

Using Assessment Data as Part of a Results-Driven Accountability System

Input from the NCEO Core Team

July 17, 2012*

* Revised August 24, 2012 to ensure consistent terminology with other OSEP documents.

This document was prepared by the National Center on Educational Outcomes (NCEO) Core Team for consideration by OSEP. NCEO is supported through a Cooperative Agreement (#H326G110002) with the Research to Practice Division, Office of Special Education Programs, U.S. Department of Education.

Dear Colleague:

On March 2, 2012, the Office of Special Education Programs (OSEP) announced we would be taking steps to help close the achievement gap for students with disabilities by moving away from a one-size-fits-all, compliance-focused approach to a more balanced system that looks at how well students are being educated, in addition to continued efforts to protect the rights of students with disabilities. OSEP's vision for the new Results-Driven Accountability (RDA) system is that all components of accountability will be aligned in a manner that best supports States in improving results for infants, toddlers, children and youth with disabilities, and their families.

The Individuals with Disabilities Education Act (IDEA) requires that the primary focus of monitoring be on improving educational results and functional outcomes for children with disabilities, and ensuring that States meet the IDEA program requirements. The current system places a heavy emphasis on procedural compliance without consideration of how the requirements impact student learning and postschool outcomes. In order to better fulfill the IDEA's requirements, a more balanced approach to determining program effectiveness in special education is necessary.

The major components of RDA will include: (1) the State Performance Plan and Annual Performance Report (APR) and State Status Determinations (required under section 616 of the IDEA); and (2) a differentiated system of monitoring and technical assistance (TA) that will support all States but especially those States with the most significant needs for improvement. Performance of States relative to other States and to national data will be determined using data on priority elements, and will be used to determine the appropriate level of monitoring and technical assistance.

As an initial step to identifying States most in need of improvement, I tasked the National Center on Educational Outcomes (NCEO) to convene a team of national experts on assessment and accountability systems to do a thoughtful examination of available data that help in understanding how students with disabilities are performing. The following document is the team's report and includes possible applications of their work. This is the first step that we have taken to use educational results in a more meaningful way. We are currently looking at other data sources that may also be useful.

Thank you to NCEO, especially Martha Thurlow, Sheryl Lazarus, Kamarrie Davis, and Laurene Christensen, for their thoughtful work on this project. I also greatly appreciate the contributions of members of the Core Team and Resource Group, and acknowledge the extensive commitment of time and expertise by all those involved.

Finally, we will continue to call on State staff, parents, TA experts, data experts, advocates and others to provide input as we move forward in the development of RDA. In addition, we will continue our broad-based communication efforts through the Department's web-site, conference calls, webinars, conference presentations and individual meetings with interested parties. I look forward to working with all of you to develop an accountability system that focuses on improving results for infants, toddlers, children and youth with disabilities while continuing to ensure that the rights of children and families are protected.

Sincerely,

Melody Musgrove
Director
Office of Special Education Programs

Executive Summary

The National Center on Educational Outcomes (NCEO) was asked by the Office of Special Education Programs (OSEP) to work with a team to provide input on measures that could be used to review states' performance results of their students with disabilities who receive special education services. This represents a significant shift toward results-driven accountability for states.

NCEO established two groups to work on this task. A Core Team was established to identify measures of the academic performance results of students with disabilities who receive special education services. A Resource Group was established to support and serve as a "critical friend" to the Core Team to critique and support the refinement of the proposed measures. Through intensive and thoughtful conversations, the Core Team first identified several framing considerations to guide its work, and then generated a document for consideration and reaction by the Resource Group (see Appendix A for member names).

Framing Considerations

The Core Team identified seven framing considerations for developing measures of student performance that could be used for OSEP's review of student performance:

1. Public transparency and understandability are critical features of an accountability system and must be reflected in measures used to review performance.
2. Multiple measures must be included. No single measure should be used in decisions about student performance results.
3. The use of measures of student performance should provide appropriate incentives to states, particularly in relation to identified values (e.g., inclusion in the general assessment).
4. The measures should provide a flag to look deeper into areas that need improvement.
5. A plan should be developed and steps taken to monitor, validate, and improve the use of measures by OSEP and others; additional variables may be appropriate to enhance the measures in the future.
6. Variables that may be related to student performance but that have inconsistent interpretations and reliability should not be included in measures that are used for reviewing states on the performance of their students with disabilities.
7. No increased burden on states to collect additional data should result from the shift to using student performance results. The developed measures need to fit within what states are doing as they review districts, and should be compatible with and reflective of the state's overall accountability system used for school improvement.

