U.S. VIRGIN ISLANDS ACCOUNTABILITY SYSTEM:

Design Decisions

September 2017

Prepared by

Juan M. D'Brot, Ph.D. | Senior Associate

Center for Assessment

National Center for the Improvement
of Educational Assessment
Dover, New Hampshire

CENTER FOR ASSESSMENT

EXECUTIVE SUMMARY

With the passage of the *Every Student Succeeds Act* (ESSA), the U.S. Virgin Islands Department of Education (VIDE) took the opportunity to develop a new ESSA-aligned accountability system. The VIDE's accountability system embodies the goals and priorities of the territory and seeks to incentivize continuous improvement for all schools. To develop this accountability system, the VIDE partnered with the Center for Assessment to facilitate meetings that included representatives from the VIDE, the Florida and the Islands Comprehensive Center (FLICC), and the VIDE's Accountability Task Force. Over the course of meetings from 2016-2017 and 2017-2018 school years, the Task Force made recommendations to the VIDE with regard to goals, design principles, accountability indicators, and how to combine those indicators in an overall system. This document summarizes the discussions, considerations, and recommendations made by the Task Force. It concludes with a description of the next steps required for the accountability system, which includes setting performance standards and documenting decisions for an operational accountability system.

INTRODUCTION

The passage of the Every Student Succeeds Act (ESSA) affords states and territories an opportunity to reexamine their accountability systems. As part of this opportunity, states and territories can promote strategies that integrate graduation requirements, assessments, and other indicators of school quality and student success to help guide schools and districts in supporting student readiness for college and careers. The following report begins by outlining the goals, guiding principles, and theory of action of the U.S. Virgin Islands' (USVI) accountability system. It then continues to describe the recommendations and decisions made by the USVI's Accountability Task Force and the Virgin Islands Department of Education (VIDE).

KEY GOALS FOR USVI'S ACCOUNTABILITY SYSTEM

Explicitly outlining a theory of action is one of the first steps in designing a coherent and effective school accountability system. A theory of action—or theory of change—defines the mechanisms by which the accountability system will accomplish its goals and identifies the assumptions that must hold in order for change agents to function properly. Articulating the theory of action for the USVI ESSA Accountability system is a crucial initial step for ensuring that the system design is aligned with, and promotes the attainment of, the educational goals of the territory, local education agencies, and stakeholder groups. Additionally, by identifying the chain of logic that undergirds the accountability system, potential unintended negative consequences can be brought to light and mitigated. Building a theory of action begins with the identification of system goals, or the purposes of the accountability system. There are four overarching goals of the U.S. Virgin Islands ESSA School Accountability System:

1. Articulate clear expectations and improve teaching and student learning to ensure all students are ready for college and career.
2. Differentiate the performance of schools and districts in meaningful ways so that those in need of improvement receive appropriate supports and interventions, and those excelling can be recognized as models of excellence.
3. Provide transparent, timely reporting of actionable data on performance results so that stakeholders at all levels can take appropriate actions.
4. Foster a commitment to innovation and improvement in teaching and instruction.

THE USVI'S ACCOUNTABILITY SYSTEM TASK FORCE

The approach to designing the new accountability system is rooted in the belief that the school accountability system should reflect the Virgin Islands' educational values and priorities. In support of truly reflecting the context, values, and principles of the territory, the Virgin Islands Department of Education (VIDE) convened an Accountability Task Force to provide input into the system design. The Task Force was purposefully selected to represent a variety of perspectives

and stakeholders. The Task Force is made up of key educational leaders in the territory and includes representatives from the following groups:

- District leaders from the St. Thomas and St. Croix/St. John districts
- Teachers and principals from the St. Thomas and St. Croix/St. John districts
- Representatives from the St. Thomas and St. Croix/St. John Federation of Teachers
- Representatives from the St. Thomas and St. Croix/St. John Federation of School Administrators
- Key members of the VIDE with expertise in assessment, accountability, school improvement, and curriculum

Prior to ESSA's passage, the Task Force convened to specify expectations and accountability rules for schools under the *No Child Left Behind Act of 2001*. Shortly following ESSA's passage, the Task Force shifted its focus to establish a locally-defined accountability system that embodied the goals of the territory. The Task Force sought to accomplish this by reviewing the requirements of ESSA to better understand the intent of the law and how it could push their own goals forward. Using the law as a guide, the Task Force defined the design principles for the territory's accountability system. After defining the design principles, the Task Force then focused on the measures that would most accurately reflect the goals and design principles of the Islands while aligning with the indicators specified in ESSA.

Ultimately, the requirements and opportunities of ESSA were used as an anchor and guide for the Task Force's recommendations and VIDE's decisions. The recommendations and decisions made by the Task Force and VIDE are presented below and will form the basis of defining performance standards for the accountability system in October of 2017.

DESIGN PRINCIPLES FOR USVI'S ACCOUNTABILITY SYSTEM

1. As a result of the first meeting focusing on ESSA, five design principles for the new accountability system emerged, serving as the foundation for the new system's theory of action and its design: **Multiple measures within the accountability system provide a more complete picture of school quality than any single metric.** ESSA specifies a set of required indicators in addition to academic achievement, which includes student growth and English Language Proficiency. The Task Force used this principle to determine how to best include additional indicators of school quality and student success in the accountability system.

2. **The indicators in the system are reflective of practice and inform system-level decisions and school improvement efforts.** The Task Force believed the system should include indicators that are closer to what is occurring in everyday teaching and learning, but not lose sight of the need to have less subjective measures of district and school progress. In this way, the system can provide feedback to schools about performance trends and practice because the results will be meaningful for educators and students. Additionally, indicators that reflect practice will help identify best practices for school improvement across the state or territory. Thus, the accountability

system should report on those measures that are critical to improving student learning and should be tied to processes and expectations that focus on student and teacher behavior, school improvement practices, and strategic connections within the school-Local Educational Agency (LEA)-State Education Agency (SEA) partnership.

