

# Virginia's K-12 School Accountability System

# Study resolution

- Directs staff to review Virginia's recently revised K–12 accountability system for public schools
  - role and purpose in ensuring quality education and incentivizing school performance
  - appropriate criteria, sufficiently transparent, and understandable
  - appropriate balance between accountability, assistance, and improving educational quality
  - alignment with federal requirements and other states' accountability systems

Commission resolution (November 7, 2024)

# Primary research activities

- Reviewed federal requirements, guidance, and Virginia's Every Student Succeeds Act (ESSA) plan
- Compared Virginia's new accountability system to Virginia's prior system and other states' systems
- Conducted extensive analysis of Virginia's new accountability system's design and results
- Surveyed local school divisions

---

## Primary research activities (cont.)

- Interviewed national subject-matter experts, school division superintendents, school principals and counselors, local school board members, VBOE members, VDOE staff
- Viewed VBOE meetings where school accountability was discussed (2022 to 2025) and reviewed public comment

VBOE = Virginia Board of Education. VDOE = Virginia Department of Education.

---

## In brief

The new accountability system is more effective and useful than the state's prior system, but refinements are needed.

The system's labels do not clearly communicate performance, nor relate on a continuous scale.

The scoring thresholds do not account for differences in how school levels are measured, benefiting high schools.

The system weights proficiency more than growth, which captures less of schools' contribution to learning.

Scoring used for the “3Es” indicator is problematic.

VDOE is redesigning its school improvement program, but it is too early to determine its effectiveness.

---

# In this presentation

## Background

Overall assessment of new system

Results from new system

Labeling school performance

Proficiency and growth indicators

Other potential refinements

VDOE's school improvement program

# K–12 school accountability systems serve several key purposes, and results are highly visible

- Measuring and reporting school performance:
  - Provides stakeholders with information about school quality
  - Holds school leaders accountable and motivates them to improve performance
  - Identifies schools and students that need additional support (assistance from state staff and access to funds)
- Accountability results are highly visible to the public and carry implications for schools

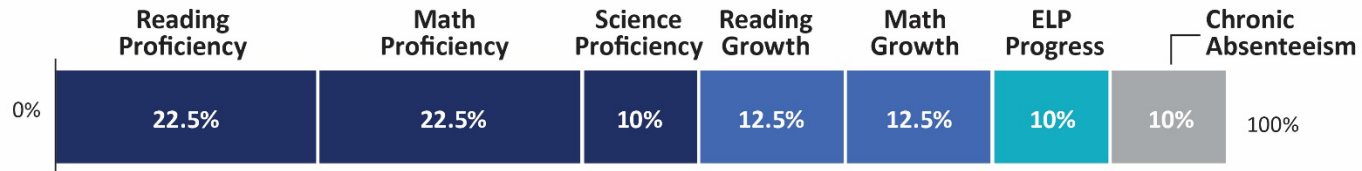
## VBOE approved a new K–12 accountability system in August 2024

- The School Performance & Support Framework (SPSF) replaced the prior system in place since 2017
- SPSF went into effect beginning with the 2024–25 school year; 2024–25 results were released in December 2025
- SPSF measures school performance using indicators such as standardized test performance, graduation rates, and attendance
- Each school receives an overall rating that is made public
- Schools with lower ratings may be eligible for state assistance and state or federal funding

