

September 20, 2017

The Honorable Betsy DeVos
U.S. Department of Education
400 Maryland Avenue, SW
Washington, DC 20202

Re: Docket ID–2017–OS–0074, Evaluation of Existing Regulations for Repeal, Replacement, or Modification to Improve the Efficiency and Effectiveness of the Department

Secretary DeVos:

I write in response to the request for comments published in the June 22, 2017, Federal Register regarding the Administration’s review of existing regulations and guidelines issued by the U.S. Department of Education (USED) for possible repeal, replacement, or modification.

USED plays a vital role in administering federal laws in education to support states and districts in improving educational opportunity and outcomes for all children, youth, and adults, and particularly to promote equity and protect the civil rights of all students. USED regulations and guidance are often critical to this work – operationalizing federal statutes in a manner that is not clearly-established in law, is poorly understood in the field, and supports effective implementation. USED regulations can promote certainty and stability in federal law – for states, districts, students, and families – through processes that ensure public engagement and transparency. USED guidance can provide clarity without diminishing state or local authority. Together, these vital functions are often critical to establishing the systems within which states and districts operate, and on which individual students depend for the freedom to learn.

The Department can play an effective role in supporting states in their efforts to successfully meet the needs of each student as they implement the Every Student Succeeds Act (ESSA). The following describes several areas where additional clarification, guidance, or modification by the Department would be useful to states in these ongoing implementation efforts, as well as areas where ongoing regulatory attention will support equity in the provision of education.

I. Maintain and continue support for regulations and related technical assistance offered by the Office for Civil Rights (OCR) aimed at ensuring nondiscriminatory action by recipients of federal funds, the provision of due process, and the effective implementation of the biannual Civil Rights Data Collection.

OCR’s regulations and related technical assistance functions are intended to help ensure the protection of student civil rights such that no person faces discrimination under any program or activity receiving federal assistance from the Department because of race, color, national origin, sex, sexual orientation, or disability status. These efforts are further aided by the existing related regulations that serve to provide clarity and consistency in the implementation and enforcement of these civil rights protections. In addition to its critical role in conducting investigations and enforcing civil rights for students, OCR provides much-needed technical assistance and resources for addressing discriminatory practices to schools, districts, and states.

Specifically, with respect to USED’s advancement and enforcement of civil rights in education regulations, guidance, TA and enforcement from the Department all play a critical role in both the

development and implementation of state and local education policies and in the everyday protection of children and students when state or local policies or practices fall short of their civil rights obligations. Federal regulations and guidance are critical to this work in several ways and should be protected and maintained in the current review process. This specifically includes the following longstanding, essential provisions:

- Regulations under Title VI of the Civil Rights Act bar recipients of federal aid from using criteria or practices that may not on their face treat individuals differently based on race, color, or national origin, but that have the effect of subjecting individuals to discrimination based on these factors. (See 34 CFR 100.3(b)(2)&(3)). Title IX regulations and regulations based on Section 504 of the Rehabilitation Act include parallel provisions prohibiting these actions based on gender and disability, respectively. There is no private right of action to challenge these criteria or methods of administration in court. Only USED's Office for Civil Rights (OCR), (and the US Department of Justice) is able to enforce against these forms of discrimination. If the Administration were to revoke or dilute these regulations, it would undermine fundamental civil rights protections regarding equal opportunity that have been embedded in the law for many decades. I note also that it is not necessarily the case that a criterion or practice that has a discriminatory effect is illegal under these regulations. Rather, the effects result in a fuller analysis of whether the criteria or practices can be educationally justified and, if so, whether there are equally effective criteria or practices that do not have that effect.
- USED's civil rights regulations play an essential role in enforcing federal civil rights laws and require action whenever those rights are violated. They do not require "standing" for a complaint to be filed with OCR. The regulations provide that any person who believes himself or any class of individuals to be subjected to prohibited discrimination may file a complaint with OCR and that OCR will investigate whenever a complaint, compliance review, report, or other information indicates a possible failure to comply federal civil rights laws. (34 CFR 100.7(b)&(c)). This longstanding policy provides a fulsome mechanism for issues of discrimination to be raised with and reviewed by USED and OCR, and it is important both to address individual incidents and identify more systemic problems. OCR still has multiple ways to dismiss cases that do not warrant its review. But diminishing this longstanding access to civil rights review and protections would significantly curtail OCR's authority and opportunity to identify and appropriately address issues of discrimination that deny educational opportunities for historically underserved children and students.
- Since 1968, more than a decade before USED was even created, USED and its predecessor (the Department of Health, Education, & Welfare (HEW)) has conducted the Civil Rights Data Collection (CRDC) to collect data on key education and civil rights issues in our nation's public schools. This collection of data is clearly authorized by 34 CFR 100.6, which applies to Title IX and Section 504 regulations as well as Title VI regulations. The information collected has been vital in setting priorities for OCR review and enforcement of the non-discrimination laws, and also for identifying specific issues and, on occasion, specific institutions that warrant compliance reviews. The data also plays an increasingly important role for research and action in the field to understand and advance equity in both education opportunity and outcomes. While the CRDC is undergoing a separate review, I want to make clear that curtailing this longstanding data collection (and its longitudinal data) would hamstring OCR in carrying out its functions and more so undercut transparency and impact in the field.