3. **The system supports the identification of low performing schools and helps determine what resources are needed to support effective school improvement practices.** In addition to promoting a meaningful identification process, the Task Force believed the accountability system should include mechanisms to encourage districts and schools to engage in self-evaluation to drive improvement efforts. School quality indicators signaling high expectations and engagement were identified and will need to continue to be linked to research-based practices and supports that facilitate school improvement. The accountability system is intended to be reciprocal in that the SEA should allocate resources and supports to help struggling schools meet performance expectations. In turn, schools are expected to use those resources and supports effectively to better prepare students to be college- and career-ready. This principle requires ongoing evaluation and revision of the accountability and improvement systems and their implementation.

4. **Results are timely, accessible, and actionable. Reporting should reflect both high-stakes indicators of the accountability system and low-stakes measures that signal more near-term progress.** The Task Force believed this will allow both the public and local educators to easily access up-to-date and clear information about school quality. The Task Force also believed results should be reported in a way that reflects information from the multiple indicators comprising the accountability system. By reporting both high-stakes and lower-stakes information, the system can help stakeholder understand the potential link between the two. This will be an ongoing focus of revisions to the accountability system.

5. **The accountability system should incentivize behaviors that include high-level policy decisions (e.g., resource allocation and professional development) and closer-to-the-classroom strategies to encourage more effective teaching and learning.** The Task Force believed high- and ground-level behaviors should be informed by the measures in the accountability system. Further, the expected strategies and practices associated with school identification should direct schools to suggested resources and supports. While the differentiation among school ratings reflects differences in lagging indicators (e.g., proficiency rates, rates of growth, graduation rates), the Task Force intended for school ratings to also specify what resources schools should receive. The differentiation of resources is based on whether schools are on the reported high-stakes indicators and is further informed by the improvement processes and expectations set forth by the system.

THEORY OF ACTION FOR USVI'S ACCOUNTABILITY SYSTEM

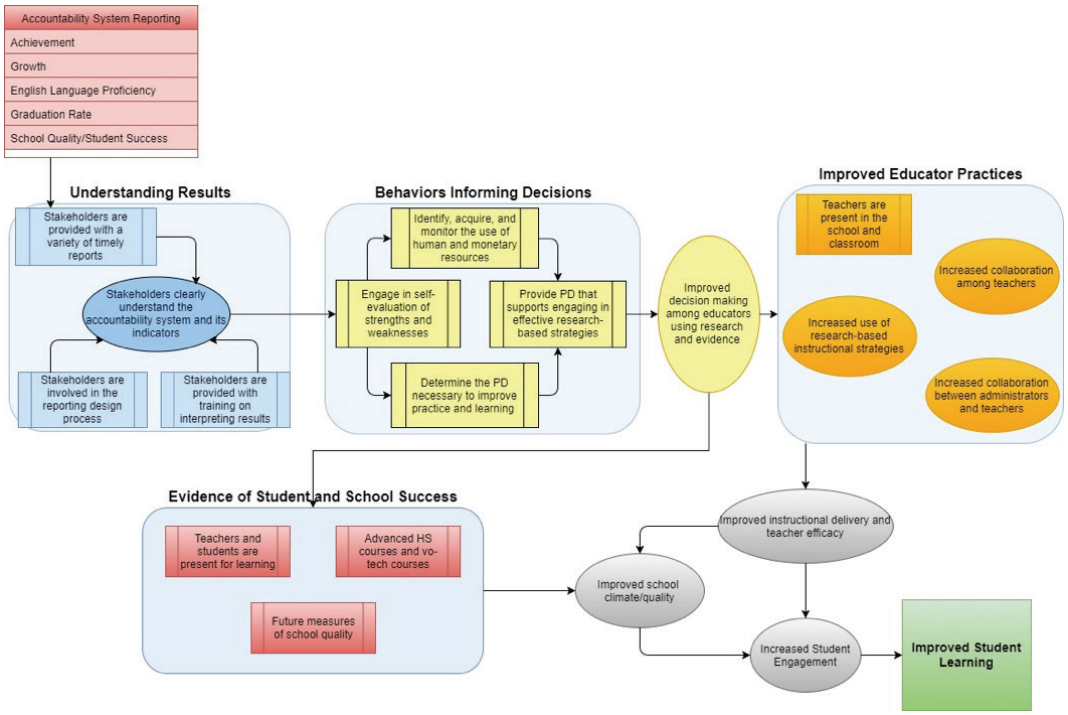
A theory of action should reflect the goals and design principles for a state or territory and its agency. The USVI Accountability Task force distilled the four goals into the following key purposes:

1. Improve student learning (Goal 1)
2. Stakeholders understand results (Goal 3)
3. Improve programs and professional development to support better teaching and learning (Goals 2 and 4)

These key goals and purposes are reflected in the theory of action for the Virgin Islands Accountability System below. The overall intended outcome is identified by the green rectangle. The ellipses represent the intermediate mechanisms by which the goals of the system will be accomplished. The boxes represent the processes/actions necessary to support those necessary mechanisms. The large background rectangles represent grouped clusters of processes, actions, and/or mechanisms. The arrows represent the flow of system information and the assumptions that must hold for the intermediate and long-term outcomes to be realized. Lastly, the chart list in the upper left represents system inputs, or in other words, accountability indicators.

This diagram is purposefully broad to allow for ongoing specification and refinement of the indicators, change agents, and goals by the Accountability Task Force and the VIDE as necessary. Additionally, the Theory of Action is intended help the VIDE check assumptions and for potentially unintended negative consequences as more detailed design decisions are made and adjusted. The Task Force and VIDE recognized that each component of the Theory of Action can become an independent source of evidence that may require evaluation and monitoring. The Task Force recommendations for the accountability system are discussed following the Theory of Action.

