Guidance for Examining District Alternate Assessment Participation Rates

Federal policy limits the number of students that a state may assess with an alternate assessment aligned to alternate achievement standards (AA-AAS) to no more than 1% of all students in the grades assessed in a state. The AA-AAS is intended only for students with the most significant cognitive disabilities. Given that many states report more than 1% of their student population taking an AA-AAS (Thurlow & Wu, 2018), there is a need to identify effective and technically defensible practices for monitoring participation rates for alternate assessments.

NCEO
National Center on Educational Outcomes

Center for Assessment

This Brief addresses important questions that arise for states as they implement federal AA-AAS participation requirements. These questions include:

- What are some approaches for dealing with the uncertainty associated with small student populations (i.e., n-sizes)?

- What are some appropriate analytical approaches for monitoring AA-AAS participation rates?

- What are the strengths and limitations of various methods?

- What criteria could be established for defining "expected" versus "exceptional" fluctuations in district participation rates?

ESSA Requirements

In 2015, Congress reauthorized the Elementary and Secondary Education Act of 1965 as the *Every Student Succeeds Act* (ESSA). With this reauthorization the requirement changed for AA-AAS. ESSA reaffirmed that an alternate assessment is the appropriate assessment for students with the most significant cognitive disabilities to demonstrate their knowledge and skills. However, ESSA changed the 1% cap to be based on the **participation rate** rather than the proficiency or performance rate as required under *No Child Left Behind.* It placed the 1% cap on the **state** participation rate for each subject, based on the total number of all students in the state assessed in the subject (34 CFR 200.6(c)(2)).

The federal law under ESSA also specifies that states cannot place a cap on the participation rates of local education agencies (LEAs[1]) in any subject (34 CFR 200.6(c)(3)(i)). This means that LEAs can exceed the 1% participation cap on an AA-AAS in a given subject, but the state as a whole cannot exceed the 1% cap in any subject. The law requires LEAs that exceed the 1% participation cap to submit information justifying the need to exceed the cap and the state must provide oversight of those LEAs (34 CFR 200.6(c)(3)(ii-iii)).

If a state anticipates that it will exceed the 1%

participation cap in any subject in any school year, the state may request that the U.S. Secretary of Education waive the cap for the relevant subject for one year (34 CFR 200.6(c)(4)). The request must include:

- Assurances from the state that it has verified that each LEA anticipated to exceed the 1% cap will address any disproportionality in the percentage of students in any subgroup taking an AA-AAS (34 CFR 200.6(c)(4)(iii)(B)).

- A plan and timeline by which the state will take additional steps to support and provide appropriate oversight to each LEA that the state anticipates will assess more than 1% in a given subject in a school year to ensure that only students with the most significant cognitive disabilities take an AA-AAS and the state must describe how it will monitor and regularly evaluate each such LEA to ensure that the LEA provides sufficient training to school staff so that all students are appropriately assessed (34 CFR 200.6(c)(4)(iv)(B)).

- Confirmation that the state will address any disproportionality in the percentage of students taking an AA-AAS as identified through the data provided (34 CFR 200.6(c)(4)(iv)(C)).

The difficulty many states will likely encounter as they address ESSA requirements is how to use data to accurately and defensibly identify LEAs that need oversight and monitoring because of their AA-AAS participation rate. The small n-sizes of students with the most significant cognitive disabilities can make annual participation rates and fluctuations in those participation rates appear much more significant than they are in reality. Additionally, the analytic approaches and underlying assumptions that work to flag districts with problematic participation rates on the state assessment do not necessarily work in the same way to flag districts with problematic participation rates on an AA-AAS.

A Unique Context

Small sample sizes present a unique challenge when attempting to identify districts because typical

[1]The term LEA and district are used synonymously in this Brief.

analytic methods may be biased with such small sample sizes. For example, a small difference in the number of students taking an AA-AAS within a district may result in a participation rate exceeding 1% or a significant spike in the participation rate from the previous year. Without knowledge of the context it is impossible to know whether these values are "expected" or "exceptional" and therefore whether the district needs additional support and oversight.