CRDC reporting also reveals disparities in historically underserved students' access to equitable and inclusive disciplinary practices, rigorous coursework, and effective teaching. These data are

vital to educators, researchers, and advocates as they work with schools and communities to close these gaps and eliminate discriminatory practices. The Department should continue to maintain and support OCR and the important data collection, technical assistance, and civil rights protection that the office provides.

I urge USED to maintain the CRDC, its critical longitudinal data on access to opportunity, its dedicated funding stream separate from the larger OCR budget and its universal data collection across all public schools.

II. Modify existing assessment regulations or issue accompanying guidance in the following areas related to assessments administered under ESSA, as well as under the Innovative Assessment and Accountability Authority.

High-quality assessments are a critical element in helping educators and students master deeper learning competencies, such as critical thinking, complex problem solving, and effective communication. States should be encouraged to pursue responsible innovation in the design and implementation of assessments and the means for determining student performance. In particular, the Innovative Assessment and Accountability Authority authorized by Congress is critically important to enabling states to improve the quality of assessments, as other countries have already done,¹ so that they assess 21st-century skills and leverage these deeper learning competencies our students need to succeed in today's knowledge-based economy.

ESSA explicitly calls for states to use “multiple up-to-date measures of student academic achievement, including measures that assess higher order thinking skills and understanding, which may include measures of student academic growth and may be partially delivered in the form of portfolios, projects, or extended performance tasks.” For this direction to be taken up by states, Departmental regulations will be needed that do not unintentionally stifle opportunities to make advances in this area. As a result, I recommend that the Department:

1. Modify the existing assessment regulations or issue guidance to states regarding Sec. 1111(b)(2)(B) of ESSA to provide additional support and encouragement to use performance assessment methods.

For example, any modification to the regulations or subsequent guidance should clarify for states that they can use assessment funds under the law to support the development and use of portfolios, projects, or performance assessments at the state and local level. In addition, non-regulatory guidance to states could provide states with examples of how funds from Title I and Title II can be used to provide professional learning opportunities to educators to develop and score such assessments, as well as to develop instructional strategies designed to help students acquire the higher-level skills the assessments call for.

2. Clarify through modification of the existing assessment regulation or subsequent guidance that rules for defining the “same” assessments should focus on measurement of common standards, not use of identical items.

The Department's existing regulations would benefit from further clarification as to how states may implement the statutory requirement that states must both use "multiple up-to-date measures of student academic achievement, including measures that assess higher order thinking skills and understanding, which may include measures of student academic growth and may be partially delivered in the form of portfolios, projects, or extended performance tasks" and "the same academic assessments to measure the achievement of all public elementary school and secondary students in the State." Portfolios, projects, performance tasks, and adaptive assessments, by definition, include different assessment items; therefore, ESSA regulations or guidance must clarify how these forms of assessment may be implemented while meeting the statutory requirement for the implementation of the "same" assessments to measure all students in the state.