FIGURE 1. U.S. VIRGIN ISLANDS THEORY OF ACTION



ESSA ACCOUNTABILITY TASK FORCE RECOMMENDATIONS

A key aspect of the Task Force's work was transforming the design principles into a framework for accountability using ESSA as a guide. The task force considered several possible measures for each indicator. The Task Force also recognized that each indicator is subject to constraints that determined how recommendations could be applied. These constraints included, but were not limited to, data entry, data accuracy, data volatility over time, variability, corruptibility, data availability over time, data availability across schools, and relevant funding. By considering these constraints, the Task Force proactively avoided some of the pitfalls associated with selecting measures that may be inappropriate for high-stakes use.

The Task Force recommendations for each indicator of the school accountability system are described below. These indicators include: (1) academic achievement, (2) student progress, (3) graduation rate, (4) English language proficiency progress, and (5) school quality and student success. Following the description of the indicators, the recommendations for aggregation and reporting are described.

USVI Indicators based on ESSA Requirements

Academic Achievement. Academic achievement can be referred to as a status or point-in-time indicator. It has historically been operationalized as either the percent of students who are proficient on the statewide assessment or based on an index that awards points to a school based on the number/percent of students who score at various achievement levels. Under NCLB, achievement was reported as the percentage of students scoring at the proficient level or higher. Percent above the cut (e.g., proficient) has been criticized for many measurement (e.g., reduction of information) and consequential (e.g., focusing on "bubble kids") reasons, but it has the advantage of familiarity and relative ease of understanding. While states and territories are still required to report percent proficient, ESSA may allow for approaches that rely on information gained from the full achievement distribution such as an index system.

The Task Force considered several possible approaches for including academic achievement in the accountability system. They began by exploring two complementary approaches with different signals. The first was the percentage of students scoring proficient and above, which prioritized ease of use and familiarity. The second was mean (average) scale score, which prioritized detail, variability (i.e., the degree to which differences could be detected by group), and specificity. Given the bluntness of percent proficient, the Task Force initially recommended using average scale scores for the achievement indicator. Upon reviewing the now retracted regulations for ESSA that disallowed the use of average scale scores, the Task Force questioned the complexity of this approach. Ultimately, they sought to find a middle ground that would retain some familiarity to the public but would provide more detail than proficiency.

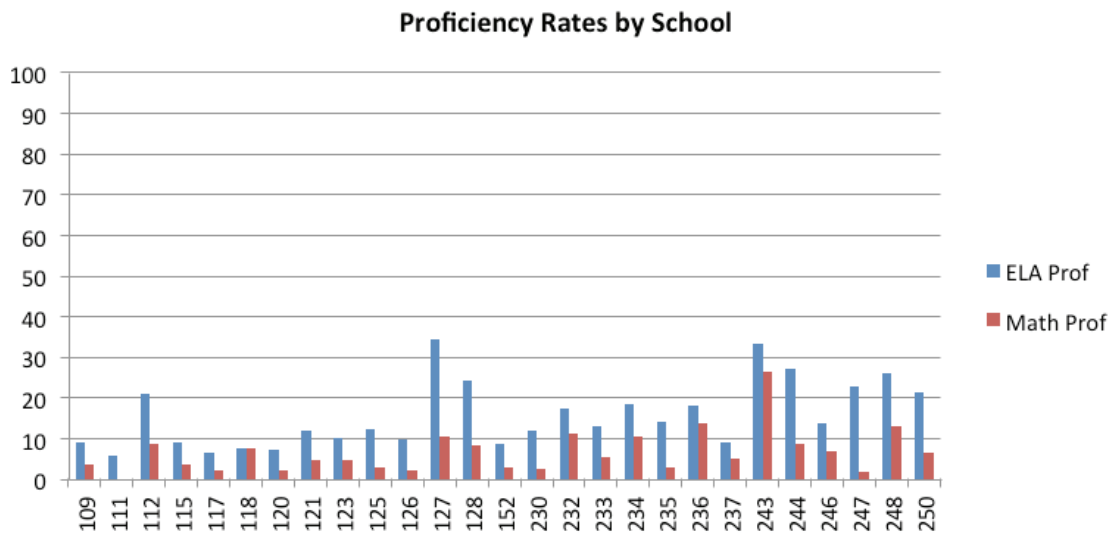
The Task Force's final recommendation for the achievement indicator was for the VIDE to use an index-style approach that awarded points to schools based on the percent of students scoring at each performance level. Depending on the percent of students in a school in a given performance level, the school would then be awarded a proportion of the possible points. The point spread for student performance is displayed in the table below.

TABLE 1. ACHIEVEMENT INDEX POINTS

ACHIEVEMENT LEVEL	POINT VALUE
Level 1	0
Level 2	0.5
Level 3	1
Level 4	1.25

The table above reflects the Task Force's desire to award points relative to how a student's performance compares to the proficiency criterion. Further, the Task Force sought to award extra points for those students who exceeded proficiency, but not so many that it would fully compensate for students who have yet to reach proficiency. This also reflects the reality of the Virgin Island's current performance. Using the most recent validated assessment data, the percent of students who are proficient is both comparatively and objectively low. These rates can be seen in the figure below.

FIGURE 2. PROFICIENCY RATES FOR THE U.S. VIRGIN ISLANDS *(2015-2016 school year)*



As seen in the figure above, the blue (i.e., left of pair) columns reflect proficiency rates in English/language arts (ELA) and the red (i.e., right of pair) columns reflect math proficiency rates for each school in the Virgin Islands. It may be difficult to interpret the degree to which students are performing on the statewide assessment using only proficiency rates. That is due to the lack of sensitivity associated with percent proficient and the generally low performance on the statewide assessment. While the statewide assessment highlights student progress toward post-secondary readiness, it reduces comparisons to a yes/no indicator.