Participation rates in the AA-AAS have gradually increased across years, from less than 1% in most states in 2000-2001 (Thurlow, Wiley, & Bielinski, 2002) to more than 1% in most states in 2015-16 (Thurlow & Wu, 2018). Thurlow and Lazarus (2017) argued that all states should be able to meet the federal 1% participation cap if appropriate policies, practices, and procedures are in place, including more professional development for IEP teams within districts or better support and monitoring. Federal policy allows for district-level variability by not permitting states to set a district participation cap, in part because there could be valid reasons for variations at the district level in the percentage of students who take an AA-AAS.

When requesting a waiver, states should explain how they identified districts with more than 1% participation in the AA-AAS, as well as how they will monitor and evaluate district participation rates. States will need a defensible approach that takes into account uncertainty due to small n-sizes in order to accurately classify districts into categories for additional monitoring and support.

There are important reasons to get this right. The consequences of making an error in these identifications can have real implications for students with the most significant cognitive disabilities. This student population needs access to the most appropriate services available. Two types of error (Type I–False Positive and Type II–False Negative) can affect the ability of districts to provide those services. Both of these types of error may result in serious consequences for serving students with the most significant cognitive disabilities.

Type I error occurs *if a state incorrectly identifies a district in need of monitoring and support when the district has not incorrectly assessed any students on an AA-AAS*. This type of error may drain time and resources that could otherwise be directed more productively in helping to provide the best educational services available to students with the most significant cognitive disabilities. Type I error could also lead to a district being erroneously flagged or concerned that they may be erroneously flagged, with the result being that it fails to take necessary action and provide appropriate services to students with the most significant cognitive disabilities because of its concerns about how such action will impact the district's participation rate and subsequent state monitoring.

Type II error occurs *if a state fails to identify a district in need of monitoring and support when the district does in fact need oversight due to potentially incorrectly assessing some students using an AA-AAS*. Students who participate in the AA-AAS work on the same grade-level content as other students, but the expectations for achievement differ in depth, breadth, or complexity. Type II error in this situation means that some students may be given an assessment not appropriately matched to their learning needs.

Guiding Principles

There is no single correct way to monitor and evaluate participation rates in an AA-AAS. Still, there are four guiding principles that characterize approaches likely to be more effective.

1. *A comprehensive solution to identify districts in need of additional monitoring and support on participation rates cannot be purely empirical*. Small n-sizes make it difficult if not impossible to use purely empirical analyses to classify districts into monitoring categories. States will need to combine data analyses with thoughtful inquiry that is tailored to the unique context of the state and districts.

For example, one approach is to combine the data analyses with careful investigation into what does not seem typical or likely based on state trends and

district context in the current year and prior years. This approach requires both collecting evidence from empirical data analyses and a follow-up inquiry that takes into account the unique local context.

2. *It is important to detect atypical or "exceptional" values*. Exceptional means that the values do not seem likely given what is known about the state or district context from former, current, and multi-year data. Detecting atypical values can leverage research-based evidence such as:

- Students with the most significant cognitive disabilities tend most often to be in three federal disability categories (intellectual disabilities, autism, and multiple disabilities);

- Students with the most significant cognitive disabilities tend to enter the special education identification process during early childhood, rarely exit during their school years, and rarely show spikes in participation in the AA-AAS even though it might be reasonable to expect participation to increase in higher grades; and

- Although the performance of students with the most significant cognitive disabilities often surprises us when they are provided access to grade-level content, they generally do not have significant score variations from year to year (e.g., score below proficient for multiple years, then jump to advanced proficiency in one year) if access to instruction does not change. This is especially atypical when it occurs for multiple students.

3. *There should be a method applied to deal with uncertainty*. Uncertainty is a reality for any inferential statistical analyses where one attempts to make judgments about the likelihood that an observed difference (such as a district's participation rate being greater than 1% in a subject area) might have happened by chance alone. Uncertainty is a critical issue in this situation because of the small n-sizes. Uncertainty can be quantified into what is probable or likely using different methods.