As is true for standardized tests that use multiple forms and for adaptive tests, I suggest that, with respect to performance assessments, the assessment regulation and any associated guidance should define the "same" assessments as those that measure the same well-defined standards using clear specifications for the student response that can be evaluated using common criteria to determine levels of performance. Taking this approach would allow for a range of performance tasks, just as standardized tests have different items and forms while still complying with the requirements under ESSA.

3. Clarify through modification of the existing assessment regulation or subsequent guidance that any requirements in Innovative Assessment and Accountability Authority for establishing comparability of results among new, innovative assessments and existing assessments should not constrain the quality of new assessments by the limitations of existing tests.

Under the Innovative Assessment and Accountability Authority, ESSA requires that states establish comparability of results between their new assessments and existing assessments. This determination of comparability should be done in a manner that does not constrain the quality of new assessments within the boundaries of existing tests. Since innovative performance assessments are typically designed to measure standards that are not fully measured on traditional state tests (for example, English Language Arts standards regarding extended research; written, oral, and graphical communication; debate; or mathematics standards regarding applications of mathematical concepts and practices), comparability cannot appropriately be established by looking for identical outcomes across the two sets of assessments.

Furthermore, many states have implemented tests that were developed prior to the adoption of more rigorous standards and/or that do not measure the full depth and breadth of the current standards. According to a RAND Corporation study examining the data available from 17 states, fewer than 2% of mathematics items and only 21% of English language arts items required higher-level processing and complex analyses.ⁱⁱ In addition, only 3% to 10% of k-12 students were assessed using extended activities that called for complex analyses and the ability to synthesize complex ideas.ⁱⁱⁱ

To require new assessments to produce the same results as existing assessments measuring different skills in different ways would limit their ability to assess the higher-order skills they are designed to evaluate. An analogy would be if a hospital had been using a thermometer to derive

scores for patients' health, but then acquired other tools to measure a broader set of health factors, such as blood pressure monitors, EKGs, or blood tests. It would not be reasonable to expect a patient's health score to be identical to the old thermometer score when these new and different kinds of information are added.

Instead, regulations and guidance should allow comparability to be established in terms of:

- comparable measurement of the standards, ensuring that assessments measure well-defined standards validly and in comparable ways across tasks as informed by a peer review process;
- comparable scoring across responses, using common criteria to determine levels of performance and systems of moderation to ensure consistency in scoring; and
- comparable results across individuals, schools, and districts, through analysis of student results on common items or tasks; for example, through continued use of state assessments in grade spans, use of embedded items from state assessments, or embedding innovative tasks into the pilot section of the state assessment.

4. Clarify through modification of the existing assessment regulation or subsequent guidance that rules for assessing student performance should allow for competency-based systems.

ESSA ostensibly allows for the implementation of competency-based systems of assessment both under its general grants for assessments and under the Innovative Assessment and Accountability Authority. The commitment to a competency-based approach in a growing number of states and communities requires that students be taught and assessed at whatever the appropriate level along a continuum of learning for a given domain of knowledge and/or set of skills. The student's learning and readiness is the driver for the assessment, including when the assessment is provided, rather than the designation of student grade level or a fixed point in time.

ESSA recognizes that single-grade-level end-of-year tests, common under NCLB, offer too little information for many students and do not permit the evolution of competency-based approaches. Section 1111(b)(2)(J) of ESSA specifically allows for the measurement of a student's level of academic proficiency and growth using items above or below the student's grade level, in addition to grade-level proficiency, including for use in the state's accountability system, specifically in the context of computer adaptive assessments. This type of out-of-grade-level assessment should not be limited to computer adaptive assessments because it is valuable to the development of innovative assessment systems permitted under the demonstration authority. In addition to whether a student is performing on grade level, information as to whether a student is performing above or below grade level, and by how much, provides useful data that can drive the appropriate level of instructional rigor and support.

For example, many performance tasks are designed to allow multiple entry points and expansive responses that ultimately provide data about student learning in ways that are not bounded by grade-level standards. These tasks should be able to be scored in ways that attend to the actual skills demonstrated, rather than in ways that are artificially bounded by grade-level standards. Clarifying that innovative assessment systems may include items and tasks that provide information above or below grade level is inherent to the concept of competency-based education. Such clarification is also aligned with the goals of the demonstration authority to develop assessment systems that are student centered and emphasize growth to and beyond

proficiency.