In an attempt to communicate more detail and to prioritize proficiency against grade-level expectations, the Task Force recommended using the percent of students who are in each performance level. While there are still many students who are not yet at Level 3 or above (i.e., the proficient target), highlighting the percent of students who are in Levels 1 and 2 begins to show a different picture with regard to performance. The percent of students in each performance level for ELA and math are shown in the two figures below.

FIGURE 3. PERCENT OF STUDENTS IN EACH PERFORMANCE LEVEL IN ELA
(2015-2016 school year)

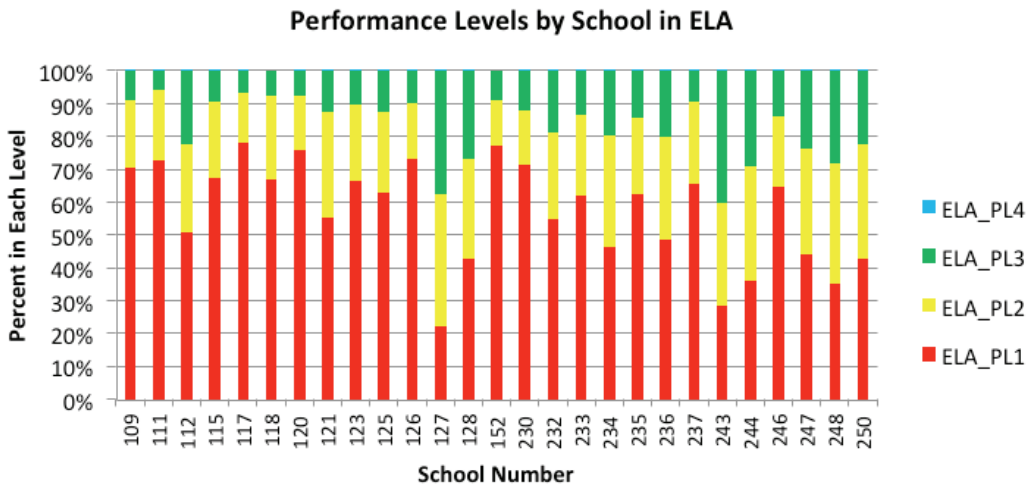
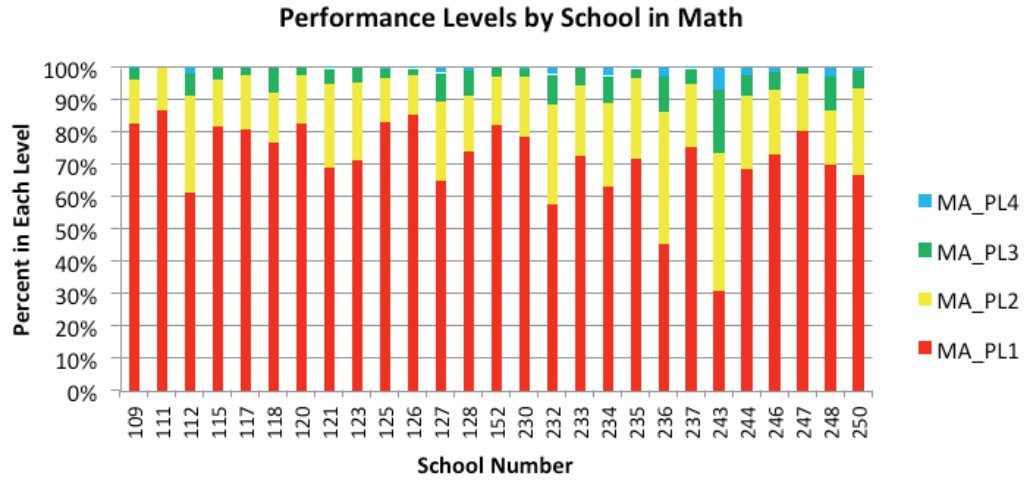


FIGURE 4. PERCENT OF STUDENTS IN EACH PERFORMANCE LEVEL IN MATH
(2015-2016 school year)



Using the number of students in each performance level, schools are then awarded points based on the point table shown earlier in this section. The points and calculation is displayed in the table below. These points are applied to the overall accountability calculation.

TABLE 2. ACHIEVEMENT LEVEL POINTS AND CALCULATION

ACHIEVEMENT LEVEL	POINT VALUE	USE
Level 1	0	$\frac{\text{(Number in Category (by school) x Point Value)}}{\text{Number of students in the school}} \times 100$
Level 2	0.5	
Level 3	1	
Level 4	1.25	
*Note: Highest level cannot compensate for lower level performance		

Student Progress. According to ESSA, another valid and reliable academic indicator must be included in the accountability system. The law offers student growth and achievement gap closure as two examples, but it is not limited to those examples. That said, measuring achievement gaps is one of the trickiest things to do well in educational measurement. Simple approaches such as computing the differences in percent proficient are almost always wrong, while more technically correct approaches such as computing the area between two performance distributions or effect sizes are a bit more challenging to explain. As difficult as it is to measure achievement gaps at any point in time, the challenges associated with measuring changes in achievement gaps are enormous. On the other hand, there are well-established methods for documenting student growth such as student growth percentiles (SGP) or mixed-effects models. However, these approaches are also subject to the same tensions between computational accuracy and ease of explanation.