4. *The culminating decision and subsequent actions based on the evidence is (a) a matter of degree, and*

(b) related to unique context and circumstances. Therefore, findings should be considered holistically because no single piece of evidence is likely sufficient to form a firm conclusion about whether the participation rate is "expected" or "exceptional." Evaluating the collection of results to describe a level of confidence along a continuum is preferred.

Suggested Analytic Approaches

There are two types of analytic approaches that we suggest: (a) methods for detecting atypical values, and (b) methods for dealing with uncertainty. Tables 1 and 2 summarize these as well as their benefits and limitations.

Methods for Detecting Atypical Values

States have submitted information on the number of students who take an AA-AAS to the federal government for years. Current and former years' data can be used to flag "exceptional" values—values that do not seem likely given what is known about the district context and what is known about students with the most significant cognitive disabilities. Three approaches for using these data are: (a) current or former year analyses, (b) multi-year analyses, and (c) performance trends (see Table 1).

Table 1. Suggested Approaches for Detecting Atypical Values

Approach	Description of Approach	Benefit(s) of Approach	Limitation(s) of Approach
Current or Former Year Analyses	Check for atypical or unusual (e.g., out of range) values. Use descriptive statistics (M, SD, Min, Max, and Range) and visual data displays (histograms, boxplots, etc.) to locate values of interest. Flag districts with participation rates that seem really high, really low, or atypical Evaluate whether students from unexpected disability categories (e.g., specific learning disabilities, speech and language, etc.) are participating in the AA-AAS. Flag districts with atypical findings or patterns.	Provides insight into any values that seem really high, really low, or atypical in some way. Provides a baseline of what seems typical in any given state context.	Due to small n-sizes, even small fluctuations can appear like larger differences.
Multi-Year Analyses: Longitudinal Trends	Compare district participation rates over the past 3 to 5 years at the state level. Compare district participation rates over the last 3 to 5 years for each district. Flag districts with the largest difference in participation rates.	Helpful to use state and district trends to contextualize findings and examine magnitude of change. Comparing participation rates within a district over time can help control for some factors that may influence district participation that may differ from district to district.	There may be reasons why there is a large fluctuation in participation rates over a period of time within a district such as the start of a school within a district that services students with the most significant cognitive disabilities. Changes in the alternate assessment used within a state could affect performance level trends. Some regions within states may have more student entry from out of state which could potentially affect results for certain regions.
Multi-Year Analyses: Cross-Sectional/ Cohort Trends	Evaluate student participation entry within a district over multiple years. Calculate the proportion of new examinees at the cohort level by subject. Flag outlier districts for further review. Evaluate performance changes within a district over multiple years. Calculate the proportion of students with large performance changes (e.g., move two classification levels in one year). Flag districts with higher than expected proportions for further review.		
Performance Trends	Evaluate district performance distributions for atypical shifts or spikes in performance (e.g., many more students scoring advanced in one year than in other years) alongside district participation rates. Flag districts with atypical findings. Compare district performance distribution with the overall state performance distribution on the same AA-AAS and examine a district's performance distribution over time on the general state assessment. Flag districts with atypical results.	Uses another source of evidence (performance) alone or combined with participation to investigate district results. Helpful to use state and district trends to contextualize findings and highlight atypical results.	Changes in the alternate assessment or general assessment used within a state could affect performance level trends.

Current or Former Year Analyses

Because the federal law requires states to apply for a waiver if they *anticipate* they will exceed the 1% participation cap on an AA-AAS, it is likely that many states will use former year data initially. The most basic and yet critical step in analyzing any data is to look for atypical or unusual (e.g., out of range) values and use descriptive statistics to get a general sense of the sample's central tendency, spread, and variability. Out-of-range values can be quickly flagged and districts can be contacted to verify data. The descriptive information can also help states locate values of interest to flag for further follow-up and review. This involves computing the mean, standard deviation, minimum value, maximum value, and range of values alongside visual data displays (e.g., histograms, box plots, etc.) to identify district participation rates that seem really high, really low, or atypical in some way.