State efforts to move toward these types of comprehensive assessment systems should be supported. The Department should provide clarification to §200.2(b) of the assessment regulations through either modification or accompanying guidance that states implementing innovative assessment systems may, as with computer adaptive assessments, “measure the student’s level of academic proficiency and growth using items above or below the student’s grade level, including for use as part of a State’s accountability system” (see ESSA Sec. 1111(b)(2)(J)(i)(bb)).

5. Clarify through modification of the existing assessment regulation or subsequent guidance that effective use of assessment audits should exclude performance assessments from calculations of testing time.

ESSA includes two provisions regarding time spent on test administration: (1) under the Parents Right-to-Know provision, states must inform parents, where the information is made available, with “the amount of time students will spend taking the assessment” and (2) under the provision providing funding for assessment audits, information for stakeholders must include the amount of time teachers spend on administering assessments. Both provisions may incentivize states to substitute lower-quality assessments that take less time to administer in place of higher-quality assessments that may take longer to administer, such as extended-performance tasks, project-based learning, and portfolios.

Guidance related to the assessment regulations should clarify that calculations of time spent on testing should exempt the time spent on in-depth performance tasks and portfolios, which are typically embedded in the curriculum and part of the instructional process. Including this clarification in the regulations recognizes the distinction between these types of activities and will serve to incentivize state adoption of higher-quality assessments.

III. Provide clarification to states and districts regarding the following issue areas addressed in the ESSA accountability regulations that were repealed. Absent these regulations, it would be helpful for the Department to provide guidance and make determinations of ESSA plan approval that derive directly from the letter of the law and are consistently applied in the way they interpret the law.

1. Clarify to states that ESSA does not require the use of a single summative score for accountability purposes, as well as that a data dashboard approach, along with a set of decision rules for identifying schools for intervention, is permissible.

Because of the language included in the previous Administration’s proposed and final accountability regulations, there is confusion among some states as to whether a single summative rating or score is required, despite the fact that this requirement is not included in ESSA.

ESSA requires that states identify at least 5% of their Title I schools for comprehensive assistance based on its new accountability systems, which will include multiple measures, such as literacy and math achievement, English proficiency gains, graduation rates, and other

indicators. The law does not prescribe a particular method for this identification, aside from noting that the four academic measures specified must each carry “substantial weight,” and the combined set must carry “much greater weight” than other measures the states add. In lieu of an index or single score, some states have developed decision rules that allow them to identify at least 5% of schools for intervention, while ensuring that the required four academic measures carry much greater weight than other indicators in the system. This should be permitted under ESSA, which does not proscribe such an approach.

Single-dimensional approaches to school ranking such as letter grades or an accountability index can lose the value that a multidimensional system can offer.^{iv} A single index masks the rich information that the multiple indicators of student progress provide; thus, it does not provide teachers and school systems the data needed to inform the basis of strategies for improvement.

Some states that previously tried a single-score grading or ranking system found this practice problematic because it failed to focus schools on specific areas for growth. This was especially true if schools ranked above a minimum cut point, hiding important information that should be informing improvement efforts. Under a single summative scoring system, some schools that are doing poorly in one area but reasonably well in others may not be identified for interventions and support. Further, when multiple indicators are aggregated together to yield a summative score, student subgroup performance can also be hidden from view. For example, as a study done by the Alliance for Excellent Education found, in one state with an A-F system, the average proficiency rate for African American students in schools that received an “A” rating was only 58%.^v In another state, 183 high schools received the highest rating within the state accountability system while having at least one student subgroup with a graduation rate in the 60s or below.^{vi}

Other methods for implementing a multiple-measure accountability system, such as data dashboards, can promote transparency, support the continuous improvement of all schools, and allow schools to more effectively measure the deeper learning skills and competencies students need to be successful in postsecondary education and the workforce. Under ESSA, many states are well along a path toward developing new accountability systems relying on a dashboard of multiple indicators that provide more comprehensive information for school identification, intervention, and improvement. California, New Hampshire, New York, Oregon, Vermont, Virginia, and West Virginia are among the states choosing to use a dashboard approach rather than a single summative score for school identification and continuous improvement. They have developed identification processes that meet the letter of the law in appropriate ways while preserving the benefits of a multiple measures system. The Department should support these states in their efforts and continue to clarify to other states that these approaches are allowable under ESSA.