Before ESSA was signed into law the Task Force and the VIDE were interested in using a student progress measure and thus made this aspect of the accountability system a focus of conversation and consideration. Several approaches for calculating student growth were proposed to the Task Force. Each approach was described based on the following characteristics:

- The interpretation supported by the model;
- The granularity/coarseness of the measure used for and results of the model;
- The degree to which status and are growth related (i.e., correlated);
- The level of mismatch between the face-value concept of growth (e.g., student progress on the standards) and the mechanics of a growth model;

- The challenges, if applicable, with aggregating growth estimates across subjects and scales ;
- The calculation and communication complexity of the model; and
- The reliability or instability of the measure used for and results of the model.

The Task Force recommended that they focus on two growth approaches to explore more carefully—student growth percentiles (SGPs) and transition tables. SGPs are conditional growth percentiles that communicate how much growth has (or should have) occurred given a student's starting point and given the performance of other students like them. For example, lower performing students (e.g., low Level 1 students) will likely demonstrate large scale score increases from year to year. However, a larger than average scale score increase in comparison to other students in their academic peer group would result in high growth. A smaller than average scale score increase in comparison to other students in their peer group would result in low growth.

Transition tables attempt to detect progress by tracking changes in performance levels from year to year. This is most commonly done by dividing a performance level into two or three sub-levels. For example, Level 1 would be divided into Levels 1a, 1b, and 1c. The degree to which a performance level can be divided is typically based on the amount of measurement error in a test, or the confidence in the consistency of a student's score if he or she were to take the test again.

The Task Force wrestled with the message and philosophy that would help determine which growth approach to use while also managing the tension between computational accuracy and ease of explanation. Over a series of four separate meetings, members actively discussed and debated the signals sent by student progress measures. While SGPs were less related to achievement and served to better (and more precisely) describe student progress, the Task Force was concerned about the low performance of the territory and how it interacted with a norm-referenced growth model[1]. That is, because progress might flat-line, very small improvements would be interpreted as high growth. This would send unintended signals of high performance in the accountability system that could result in diminished urgency of school improvement efforts. While SGPs address this concern by calculating the necessary amount of growth required to reach proficiency, the Task Force did not believe their constituents would be ready to interpret and use both the norm- and criterion-based growth estimates.

Transition tables are more related (i.e., correlated) to proficiency rates and are based on changes in performance levels over two years. This makes them very straightforward and easy to understand, but less sensitive to progress and perhaps duplicative of information provided by proficiency rates. However, the Task Force was intentional about making signals across indicators consistent and wanted to incentivize growth to proficiency above any other kind of progress over time. Thus, the Task Force's recommendation for student progress was to use transition tables for the statewide assessment. The tables below describe the performance sub-levels and how schools are awarded points.

[1] It is important to note that many growth models are normative and normative models tend to be less related to achievement—a criterion referenced point-in-time indicator.

TABLE 3. TRANSITION TABLE SUB-LEVELS AND GROWTH CATEGORIES

PRIOR GRADE		CURRENT GRADE							
		Level 1		Level 2		Level 3		Level 4	
		Low	High	Low	High	Low	High	Low	High
Level 1	Low	M	I	SI	VSI	VSI	VSI	VSI	VSI
	High	D	M	I	SI	VSI	VSI	VSI	VSI
Level 2	Low	SD	D	M	I	SI	VSI	VSI	VSI
	High	VSD	SD	D	M	I	SI	VSI	VSI
Level 3	Low	VSD	VSD	SD	D	M	I	SI	VSI
	High	VSD	VSD	VSD	SD	D	M	I	SI
Level 4	Low	VSD	VSD	VSD	VSD	SD	D	M	I
	High	VSD	VSD	VSD	VSD	VSD	SD	D	M

Based on the table above, each of the performance levels is divided into two levels, resulting in eight total sub-levels used for tracking student progress across two years. Depending on the number of sub-levels a student progresses, they are identified as one of seven growth categories. Each one of these categories corresponds to a point value that is used to award points to a school for growth indicator. These categories and point values are presented in the table below, which the VIDE will evaluate once data are available.

TABLE 4. TRANSITION TABLE GROWTH CATEGORIES AND POINT VALUES

CATEGORY	POINT VALUE	USE
Very Significant Improvement	1.50	$\frac{\text{(Number in Category (by school) x Point Value)}}{\text{Number of students in the school}} \times 100$
Significant Improvement	1.00	
Improvement	.50	
Maintain	0	
Decline	-.50	
Significant Decline	-1.00	
Very Significant Decline	-1.50	

Graduation Rate. Under NCLB and now under ESSA, graduation rate must be part of the accountability system for high schools. While historically allowed, ESSA specifies that an extended graduation rate can be used at the SEA's discretion in addition to the required four-year adjusted cohort graduation rate (ACGR). States and territories have been using the ACGR since approximately 2010 (based on 2008 regulations) to support comparisons across the U.S. and its territories. The ACGR is based on the percent of students who enter grade 9 in a given year (i.e., the cohort year) and graduate with a regular high school diploma. This cohort can be adjusted each year based on certain conditions that add (e.g., transfer in, move in-state) or remove (e.g., transfer, move out-of-state, death) students from the cohort. The extended graduation rate can vary from a five- to seven-year ACGR.

The U.S. Virgin Islands has been using the four-year ACGR under NCLB. The Task Force recommended including the five-year ACGR as part of the accountability system. The Task Force believed that including the extended ACGR sends a signal to high schools that even if students do not graduate in four years, schools continue to prepare them for post-secondary readiness. Thus, schools should be rewarded, rather than penalized, for continuing to support students who may not have the foundational experiences to be post-secondarily ready in four years. The Task Force recommended that the graduation rate indicator should be composed of both the four- and five-year ACGR. For the accountability system, the Task Force recommended that the four-year ACGR should be weighted slightly more heavily than the five-year ACGR, which will be determined during accountability standard setting. The ACGR calculation approach is described in the table below.