Visual data displays also provide important evidence about the distribution of LEA participation rates in the former year and in the current year. Do the distributions fit what the state expected? Do the distributions look really different from year-to-year? Is there a wide range of LEA participation rates or are they mainly clustered around some percentage?

Because available data indicate that most students who participate in an AA-AAS are in the disability categories of intellectual disabilities, autism, and multiple disabilities (Kearns, Towles-Reeves, Kleinert, Kleinert, & Thomas, 2011; Thurlow, Wu, Quenemoen, & Towles, 2016), it is also useful for a state to evaluate whether students from unexpected disability categories are participating in the AA-AAS. Thurlow and Lazarus (2017) recommended that even though disability category cannot determine whether a student is one with a significant cognitive disability, a state can investigate particular cases (or patterns of cases) especially where students with speech-language impairments, specific learning disabilities, emotional disabilities, other health impairments are participating in an AA-AAS in unexpected numbers to confirm that it is the most appropriate assessment for each student.

Benefits. It is always important to start any data analyses by looking at the values of interest for anything that seems off, and to ground other analyses in the general characteristics of the data. This approach provides a quick way to identify any values that are out of range, how much variability there is around the mean, the spread of the data, and what the distribution looks like. Furthermore, the examination of participation rates by disability category allows a state to evaluate the extent to which districts are potentially mis-identifying some students who typically do not have a significant cognitive disability even though they may have a federally-recognized disability. This information can provide the state an opportunity to provide districts with more professional development and support about how to correctly identify students with the most significant cognitive disabilities.

Limitations. Small total student populations can make it appear like there are larger differences in LEA participation rates between districts than is actually the case. This may be especially true for districts whose total student population is smaller. For example, if a district has 10 students out of 500 total students who take an AA-AAS then that district's participation rate would be 2%. If a district has 10 students out of 5,000 total students who take an AA-AAS, then that district's participation rate would be 0.2%. The first district would be above the expected participation rate and the second district would be below only because the total student population is larger. There could be many reasons why a smaller district would have the same (or more) students with the most significant cognitive disabilities than a larger district. For example, a public, special education school that specializes in serving students with the most significant cognitive disabilities may be situated in the small district. This is why thoughtful inquiry and qualitative follow-up is so important in this context.

Multi-Year Analyses: Longitudinal Trends and Cross-Sectional/Cohort Trends

The use of multi-year longitudinal data can reduce issues related to small sample sizes and can also allow

for the examination of the magnitude of change over time at both the state- and LEA-level. For example, one can compare LEA participation rates over the past 3 to 5 years using a line graph. What is the general trend? Are participation rates going up or down? Is the year-to-year difference holding steady or is it changing more rapidly across years?

This approach provides insight into the magnitude of change within LEAs over time and which LEAs have the largest difference in participation rates over an identified period of time (e.g., from Year 1 to Year 3 or from Year 1 to Year 5). A flagging rule could be applied wherein districts are flagged if they have atypical longitudinal trends, unusually high or low values, or large differences in participation rates over time.

Cross-sectional/cohort trend analyses leverage some key assumptions about students with the most significant cognitive disabilities. These assumptions are that most of these students tend not to enter the special education identification process after starting school and do not typically have significant score variations from year-to-year (e.g. score below proficient for multiple years and then jump up to advanced proficiency in one year) if access to instruction does not change across years. These two assumptions allow a state to use cross-sectional/cohort trend data to evaluate student participation entry on an AA-AAS at the district level. For example, the proportion of new examinees at the cohort level per subject area can be used to flag LEAs with higher than expected or unusual spikes in proportions of new examinees for further follow-up and review. It may be reasonable to expect that participation in an AA-AAS will increase in the higher grades, especially if students are encouraged to take the general assessment in elementary school rather than immediately making the decision for them to take the alternate assessment.