- 2. Clarify that states may include the use of scale scores or index scores that measure performance across the achievement continuum in their reporting on student performance, in addition to the percentage of students who have met a single cut point.***

ESSA requires that states measure at a minimum each student’s academic proficiency based on the challenging state academic standards for the student’s grade level and growth toward such standards. The law requires states to use an indicator of academic achievement “as measured by

proficiency on the annual assessments in mathematics and reading or language arts.”^{vii} ESSA further requires states to establish long-term goals for all students and separately for each subgroup of students in academic achievement, as measured by proficiency on the annual assessments, and to report this information.

ESSA does not, however, require a particular method of tracking students’ proficiency levels. Research demonstrates that a focus on the percentage of students who reach a particular cut point or proficiency standard incentivizes schools to focus only on a selected few students hovering around the proficiency cut point, rather than paying attention to all students at all levels of achievement. Measures that rely upon moving a select and small group of students across a threshold—for example, from “Basic” to “Proficient”—create the incentive to overdirect attention and resources to students on the bubble at the expense of others.^{viii}

Further, studies during the NCLB era characterized this well-documented practice as “educational triage,” which resulted in focusing especially on “students near proficiency and emphasizing test-specific rather than generalizable skills.”^{ix} Research demonstrates that the improvement gap during this era was largest in the low-achieving schools in which focusing on students near the proficiency cut score came at the expense of attention to the lowest-achieving students.^x The sole use of “percent proficient” fails to make distinctions among students or schools that are further away from or closer to the cut points, and those that have made significant progress or have largely stagnated in their progress.

Many states are moving toward using scale scores or movement across performance levels in their reporting and accountability systems to demonstrate student and subgroup growth across the entire achievement continuum,^{xi} no longer using “percent proficient” at a single cut score to describe achievement levels. They are including a focus on both student-level achievement and growth over static measures for the well-understood reason that all students arrive at school with varying levels of preparedness and that schools should be recognized for having increased student learning.

This clarification by the Department was requested by a group of assessment experts in [a letter](#) sent to the previous administration asking the Department to clarify that the use of average scale scores is permissible under ESSA. The Department should support the use of significantly more informative ways to report performance and growth, including progress along the entire scale used to reflect scores. For example, such reporting can reveal that overall students moved, on average, from a score of 234 to 250, while English learners moved from 208 to 240, a rate of improvement twice as great. All of these changes could occur without being evident on a measure of the “percent proficient,” masking information about growth and progress.

Similarly, a school that is 1% below an arbitrary target is not substantially different from a school that is 1% above the same target. However, a school that falls 1% below the target is likely substantially different than one that falls 30% below the target, yet both would be treated in the same way under a “percent proficient” reporting system. As opposed to cut scores, using scaled scores can help reveal actual performance and how far students have progressed toward proficiency.

The Department should make clear that states are allowed to report achievement in more productive ways by encouraging methods—such as the average scale score or movement across categories representing the achievement continuum—that provide increased accuracy and usefulness with information that shows status, progress, and improvement across the full range of proficiency levels.

3. *Clarify that states may include criteria beyond the English language proficiency assessment to determine language proficiency and reclassification.*

It is unclear due to the repeal of the accountability regulations whether states may include criteria beyond the English language proficiency assessment to determine language proficiency. Either through guidance related to accountability or through modification to or guidance for the assessment regulations, the Department has an opportunity to provide greater clarification.

Individual student performance on the state English language proficiency (ELP) assessment is used by many districts and schools to determine readiness for English learners to be reclassified and exit the subgroup. In fact, according to a working paper from the ELL Working Group and the Council of Chief State School Officers, 29 of the 50 states and the District of Columbia exclusively use the ELP to determine a student's reclassification to former English learner status.^{xii} However, because tests are necessarily a limited measure, a number of the states use additional criteria to assist with reclassification. For example, Montana, New Jersey, New York, Pennsylvania, and Utah consider teacher input and evaluation in addition to the ELP, while California, Michigan, Rhode Island, and Wisconsin use individual performance on the academic content assessment in addition to the ELP.