TABLE. 5. GRADUATION RATE CALCULATION

ACGR TYPE	CALCULATION METHODOLOGY
4-Year AGCR	$\frac{\text{The number of students who graduate in 4 years with a regular high school diploma}}{\text{The number of students who form the adjusted cohort for the graduating class}}$
5-Year AGCR	$\frac{\text{The number of students who graduate in 5 years with a regular high school diploma}}{\text{The number of students who form the adjusted cohort for the graduating class}}$

English Language Proficiency Progress. Measuring and reporting English language proficiency (ELP) has been a requirement under NCLB but has traditionally been a Title III requirement. Under ESSA, ELP is a now part of Title I accountability. This is largely because Title III accountability has now been rolled into Title I. One of the key tenets of accountability design is that the results of applying the accountability rules should not privilege or reward schools based on the demographic characteristics of the school. Given that ELP is a relevant indicator in only a certain percentage of schools, there are additional design considerations that must be examined to ensure that schools responsible for developing English language proficiency in their students are held accountable, but that the presence of this indicator does not automatically disadvantage the school in accountability determinations.

Additionally, ESSA specifies that the ELP indicator should be based on progress toward proficiency in the English language. This is an important distinction because it changes the number of factors that need to be considered. Historically, reporting and accountability has dealt only with the number and percent of students who exceeded a certain threshold on the ELP assessment (i.e., point-in-time status measure). Under ESSA, states and territories must now consider progress toward ELP, which at a minimum requires three information points: a student's (1) EL starting point, (2) ELP point, and (3) the time-bound expectation to close that gap. ESSA also suggests examining state data to determine what demographic characteristics may be important to consider when informing targets (e.g., age of entry into the U.S. and its territories, grade level when first assessed).

The Task Force was charged with three key decisions regarding ELs and how the EL subgroup should be included in the accountability system:

1. To assess students on the ELA assessment their first year of EL identification (to allow calculation of ELA growth scores in year 2) or to exempt them from the first year of ELA and test in year 2;

2. To specify recommendations for the measure in the ELP progress indicator; and

3. To recommend how the ELP progress indicator should be included in the accountability system in the face of some schools not having sufficient students in a school to include the measure.

For the first key decision, the Task Force recommended that first-year EL students should be assessed on the ELA assessment. While the initial year of ELA testing for a non-native speaker is perhaps the most difficult given language barriers, it does afford schools a chance to detect how much progress students are making in ELA in their first two years of EL instruction. The Task Force went on to recommend that those students' scores for achievement should be counted in year 3 to ensure the achievement score more accurately reflects ELA performance without the noise associated with being unable to access the content linguistically.

For the second key decision, the Task Force recommended that ELP progress use the current statewide ELP assessment. Furthermore, they recommended that the reclassification criteria (i.e., no longer classified as an English learner), time line for exit, and key student characteristics be informed by the ongoing work of WIDA (formerly World-class Instructional Design and Assessment). WIDA is the state membership organization that developed and maintains the territory's ELP assessment, ACCESS 2.0.

For the third key decision, the Task Force recognized that many of the schools that have large EL populations face different challenges than those that do not. Thus, the Task Force recommended that instead of imputing data for schools without the EL subgroup or changing the total possible points, a bonus/penalty point structure should be applied. That is, for schools with EL students, schools with exceptionally high progress toward ELP should be awarded bonus points and schools with exceptionally low progress toward ELP should be penalized in their score. This is described in more detail in the aggregation section that follows.

School Quality and Student Success. ESSA also requires the use of an indicator(s) of school quality or student success (SQSS) that meaningfully differentiates and is valid, reliable, and comparable. It is clear that the authors of ESSA wanted to broaden notions of school quality by including indicators in the system other than those based on test scores. However, the Task Force recognized that it needed to be especially thoughtful about this indicator while also considering potential measures carefully. The overall system should align with the Task Force's conceptions of school quality and the roles and purposes of accountability within the territory's educational system. Further, the previously mentioned constraints on data are especially critical here. The Task Force thought carefully about possible measures while considering the following four dimensions:

- ***Information provided to system.*** How much does the measure add to the understanding of school quality? Does it deepen the construct of performance or progress, or is it more focused on broadening the construct of school quality in general?
- ***Level of inference.*** The extent to which the data are direct measurements of some aspect of the school. Do the data require us to make informed interpretations about attitudes or behaviors (e.g., engagement surveys, curriculum audits) or do the data simply reflect counts of something (e.g., attendance rates, absenteeism rates, grades)?
- ***Potential corruptibility.*** How easily can the measure be gamed? Does collecting the data require subjective judgments or can they be based directly on objective sources?
- ***Level of data burden.*** Are the data already collected and is a reporting system in place? If not, what level of effort is required to develop a corresponding policy, collection method, analysis approach, and reporting system?

The Task Force considered these dimensions carefully and discussed both existing measures based on available data and newer conceptualizations of school quality or student success. Early recommendations sought to expand the construct of school quality to represent key characteristics like engagement, community involvement, post-secondary preparation, advanced coursework enrollment and performance, and extra-curricular activities.

Once Task Force members began running each proposed measure through the four dimensions listed above, they recognized that each carried inherent risk from a data quality, corruptibility, and sustainability perspective. While other data sources may be included in the future, the task force recommended the following measures be used for the School Quality/Student Success indicator. The measures are described in the table below for elementary/middle school and high schools.

TABLE 6. SCHOOL QUALITY/STUDENT SUCCESS MEASURES

ELEMENTARY/MIDDLE SCHOOLS		HIGH SCHOOLS	
Indicator	**Description**	**Indicator**	**Description**
Chronic absenteeism	The count and percentage of students who are chronically absent, or miss more than 10% of the school year.	Chronic absenteeism	The count and percentage of students who are chronically absent, or miss more than 10% of the school year.
Educator absenteeism	The count and percentage of educators who are chronically absent, or miss more than 10% of the school year.	Teacher absenteeism	The count and percentage of educators who are chronically absent, or miss more than 10% of the school year.
		Post-secondary readiness	A measure of student readiness for either college or careers. It includes the count and percentage of students who are successful in advanced placement courses, dual enrollment, or in career-technical education concentrations.