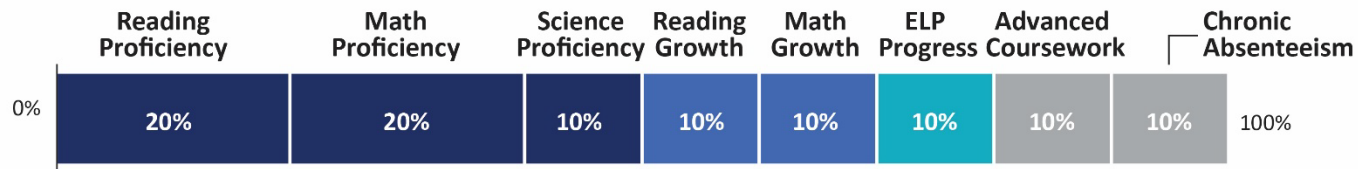


# SPSF uses a mix of indicators with varying weights for the 2024–25 year

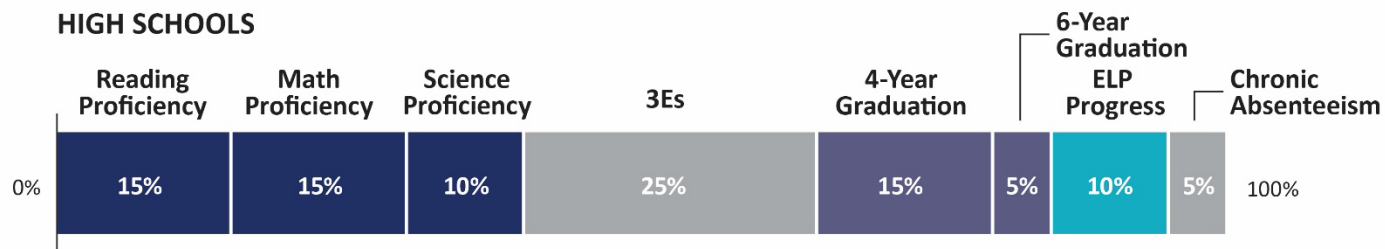
## ELEMENTARY SCHOOLS



## MIDDLE SCHOOLS



## HIGH SCHOOLS



“3Es” = Enrollment, employment, and enlistment indicator (college, career, and military readiness ). ELP Progress = English learners’ progress toward English language proficiency.

# Schools receive a score for each indicator, which combine into a summative score (example)

	Unweighted indicator score		Percent weight		Indicator score
Reading Proficiency	80	X	22.5%	=	18.0
					+
Math Proficiency	85	X	22.5%	=	19.125
					+
Science Proficiency	100	X	10%	=	10.0
					+
Reading Growth	85	X	12.5%	=	10.625
					+
Math Growth	70	X	12.5%	=	8.75
					+
ELP Progress	65	X	10%	=	6.5
					+
Chronic Absenteeism	90	X	10%	=	9.0
Summative score (sum of indicator scores)					82.0

# SPSF assigns each school a performance label and, for some schools, a federal designation

<b>Performance label</b>	<b>Summative score</b>
<i>Distinguished</i>	$\geq 90.0$
<i>On Track</i>	80.0-89.9
<i>Off Track</i>	65.0-79.9
<i>Needs Intensive Support</i>	$< 65.0$

- Scores used to assign a federal designation to some schools with low performance

NOTE: Federal designations can affect schools' performance labels. According to VBOE regulations, Targeted Support and Improvement schools (TSI) and Additional Targeted Support and Improvement schools (ATSI) have their performance label lowered by one level, and Comprehensive Support and Improvement (CSI) schools cannot be labeled as *Distinguished*.

# VBOE is considering several adjustments to the SPSF

- As of December 2025, board is considering several proposed changes to the framework
  - Three of the changes would implement JLARC recommendations or a policy option
  - Several would represent substantial changes but were made too recently for JLARC to evaluate
  - Some are narrow or technical
- Most of these changes would be for future school years

---

## SPSF results will be affected by changes to system and when SOL “cut scores” increase

- Changes currently being considered by VBOE *could* substantially affect results
- VBOE’s recent action to raise SOL “cut scores” will result in substantial declines in SPSF’s overall school ratings once fully implemented if no additional changes are made