States should be supported by the Department in using criteria in addition to the ELP to determine language proficiency, growth, and reclassification. It is important that reclassification rate measures be carefully designed; otherwise, incentives to reclassify could overestimate student capacity to meet all the challenges they may face in academic settings, while leaving them without needed services for academic success.

Assessing language proficiency is a complex undertaking, as students must demonstrate proficiency in written formats and orally, as well as in multiple subjects such as mathematics, social studies, and science, which often require discipline-specific language. Supporting states in using multiple means of assessing proficiency that include using the ELP ensures comparability, while also acknowledging the importance of educator input and performance on state content assessments. Since the ultimate reflection of proficiency in English in the real world is the capacity to learn and demonstrate learning in English, performance on content assessments provides valuable information. The Department could clarify through guidance that states may include measures beyond the ELP to develop this indicator.

These and other key USED regulations and guidance related to civil rights in education should not be undercut. This does not mean that regulations or guidance should not be revised and continuously improved, particularly if they are outdated, but this should not taken longstanding, critical guidance and regulations that undergird our nation's education system and are crucial to USED's role and student rights. I urge you to maintain and affirm these critical items.

If any additional information would be useful as you consider issuing or drafting modifications to existing regulations or accompanying guidance, please let me know. I look forward to the shared work ahead and the opportunity to support states in the successful implementation of ESSA for all students.

Sincerely,
Beth Glenn
Education Policy Consultant

ⁱ Darling-Hammond, L., & Adamson, F. (2015). *Beyond the Bubble Test: How Performance Assessments Support 21st Century Learning*. San Francisco: Jossey Bass.

ⁱⁱ Alliance for Excellent Education analysis of Tables 4.1, 4.2, and 4.3 in Yuan, K., & Le, V. (2012). *Estimating the percentage of students who were tested on cognitively demanding items through the state achievement tests*. Santa Monica, CA: RAND Corporation.

ⁱⁱⁱ Alliance for Excellent Education analysis of Tables 4.1, 4.2, and 4.3 in Yuan, K., & Le, V. (2012). *Estimating the percentage of students who were tested on cognitively demanding items through the state achievement tests*. Santa Monica, CA: RAND Corporation.

^{iv} Rothman, R. & Cardichon, J. (2015). *Data dashboards: Accounting for what matters* Washington, DC: Alliance for Excellent Education.

^v See Ushomirsky, N., Williams, D., & Hall, D. (2014). *Making sure all children matter: Getting school accountability signals right*. Washington, DC: The Education Trust. <https://edtrust.org/resource/making-sure-all-children-matter-getting-school-accountability-signals-right/>.

^{vi} Alliance for Excellent Education analysis of accountability data for Colorado.

^{vii} ESSA Section 1111(c)(4)(B)(i)(I).

^{viii} Jennings, J., & Sohn, H. (2014). Measure for measure: How proficiency-based accountability systems affect inequality in academic achievement. *Sociology of Education*, 87(2), 125–141.

^{ix} Jennings, J., & Sohn, H. (2014). Measure for measure: How proficiency-based accountability systems affect inequality in academic achievement. *Sociology of Education*, 87(2), 125–141.

^x Lauen, D. L., & Gaddis, S. M. (2015). Accountability pressure, academic standards, and educational triage. *Educational Evaluation and Policy Analysis* 38(1), 127–147; Sparks, S. D. (2012, March 9). Study finds “bubble student” triage a gut reaction to rising standards. *Education Week*. http://blogs.edweek.org/edweek/inside-school-research/2012/03/study_finds_bubble_student_tri.html.

^{xi} For example, states using the Smarter Balanced Assessments are measuring students with both scale scores and at least four levels, which each state can name differently.

^{xii} Liguanti, R., & Cook, H. G. (2015). *Re-examining reclassification: Guidance from a national working session on policies and practices for exiting students from English learner status*. Washington, DC: Council of Chief State School Officers.

http://www.ccsso.org/Documents/EL%20Reclassification%20Working%20Paper_11%2005%2015%20Final.pdf.