the question comes down to, "What is the best minimum n-size?" when really the question should be, "What do we need from our data?"

The Every Student Succeeds Act (ESSA) opens the door for states to revisit their minimum n-size. With the elimination of the Department of Education's (ED) final regulations on school accountability, advocates will play an increasingly important role in making a case for n-sizes that ensure as many students as possible are counted, while student privacy is protected.

WHAT DOES ESSA REQUIRE?

States **MUST**:

- Determine a minimum n-size, which will be used for all students and subgroups
- Explain how they chose that n-size, including how they included teachers, principals, parents, and other stakeholders in the decision-making
- Explain how they know the n-size will sufficiently protect students' personally identifiable information

States **CAN**:

- Maximize transparency and set a lower minimum n-size for public reporting than the n-size for accountability.

Challenges with N-Size

To help states set and use their minimum n-size effectively, it's important to understand its limitations, including:

- **Incorrect use of minimum n-size can actually compromise privacy.** If information on a particular student subgroup is not reported (often referred to as suppressed), it is possible to calculate the values for that missing information by subtracting the reported values from the total. To address this, states sometimes suppress the next largest subgroup to make the reverse calculations impossible—this is known as complimentary suppression. This however, does not always happen, leaving states at risk of accidental disclosure.

HOW CAN N-SIZE ACTUALLY COMPROMISE PRIVACY?

In this basic example, the data for Hispanic students was suppressed because there were fewer than five students, the minimum n-size, in the subgroup. Because the data about Hispanic students is the only suppressed data, you could work backwards to determine that three Hispanic students were chronically absent. Depending on student body demographics, it is possible that those three students could be personally identified.

Number of Chronically Absent Students by Student Subgroup

Minimum n-size = 5

Black	White	Hispanic	Asian	Total # of students
10	15	-	22	50

- **When n-sizes are too large, student groups disappear and transparency suffers.** [A 2016 report by Policy Analysis for California Education](#) showed how easily an inflated n-size can make a school's lowest performing racial or ethnic group disappear, substantially changing school performance outcome and blurring the picture of a school's performance. Inaccurate information about student performance at the school or state level hurts transparency and breaks down public trust in state information and resources.
- **Inaccurate data leads to ineffective policies.** If states do not have an n-size that both provides transparency and protects student privacy, they will not have an accurate picture of how all students are performing and cannot make informed and strategic choices. Publicly reported data is fundamental to how families make decisions about their child's education and how policymakers govern, thus it is essential that data quality be put on equal footing with privacy.

Useful Data Terms to Know¹

When advocates ask for data from the state, they are often told they cannot have the information due to disclosure fears and minimum n-size restrictions. But, there are other strategies, detailed below, that states can take to protect student data privacy without unnecessarily restricting access.

Method	Description	Benefits	Challenges
Suppression	Removing data to prevent the identification of individuals in a small group. Generally also requires suppression of non-sensitive data to ensure that the values of the suppressed cells may not be calculated by subtracting the reported values from the row and column totals.	Increased precision	Less coverage of subgroups

¹ Information taken directly from PTAC resources "[Frequently Asked Questions—Disclosure Avoidance](#)" and "[Data De-identification: An Overview of Basic Terms](#)"

Method	Description	Benefits	Challenges
Perturbation	Introduces uncertainty (or noise) into the data set so that the user isn't sure whether the value actually corresponds to its given category. This can be done in a number of ways, including switching data points between different cells.	Minimizes loss of data utility as compared to other methods. Low risk of disclosure.	Reduces the transparency and credibility of the data. Generally considered inappropriate for public reporting of program data, especially from an accountability perspective.
Blurring	Reduces the precision (exactness) of the disclosed data to minimize the potential of identification. Examples of blurring include rounding, aggregating across different populations (e.g., report information on "gap" populations which includes subgroups that aren't otherwise reported) or geographies, and reporting percentages and ranges instead of exact counts (e.g., reporting that Hispanic students were absent 10-15 days this year as opposed to giving exact number).	Excellent coverage of subgroups Generally low risk of disclosure	Reduces the precision of the data Can affect the utility of the data by reducing one's ability to make inferences about small shifts in the data, since the precision of the data has been decreased.