Each of these considerations is discussed briefly in the report.

Core Team Suggestions

The Core Team identified currently available, assessment-based data sources that represent variables for understanding student performance. The team suggested that the variables be presented in several tables. The tables are presented in a reporting format to ensure that multiple measures are considered. The Core Team also suggested that data be provided separately for reading and mathematics, and that participation on both the general assessment and alternate assessments should be included in tables.

The Core Team suggested that the performance of students with disabilities on the general assessment be included as a variable, along with a variable to indicate the relative difficulty of

each state's assessment. The team also suggested that a variable be included that shows the gap in performance on the general state assessment between students with disabilities and students without disabilities, a variable that shows the difference between performance of all students with disabilities and the state target for their performance, and a variable showing improvement over time. The team noted that some of these variables should be treated as contextual variables and others as primary variables for consideration. The team also indicated that variables that are not included at this time (e.g., LRE, growth) might be added in the future as accurate data become available.

Suggested Reporting Format

The reporting format suggested by the Core Team includes six tables. The first two tables provide a view of the performance of students with disabilities on the reading and mathematics general assessment in each state. The third table provides an indication of the performance of all students with disabilities in relation to state-defined targets. Tables 1 and 2 (and Table 3 as data improve) would provide the primary information for OSEP to consider in reviewing states on the performance of students with disabilities.

Table 1 (reading) and Table 2 (mathematics) include the following variables: (a) percentage of students receiving special education services who scored proficient or above on the general assessment; (b) participation rates for the general assessment; (c) the gap in performance (in terms of students scoring proficient or above) between students without disabilities and students with disabilities on the general assessment, (d) the change over a two year time span in the percentage of students receiving special education services who scored proficient or above on the general assessment, and (e) the relative difficulty level of each state assessment, in terms of a comparison to a common metric (e.g., NAEP scale). Table 3 provides information on each state's reading and mathematics assessment targets (in the State Performance Plan, Indicator 3C), the actual proficiency rates reported by states in their Annual Performance Reports, and the difference between the two.

Tables 4 and 5 provide a contextual view of the performance of students who participate in alternate assessments in reading and mathematics in each state. These tables summarize data for the alternate assessment based on alternate achievement standards (AA-AAS) and, if the state had one, the alternate assessment based on modified achievement standards (AA-MAS). Table 6 provides detailed disaggregated information about the participation rates for each assessment option (e.g., general assessments, AA-AAS, and AA-MAS).

Summary and Conclusions

The Core Team believes that OSEP should use multiple measures related to the performance of students with disabilities as part of a result-driven accountability system because no single available measure adequately represents the full range of important performance variables to consider. The team agreed not to suggest a single formula that combines variables into a single measure of student performance for several reasons, including the team's view that such an approach at this time is unlikely to provide an understandable or valid view of how students with disabilities are performing in a state, due primarily to the lack of proven measures and the many complex ways that multiple variables may interact. Based on its knowledge and expertise, the Core Team identified what it believes are the most important and reliable variables publicly available now to include in reviewing state performance. The Core Team's input reflects an emphasis on public reporting, and on rich indicators that could be weighed via good professional judgment to inform interpretations of the quality of state performance of students with disabilities.

In the experience of Core Team members, creating a formula to make decisions that prompt different levels of ratings and support is a complex and challenging endeavor. If OSEP decides to develop such a rating-and-differentiated accountability system, the Core Team strongly recommends that stakeholders, experts, and OSEP be involved in the discussion about policy judgments and technical decisions on how these multiple measures are used for making decisions related to student performance.