As noted in the table above, there are two common measures across all schools and an additional set of measures included for high schools. The two common measures of student and educator absenteeism can be conceptualized as a way to track lost instructional time. For students, missing 10% of the year (i.e., 18 days or nearly one month of school) is a common gateway to a host of academic problems. The Task Force believed that this indicator could help schools highlight students of concern and spark behavioral change among educators and administrators. While the educator attendance may be more structural in nature due to leave policies, the Task Force also believed that any time an educator is not present, there is a loss of instructional time. As such, the Task Force recommended that the VIDE further study the educator absentee rates and make a decision that is most conducive to communicating high expectations for educators and students. This led the VIDE to include educator absenteeism as a function of any out-of-classroom days (including professional development) to capture the notion of lost-instructional time.

For high schools, the Task Force believed there should be a reflection of post-secondary readiness that supplements the state's assessment for accountability. After examining multiple sources of data and considering data quality, corruptibility, and sustainability, the Task Force recommended the use of CTE performance, AP performance, and dual enrollment performance. The Task Force recognized the risk of basing high-stakes accountability decisions on enrollment data because of the risks associated with enrolling students for the sake of accountability rather than for the educational benefit of the student. By focusing on performance, this ensures that outcomes are the focus and that schools are expected to hold students to a high standard. However, the Task Force noted that grades are subject to grade inflation and carefully

considered how the post-secondary measures should be operationalized. The way in which the three measures are included in the system are based on:

- Career and technical education course grades are based only on those courses that can lead to national certification. This measure is based on the percent of CTE concentrators who pass the following course concentrations:
    - Architecture and Construction
    - Finance
    - Health Science
    - Hospitality and Tourism
- The Advanced Placement measure is based on those students who are enrolled in AP courses and who score 3 or higher on the AP exam
- The dual enrollment measure is based on those students who are enrolled in dual enrollment courses and receive a passing grade. Due to the low n-count of students who are enrolled in dual enrollment courses in both the St. Thomas and St. Croix/St. John districts, these students will be combined with AP student performance for schools. Once a sufficient number of students are identified using a rolling three-year average, this measure will be included as a stand-alone measure and reported accordingly.

The measures for the SQSS indicator will be combined equally depending on the number of measures available (i.e., 2 for elementary/middle schools and 5 for high school). As additional measures are considered, data will be examined, presented to members of the Task Force and VIDE, and potentially included in the SQSS indicator.

AGGREGATION METHOD

After the Task Force made recommendations on the individual components of the accountability system, the members turned their attention to the indicator weights and aggregation approach. The Task Force considered two primary approaches for how the system components should roll up: a decision matrix approach and an index-based approach.

A decision matrix is a series of decision rules that specify how system indicators interact to define school performance and quality. These rules and indicator interactions result in profiles of school performance that demonstrate certain levels of school quality. While the Task Force would provide an initial set of recommendations, a subsequent committee would specify school performance standards by (1) reviewing school profiles of performance, (2) determining the decision rules related to each school profile, (3) articulating guardrails for the system, and (4) specifying any requirements for meeting expectations. An index approach uses numerical aggregations of performance across indicators that result in an overall school score for each school. This would require agencies to determine performance standards for each of the individual indicators as well as the aggregation rules across indicators. Index approaches may also include conjunctive or compensatory rules, or a combination of both.

The Task Force discussed the pros and cons for each aggregation approach and recommended the use of an index for schoolwide accountability ratings. This recommendation was based on a desire to promote school differentiation, ranking, and to establish high expectations that can be adjusted over time. The Task Force also wanted to have the option of setting specific targets for each indicator.

Aggregation weights. The Task Force recognized that when specifying an index, the territory's theory of action and desired signals should drive the indicator weights. The members further recognized the challenges associated with resource constraints, the limited number of schools in the territory, and existing performance challenges. Thus, the Task Force recommended that student progress and graduation rates should be the highest priority signals for elementary/middle schools and high schools, respectively. However, the Task Force also believed that academic achievement should not be overshadowed in the system. Thus, the Task Force recommended that academic achievement be heavily weighted, only slightly less than growth and graduation rates. The recommended weights are presented in the table below.

TABLE 7. ACCOUNTABILITY INDEX WEIGHTS.

ELEMENTARY/MIDDLE SCHOOLS	HIGH SCHOOLS	WEIGHT VALUE
Student Progress	Graduation Rate	40%
Academic Achievement	Academic Achievement	35%
School Quality/ Student Success	School Quality/ Student Success	25%
ELP Progress	ELP Progress	+/- 10% of Index Points Earned

As indicated in the table above, the SQSS indicator comprises 25% of a school's score. The Task Force recognized that the number of SQSS measures for elementary/middle schools and high schools would differ and that each measure for high school would represent fewer points because there are more measures included (i.e., 5 for high schools and 2 for elementary/middle schools). However, the Task Force recommended that the VIDE continue to explore possible measures for elementary and middle schools to include in the accountability system in the future.

The Task Force discussed how to best represent the ELP progress indicator in the index. The members were concerned about the unintended negative impact that ELP progress could place on schools that had an EL subgroup compared to schools that did not. For example, a school without an EL subgroup would not have the opportunity to receive the ELP progress points, thereby reducing the total number of points necessary to receive a certain rating. Conversely, a school with an ELP subgroup would have to earn more points to receive the same rating. While this initially appears to be a comparable proportion of necessary points, in practice it makes it more difficult for a school with an EL subgroup to receive the same rating. This example is shown in the table below.