A similar process can be used to evaluate the proportion of students with large performance changes (e.g., move two classification levels in one year). In districts that historically have not provided AA-AAS participants with access to grade-level content, but then shift to providing access to the curriculum, a performance increase is expected. Still, we would not expect to see spikes in the proportion of students with large performance changes every year, especially for multiple students. Large variation in performance from year to year could indicate a myriad of issues within districts that the state could investigate to provide the necessary support and professional development around identifying students with the most significant cognitive disabilities as well as appropriate instruction related to grade-level content standards.

Benefits. It is helpful to use state and LEA trends to contextualize findings. LEAs differ from one another for many reasons and comparing participation rates and performance results within a state through longitudinal trend analyses and within LEAs through cohort analyses over multiple years controls for some factors that likely influence student participation and performance on an AA-AAS that may differ from district to district. Multi-year analyses are likely to be the most robust analyses and could be weighted more heavily in the collection of evidence used to classify districts into monitoring categories.

Limitations. First, there may be valid reasons to explain large fluctuations in participation rates over time within an LEA such as the opening of a school that specializes in serving students with the most significant cognitive disabilities. Follow-up inquiry with districts would likely uncover these nuances. A second limitation applies to many state assessment systems at present—the comparability of results between different assessments used within a state. Changes in the test administration such as a switch to online, computer-based administration or a change in the actual assessment itself could affect performance level trends and bias findings. These effects on performance level trends generally would be similar across LEAs, but that may or may not be true. Another limitation is that some regions within states may have more student entry from other states or from other districts, which could affect the results of those analyses for certain regions more than others. Context matters and it is imperative to add thoughtful

inquiry and qualitative follow-up investigations to any statistical analyses to ensure the most accurate interpretations and uses.

Performance Trends

It may also be important to look at other variables as well as participation. Atypical shifts or spikes in district performance on an AA-AAS could be a signal of something unusual that merits additional scrutiny. For example, substantial shifts in a district's performance distribution (e.g., many more students scoring advanced in one year than in another year) combined with an increase in that district's participation rate on an AA-AAS may indicate a finding that needs to be explored more carefully.

It also may be useful to compare a district's performance distribution on the AA-AAS with a norm-referent such as the overall state performance distribution on the AA-AAS. This comparison could highlight unusual or atypical district performance distributions on the AA-AAS, which combined with other evidence may indicate a need for additional follow-up and review.

Comparison of performance distributions is not limited to the alternate assessment. The state could also examine a district's performance distribution over time on the general state assessment to see whether there are any atypical shifts or spikes.

Benefits. This approach provides another source of data that can be used to identify districts that are potentially incorrectly assessing some students using an AA-AAS. Because suggested analyses compare performance within districts across years or compare district performance using the state as the norm-referent, the results are contextualized and control for other factors likely to bias results.

Limitation. The usefulness of examining performance trends is limited if there have been changes in the alternate assessment or the general assessment used within the state. Changes in test administration or in the actual assessment could affect performance trends and bias findings.

Methods for Dealing with Uncertainty

A difficulty in identifying "expected" versus "exceptional" district AA-AAS participation rates is the small number of students with a most significant cognitive disability who typically take an AA-AAS in any given year. Small sample sizes can be incompatible with some inferential statistical approaches that use sample data to assign the probability or likelihood of an event occurring by chance alone. Still, there are methods for addressing the uncertainty associated with the data states receive from districts: (a) confidence intervals, (b) and multi-year averages (see Table 2).

Confidence Intervals

A confidence interval (CI) is used to provide an indication of how likely it is to observe a value in a distribution. For example, a 95% CI might depict individual LEA participation rates from current or former years to give a range of values likely to occur given previous observations. Values within the interval are plausible, whereas values outside the interval are exceptional (e.g., occur less than 5% by chance). This level of certainty can be adjusted up or down—up to 99% confidence (1% uncertainty) or down to 90% confidence level (10% uncertainty)—based on the state's tolerance for error (Type 1 or Type 2). Higher CIs estimate a smaller range of exceptional values and therefore attempt to minimize Type 1 error (incorrectly identifying a district), whereas lower CIs estimate a larger range of possible exceptional values and therefore attempt to minimize Type 2 error (failing to identify a district in need of support and oversight).