Tips for Advocates

When engaging with states, advocates should assume states' good intent, be prepared to ask informed questions, and have a broader understanding of the issues wrapped up in student data privacy. The tension between privacy and transparency extends well beyond n-size, so being equipped to talk about the broader issues will help advocates when they hit a privacy barrier, including high n-sizes.

■ Know What FERPA Says

FERPA is not a blanket "no" to data sharing or reporting. FERPA provides appropriate limits for sharing and using students' personally identifiable information (PII). People often misunderstand what is allowable under FERPA, and use it as a reason to withhold important information. Understanding whether your data "ask" falls within the allowable sharing scenarios outlined in FERPA is crucial.

Resources

- [FERPA Sherpa](#): This website, a project of the Future of Privacy Forum and the Data Quality Campaign, provides service providers, parents, school officials, and policymakers easy access to laws, best practices, and guidelines that are essential to understanding education privacy.
- [Stoplight for Student Data Use](#) (DQC): This tool summarizes some of the main provisions of FERPA and related regulations and identifies when students' personally identifiable information may be securely shared under the law.
- [State Policymakers Guide to Complying with FERPA](#) (DQC): This resource outlines the roles and responsibilities state leaders have in protecting student privacy and complying with FERPA.
- [Privacy Technical Assistance Center](#): The U.S. Department of Education established the Privacy Technical Assistance Center (PTAC) as a "one-stop" resource for education stakeholders to learn about data privacy, confidentiality, and security practices related to student-level longitudinal data systems and other uses of student data. States can request guidance and technical assistance from PTAC on topics like FERPA compliance.

■ Understand Your State's Laws

Over the last three years, states have passed a large numbers of student data privacy laws, and sometimes more than one in a given state. In many cases, the new laws have changed what and how information can be accessed and

shared. Privacy concerns from parents and the community will continue to present challenges to data sharing, even when allowed by law, so understanding the full extent of your state’s laws will be critical to your ability to effectively advocate.

Resources

- [Student Data Privacy Legislation: A Summary of 2016 State Legislation](#) (DQC): This resource provides a summary and analysis of student data privacy legislation that was introduced and/or passed by states in 2016.
- [Policy Update: Trends in Student Data Privacy Bills in 2016](#) (National Association of State Boards of Education): This document provides additional analysis and information about student data privacy legislation in 2016.

■ Understand How Your Work Advances State Priorities

States get endless requests for data and are always under pressure to protect student privacy. If you are requesting data, or even advocating that the state make more information publicly available, understanding and explaining how your priorities help advance the state’s priorities is a great way to move your request to the front of the line. Consider drafting your own position and ideas specific to n-size to help states understand how that decision will effect important work on the ground. The more information you can provide about why you need the information and how it will help, the less gray area exists to make state leaders skeptical.

■ Push States to Be Transparent

States have an obligation to be transparent about how students perform, and this includes clarity around why and how minimum n-sizes are chosen. Regardless of a state’s choice, they should communicate throughout the decision-making process and seek input from key stakeholders.

Privacy and transparency are not mutually exclusive. States can keep student information safe while also providing a clear and accurate picture of how all students perform. States can start to do this by setting two different n-sizes, one for accountability and one for public reporting. Accurate data and transparency are essential to meeting state equity goals and building trust with the public. States will find it easier to establish and use smaller n-sizes if the public trusts the state as a good steward of student information.

Resources

- The Ohio Department of Education produced a [webinar](#) series to help the general public understand the n-size issue, showing specific examples of what level of information is available when different n-sizes are used.
- [Making Students Visible: Comparing Different Student Subgroup Sizes for Accountability](#) (Policy Analysis for California Education): Using data from California’s CORE Districts, this policy memo explores the effects of using different subgroup sizes. The authors identify the 20+ subgroup size as optimal for making sure that historically underserved student populations are visible.
- [Ensuring Equity in ESSA: The Role of N-Size in Subgroup Accountability](#) (Alliance for Excellent Education): This report details why states should consider smaller n-sizes to ensure that student subgroup performance is not inadvertently masked.

In 2016, **only 15 states** included an explanation of their n-size on their state report card. Was your state one of them?
For more information on state report cards, check out [DQC’s latest report](#).