TABLE 8. ELP PROGRESS EXAMPLE.

INDICATOR	POINTS POSSIBLE	SCHOOL A: EL SUBGROUP	SCHOOL B: NO EL SUBGROUP
		Points Earned	Points Earned
Growth	30	20	20
Achievement	30	17	17
SQSS	25	22	22
ELP Progress	15	7	NA
Total Points	100	66 out of 100	59 out of 85
% of Points		66%	69.4%

As seen in the table above, each indicator is assigned a point value out of a possible 100 point index score. Additionally, School A has an EL subgroup and School B does not. Both schools performed identically on the indicators of the system. However, School A would have to obtain at least 10 of the 15 points on the EL indicator to receive the same proportion of points as School B.

This discrepancy could be addressed by adjusting the proportion of the points possible or by recognizing particularly high or low progress on the ELP indicator. The Task Force preferred the latter approach by identifying differentiating between average, extremely high, and extremely low progress when preparing ELs for English proficiency. Thus, the Task Force recommended that all schools should have the same number of points available in the index, but schools with extremely high ELP progress are awarded bonus points and that schools with extremely low ELP progress are deducted points. Ultimately, this results in either a 10% bonus or penalty to the number of index points a school has earned.

TABLE 9. REVISED ELP PROGRESS EXAMPLE.

INDICATOR	POINTS POSSIBLE	SCHOOL A: EL SUBGROUP	SCHOOL B: EL SUBGROUP	SCHOOL C: NO EL SUBGROUP
		Points Earned	Points Earned	Points Earned
Growth	40	25	25	25
Achievement	35	17	17	17
SQSS	25	23	23	23
ELP Progress	10%	High Progress (+10% or 6.5 pts)	Low Progress (-10% or -6.5 pts)	NA
Total Points	100	72.5 out of 100	58.5 out of 100	65 out of 100
% of Points		72.5%	58.5%	69.4%

REPORTING APPROACH

The final set of Task Force discussions revolved around reporting performance on each indicator and the overall system. ESSA requires that states report and disaggregate each indicator in the accountability system as well as the overall designation for schools. The Task Force recognized the value of this requirement and sought to make recommendations that supported the public's understanding of school and subgroup progress. The Task Force reviewed several examples of state accountability reports to determine the types of features territory reports should possess. These were translated into recommendations for the VIDE to consider as the department develops accountability reports.

The Task Force was asked to consider several characteristics of reporting, which included but were not limited to:

- The kind of signals to send to the public and educators when describing schools
- How to communicate school designations
- Whether ESSA required data be supplemented with lower-stakes information based on the theory of action
- Whether users should have the ability to zoom in or out to make comparisons across districts and the territory

After reviewing a number of reports, the Task Force believed that a degree of comparison is important to help the public understand performance relative to both a criterion and other schools in the territory. That is, reporting should showcase the progress students and school are making toward high expectations (e.g., proficiency, college-and-career-readiness, English language proficiency) and how scores compare to the highest, lowest, and average performing schools in the territory.

The Task Force also recommended that using reporting tools similar to those already being used by districts and schools would facilitate ease of use and a greater understanding of a new accountability system. This should be extended to grade, subgroup, school, and district reporting of high-stakes indicators. Recommendations also included that school- and district-specific reporting tools be developed to support local and contextually-dependent reports, highlights, and descriptions to help tell a school's story. The VIDE is expected to continue exploring options to develop reports that inform the public and educators within the scope of resources and department capacity.

NEXT STEPS

The Task Force provided a great deal of guidance in their recommendations to the Commissioner and the VIDE in developing the accountability system. The VIDE is now translating those recommendations into data tables, analyses, and draft accountability ratings. These ratings will be taken to a Standard Setting Committee that will recommend performance standards for the Virgin Islands Accountability System through a series of data reviews and recommended decision rules. The standard setting recommendations will then be translated into accountability system business rules that will be applied to 2015-2016 school year data for the Commissioner's review and approval. The 2015-2016 data are being used because the data have been validated for use in accountability.

Once approved, the preliminary ratings (using 2016-2017 data) will be released to educators and administrators in the territory. The preliminary ratings will not be used to determine supports or interventions and will not be communicated to the public. The preliminary ratings are instead used to introduce and educate district and school staffs on VIDE's accountability system. While low-stakes in nature, this first round of dissemination will be critical to ensuring educators and administrators in the territory understand the vision, guiding principles, and results of the accountability system.

Throughout the 2017-2018 school year, the VIDE will also reach out to the public to educate them on the new accountability system using 2016-2017 data. Beginning in the 2018-2019 school year, the VIDE will replicate accountability analyses and deploy them using 2017-2018 data year and start "high-stakes" reporting and improvement activities. Schools will be classified on a 5-star rating system. These policy descriptors (PLDs) associated with each star rating will be developed over the next several months based on the guidance of VIDE leadership. These policy level descriptors will be used as the basis for reporting and to determine performance standards for the accountability system. The PLDs may be refined using the VIDE leadership team's feedback. A draft timeline of remaining accountability tasks is provided below.

TABLE 10. REMAINING ACCOUNTABILITY TASKS AND TIMELINES*

ACTIVITY	DATE
1. Conduct a performance standard setting meeting for the accountability system	Quarter 4, 2017
2. Confirm standard setting design decisions and obtain VIDE leadership approval	Quarter 1, 2018
3. Update Task Force on final design decisions	Quarter 1, 2018
4. Model accountability design decisions on 2016-2017 data (finalized in December 2017) to confirm fit and make any necessary changes	Quarter 2, 2018
5. Develop and deploy a reporting system for 2016-2017 data	Quarter 3, 2018
6. Document decisions and develop ESSA-style accountability plan	Quarter 3, 2018

*Dates have been adjusted due to the impact of Hurricanes Irma and Maria.