Benefit. CIs provide a flagging criterion using current or previous data. In this way the CI is based on the local context and if any LEA participation rate falls outside of the CI then that LEA can be flagged for additional follow-up and review.

Limitations. To estimate a CI an appropriate sampling distribution must be chosen and there must be a suitable n-size to support inferences. Given that CIs would be estimated from LEA participation rate data,

Table 2. Suggested Approaches for Examining Uncertainty

Approach	Description of Approach	Benefit(s) of Approach	Limitation(s) of Approach
Confidence Interval	Apply a confidence interval to district participation rates from the current or former year to give a range of values that one can be certain contains the true participation rate for a state. Flag districts with rates outside of the confidence interval.	Confidence interval is based on current or previous year state data and is therefore contextualized to the state.	Distribution chosen to estimate confidence interval may not be appropriate. A suitable n-size is needed to support inferences.
Multi-Year Averages/ Rates	Calculate a multi-year average participation rate for each district by subject. Flag districts with higher than expected findings.	Helps mitigate uncertainty due to small n-sizes. Identifies districts that are consistently identifying more students with the most significant cognitive disabilities than is expected given the district student population. More robust to annual fluctuations in district participation rates.	There may be other factors that influence participation rates in districts. If those factors are not included in the way that results are analyzed then some districts could be erroneously flagged in either direction.

determining a suitable sampling distribution could be challenging. Furthermore, states with smaller numbers of LEAs should be careful about the conclusions they draw from the CI estimates because there may be more random error associated with small n-sizes (typically less than 30) and the confidence interval estimated may be so wide it will not be informative. The accuracy of CIs is context dependent and should be interpreted in light of the characteristics of the sample data.

Multi-Year Averages/Rates

A multi-year average/rate typically involves using 3 to 5 years of data from an LEA's participation rates to create a more stable basis on which to base inferences about participation patterns. A state could use multi-year averages to identify "exceptional" LEA values in light of longitudinal patterns and flag those districts for further follow-up and review.

Benefits. Multi-year analyses help mitigate uncertainty due to small sample sizes. Multi-year analyses can flag districts that are consistently identifying more students with the most significant cognitive disabilities than is expected given the LEA total student population. This approach is more robust to annual fluctuations in LEA participation rates.

Limitations. There may be other factors that affect participation rates in districts, such as total student population, regional effects related to student mobility, special education schools, and so on. If these contextual factors are not included in the way that results are analyzed, then some districts could be erroneously flagged in either direction. Thoughtful inquiry and follow-up is needed to investigate higher than expected multi-year averages.
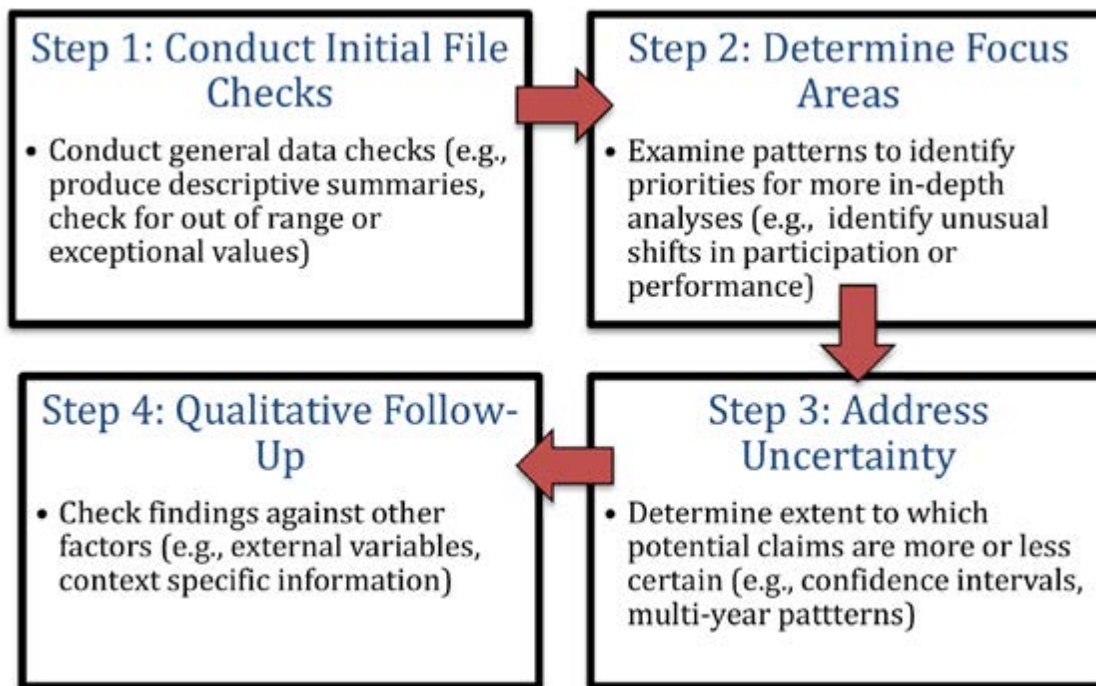
Bringing It All Together—A Guide to Implementation