Using Assessment Data as Part of a Results-Driven Accountability System *Input from the NCEO Core Team*

Background

The National Center on Educational Outcomes (NCEO) was asked by the Office of Special Education Programs (OSEP) to work with a team to provide input on measures that could be used to review states on the performance results of their students with disabilities receiving special education services under the Individuals with Disabilities Education Act (IDEA).¹ This work represents a significant shift from a compliance-based to a results-driven accountability system for states.

NCEO established two groups to work on this task. A **Core Team** was established to identify measures of the academic performance results of students with disabilities who receive special education services. A **Resource Group** was established to support and serve as a "critical friend" to the Core Team to critique and support the refinement of the proposed measures. Members of the Core Team and Resource Group are presented in Appendix A, along with the NCEO staff members who worked with the teams.

Assumptions Provided to the Teams

In addition to the assumption that OSEP would be moving to a results-driven accountability system, the Core Team and the Resource Group worked under a set of assumptions about the measures that would be used to define "results":

- Existing, publicly available, data will be used.
- All IDEA-eligible students with disabilities (on IEPs) should be considered for the measures (students in regular assessment and alternate assessments).
- The measures should be understandable by those being reviewed.
- The measures may change over time as more or different data become available.
- States could use similar measures to review their districts.
- Suggestions should be supported by a rationale.

Process for Generating Suggestions

The Core Team met on three occasions and communicated in the intervals between the meetings. Through intensive and thoughtful conversations, the Core Team first identified

¹ Throughout this document, the term "students with disabilities" will be used to refer to students who are IDEA-eligible. It does **not** include former students with disabilities, even though those students may be included by states in their accountability calculations for the Elementary and Secondary Education Act (ESEA), as allowed under current ESEA regulations. It **does** include English Language Learners with IDEA-eligible disabilities.

TGRG clarification (OEC) 12-6-12

How should accommodations be utilized on diagnostic assessments under the Third Grade Reading Guarantee?

The assessment manual for the diagnostic assessment used must be consulted to determine allowable accommodations. An accommodation should not be used if it interferes with the accurate assessment of reading (the intent of the diagnostic assessment). (See 4 criteria below.) Therefore, some accommodations might be allowed if it does not change what the test is intended to measure. For example, if a visually impaired child's IEP specifies large print, that would be allowable because it meets the 4 criteria below.

There are four criteria for accommodations on the OAA and OGT state assessments specified within (OAC 3301-13-03(H); see below. These four criteria should be considered when determining accommodations for the reading diagnostic assessment for the TGRG in K, 1, 2, and 3.

From the page 38 in the Rules Book (PDF) (Updated 9/27/12) posted on ODE:

“Accommodations are changes made in how a student has access to the curriculum or demonstrates learning. Accommodations provide equal access to learning, do not substantially change the instructional level or content, are based on individual strengths and needs, and may vary in intensity or degree. In testing situations, accommodations are changes in format, response, environment, timing, or scheduling that do not alter in a significant way what the test measures or the comparability of the scores.” (Batshaw, M. “Children With Disabilities”, 2002, p. 603)

Accommodations on State Assessments

OAC 3301-13- 03(H) specifies these four criteria for allowable accommodations:

- The accommodation is specified in a student with disabilities' IEP (or 504 Plan) and is provided for classroom and district-wide assessments. The accommodation must be documented in writing on the testing page of the IEP or 504 Plan BEFORE the student takes a test.
- The accommodation does not change the content or structure of an assessment. The accommodation does not change what type of knowledge or skill that an assessment is intended to measure. For example, the test administrator may not convert open-ended questions to multiple-choice questions.
- The accommodation does not change what type of knowledge or skill that an assessment is intended to measure. For example, the test administrator is not permitted to read the passages from a reading test, because this would change the test from a measure of reading skills to a measure of listening skills. Similarly, calculators are not allowed on the grades 3 and 4 mathematics test. Grammar checkers and spell checkers are not allowed on the OGT writing test.
- The accommodation does not change or enhance a student's response as to what type of knowledge or skill is intended to be assessed, but it facilitates how the response is provided or accessed. For example, a scribe may record only the actual response provided by a student. A scribe may not enhance, clarify or elaborate on a student's response.

To find the Rules Book (above excerpt is on page 38) please go to ODE's website at www.ode.state.oh.us and type "Ohio Statewide Assessment Program Rules Book" into the search box.