



Cell Suppression Policy for *IDEA* Data December 22, 2009

Beginning with the 2003-04 data, the data tables on www.ideadata.org contain cells in which the numbers are suppressed. Certain data are suppressed to limit disclosure of information consistent with Federal law, which, according to 34 CFR 99.35(b)(1) “requires that information collected by the Secretary under this authority [34 CFR 99.31(a)(3)(iii) and 99.35] be protected in a manner that does not permit personal identification of individuals by anyone except those officials [i.e., authorized representatives of the Secretary in connection with an audit or evaluation of Federal or State supported education programs, or for the enforcement of or compliance with Federal legal requirements which relate to those programs].”

It is the policy of the U.S. Department of Education (Department) to be consistent with the provisions of privacy statutes. Each office in the Department has different purposes for its data collections. Therefore, each office develops its own approach to data presentations that ensures the protection of privacy while meeting the purposes of the data collection and the Department’s Information Quality Guidelines. The Office of Special Education Programs’ (OSEP) cell suppression policy was developed in consultation with:

- statisticians from the National Center for Education Statistics (NCES);
- the *EDFacts* Data Governance Disclosure Risk Working Group;
- Budget Service in the Office of Planning, Evaluation and Policy Development (OPEPD); and
- the Office of Communications and Outreach (OCO).

In preparing the *IDEA* data for the public, OSEP determined that certain numbers (cells) in the tables would be suppressed in order to prevent the identification of children and students through data publication. All counts of fewer than five children, excluding counts of zero children (i.e., all counts of one through four) are suppressed. Additional cells are suppressed when necessary to prevent the calculation of another suppressed cell. Note that when cell numbers are suppressed, the associated percentages are not calculated and displayed in the tables. However, national totals in tables broken down by state are not suppressed because there is no real risk of individual identification. The data collected on assessment, dispute resolution and personnel are not suppressed, regardless of the cell value, because these data are not counts of individual children or students.

In recent years, OSEP has explored alternative cell suppression treatments in order to reduce the loss of information via cell suppression. In 2008, OSEP decided to eliminate the suppression of zero cell values and applied this revised methodology to all data retroactively, starting with the 2003-04 data. The justification for this change is that more data and important data (e.g., the absence of children with a particular disability) are available to the public, and privacy is not significantly compromised.

Cell suppression affects some tables, such as those that present multiple variables, more heavily than others. Usefulness of tables more heavily affected by cell suppression (e.g., infants and toddlers ages birth through 2 served under *IDEA*, Part C, by race/ethnicity, early intervention setting and state) must be balanced against children's and students' right to privacy.

It is important to note that cell suppression is at the database level. Once the cell suppression programs (algorithms) have been applied to the data, there are two rounds of quality control (QC). The first round ensures that all cells that ought to be suppressed are suppressed; a check for undersuppression. The second round of QC looks for oversuppression, instances where cells were suppressed that should not have been.

The importance of making the *IDEA* data available to the public must be balanced against the legal responsibility of protecting individual privacy. In light of these two obligations, OSEP reviews the data and the procedures used to suppress data on an ongoing basis to ensure that the public information needs are optimally served.