Here we describe how the analysis process could be implemented. There is not a "one-size-fits-all approach" that will work for all contexts, even though any approach used should be consistent with the guiding principles. Figure 1 illustrates the recommended analysis process.

Step 1: Conduct Initial File Checks

Begin by conducting initial data file checks to verify

Figure 1. Illustration of Analysis Process

Step 1: Conduct Initial File Checks	Step 2: Determine Focus Areas
• Conduct general data checks (e.g., produce descriptive summaries, check for out of range or exceptional values)	• Examine patterns to identify priorities for more in-depth analyses (e.g., identify unusual shifts in participation or performance)

Step 4: Qualitative Follow-Up	Step 3: Address Uncertainty
• Check findings against other factors (e.g., external variables, context specific information)	• Determine extent to which potential claims are more or less certain (e.g., confidence intervals, multi-year pattterns)

data quality and to flag exceptional values. In this initial step, the main purpose is to determine whether the data are complete and correct. This also helps get a sense for the general characteristics of the data file. For accuracy, the presence of unusual or exceptional values may signal an error in the data file, which may be due to the source data, analysis, or both. Findings to attend to include: atypical numbers of missing values, any out-of-range values (e.g., an ethnicity code that is not defined, a test score that exceeds the maximum on the scale), or values that are highly unlikely (e.g., all or most students have the same test score, the demographic characteristics in one grade are sharply different than another grades).

Occasionally, an exceptional value (or values) will be accurate, but it signals something important to know about the file. For example, if district enrollment numbers sharply increased from one year to the next because a school merged with another school or program from another district, that is important information to use in the judgment process. Often investing time upfront in some rather straightforward analyses such as producing frequency tables, cross-tabulations, and descriptives (e.g., means, ranges) will yield important insights about the data

to inform subsequent steps. We also recommend visual inspection of the data to help identify any extraordinary findings to flag for further review. For example, producing a simple histogram is an easy way to check for values that are conspicuously prominent or absent.

Step 2: Determine Focus Areas

The next step is to determine focus areas. This typically involves using the suggested analytics to examine patterns to identify a finding that merits additional scrutiny. In this step the guiding question is: to what extent are the outcomes typical? For example, one could examine patterns of participation rates for districts within year and across years to detect trends that are uncommon. This step might be thought of as the basis for forming some hypothesis to study.

Step 3: Address Uncertainty

It is important to obtain a better sense of the certainty of the finding. A central question in this phase is: how likely is it that one would observe an outcome by chance alone? We recommend conducting additional analyses on the initial findings to obtain a better

sense of the accuracy or credibility of any claims. Analytic approaches may include confidence intervals and multi-year averages or rates. Confidence intervals can be very useful when n-sizes are adequate (typically 30 or more) because they can identify a range within which variation is expected. Confidence intervals are not useful when n-sizes are low. In such circumstances, the ranges produced will be so wide they will not be informative. Combining data across years will help to increase n-sizes and determine whether the outcomes are consistently observed or whether a single year is unusual.

Step 4: Qualitative Follow-Up

Empirical analyses alone will be insufficient to support a comprehensive investigation (see Guiding Principles). We suggest a qualitative follow-up with districts. This follow-up might involve inquiring about any relevant information that would help explain or contextualize "exceptional" values. By so doing, the state can then combine the empirical data analyses with the district-supplied information to make judgments about the veracity of evidence that may support classification categories.

Judgment Process

There may not be one piece of evidence that supports placing a district into a particular classification category. We recommend that the state evaluate the collection of evidence (both quantitative and qualitative) for each subject area in order to categorize districts. Possible classifications for additional monitoring and support might be: (a) weak, (b) moderate, and (c) strong evidence. These classification categories are intended to illustrate how the evidence can be compiled and evaluated in order to make a judgment about districts in need of support and oversight as stipulated in federal law, as follows.

Weak evidence for additional monitoring and support means that there is little to no evidence that suggests a district's participation rates are atypically high or unusual given state or district trends on the AA-AAS. Participation rate values are best described as "expected." Typical patterns of weak evidence would

include districts that received no flags or the flags were for minor issues easily resolved through follow-up inquiry and discussion. It is likely that most districts in a state will fall into this category.

Moderate evidence for additional monitoring and support means that there is some evidence that suggests a district's participation rates are atypically high or unusual given state or district trends on the AA-AAS. Moderate evidence would likely include few atypical values, but the results are not clear in light of subsequent investigations of uncertainty. The district may provide some contextual information that helps explain findings, but there are still some lingering questions.

Strong evidence for additional monitoring and support means that participation rates are atypically high or unusual given state or district trends on the AA-AAS. Participation rate values are best described as "exceptional." Strong evidence would likely include multiple atypical values that can be verified with certainty. The district may provide contextual information, but that information does not alleviate concerns around outcome results.

Evidence for monitoring and support should be re-evaluated each year using the most recent year of data. States could apply a decision rule that if a district is in the "moderate evidence" category for a certain number of years in a row then it will automatically move to the "strong evidence" category.

States can include the collection of evidence used to categorize districts into the "strong evidence" category in their waiver application to the U.S. Department of Education. States also can supply this information to districts to provide them with feedback on areas of concern and places where additional training, support, and monitoring are necessary.

Conclusion

Effective and technically defensible practices for monitoring AA-AAS participation rates are critical. The need for analytical approaches is unique for many

reasons, including the small sample sizes of students with the most significant cognitive disabilities. Small n-sizes make typical analytic approaches unfeasible and require a more nuanced and tailored-to-context approaches. The unintended consequences that may result for this student population from misidentifying or failing to identify LEAs in need of support and oversight are disconcerting in both directions.

Appropriate data analyses combined with thoughtful follow-up inquiry should support judgments about support and monitoring. Information generated from the analysis process described here can be used to help IEP teams within districts to better understand how to use participation guidelines to identify students with the most significant cognitive disabilities and make the most appropriate and accurate decisions for students.

References

Kearns, J. F., Towles-Reeves, E., Kleinert, H. L., Kleinert, J. O., & Thomas, M. K. (2011). Characteristics of and implications for students participating in alternate assessments based on alternate academic achievement standards. *Journal of Special Education, 45*(1), 3–14.

Thurlow, M. L., Wu, Y.-C., Quenemoen, R. F., & Towles, E. (2016). *Characteristics of students with significant cognitive disabilities: Data from NCSC's 2015 assessment* (NCSC Brief #8). Minneapolis, MN: University of Minnesota, National Center and State Collaborative.

Thurlow, M., L., & Lazarus, S. (2017). *Strategies for meeting the 1% state-level cap on participation in the alternate assessment* (NCEO Brief #12). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

Thurlow, M. L., Wiley, H. I., & Bielinski, J. (2002). *Biennial performance reports: 2000-2001 state assessment data.* Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

Thurlow, M. L., & Wu, Y.-C. (2018). *2015-2016 APR snapshot #18: AA-AAS participation and performance.* Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.