---

# In this presentation

Background

**Overall assessment of new system**

Results from new system

Labeling school performance

Proficiency and growth indicators

Other potential refinements

VDOE's school improvement program

# Complexity of K–12 education and varied priorities mean there is no “best” system

- Designing an accountability system requires balancing state policy priorities, stakeholder concerns, and transparency
- System is limited to data that is produced consistently and reliably statewide
- Systems have inherent tradeoffs:
  - which areas of performance to measure
  - decisions about weighting, measurement, and scoring of each indicator
  - balancing simplicity vs. precision

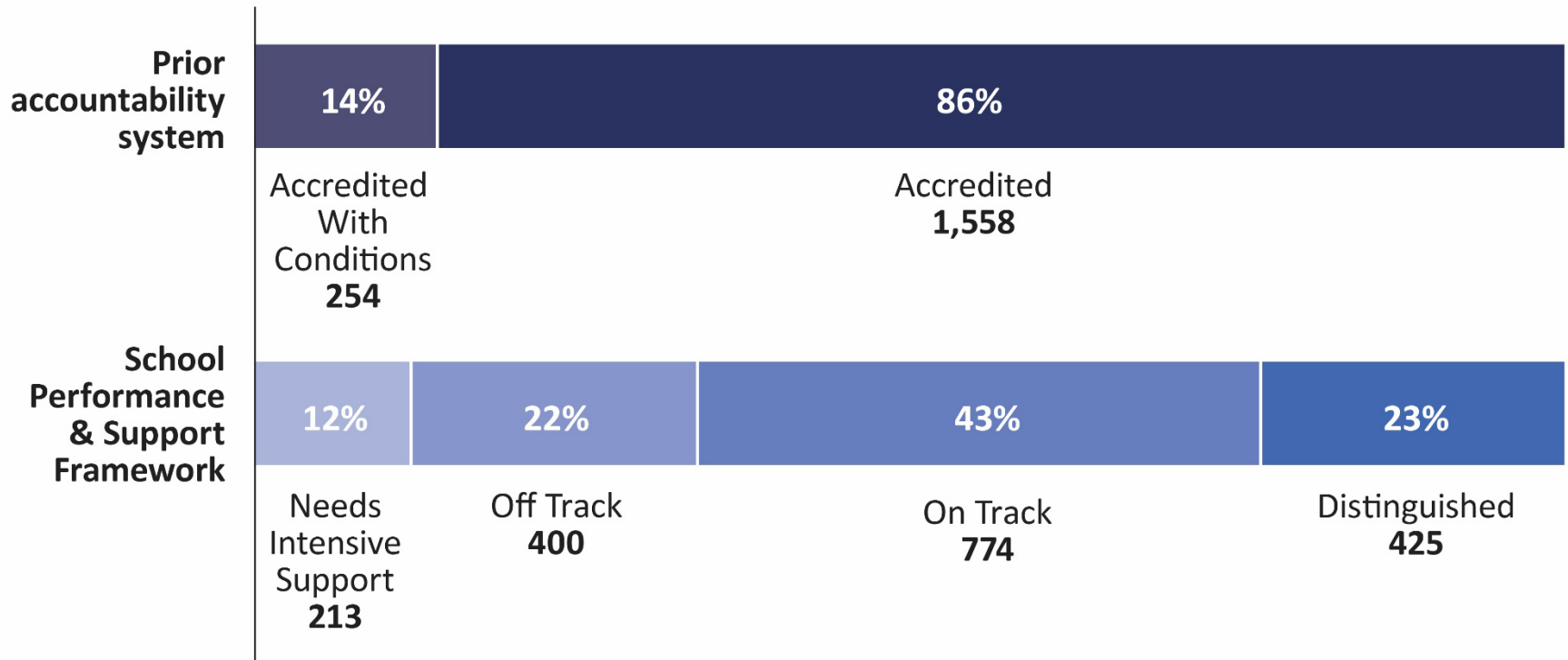
---

## Finding

Virginia's new accountability system is more effective and useful than the state's prior system.



# SPSF better differentiates school performance than the prior system (2023–24 to 2024–25)



2023–24 school year results for the prior system compared to 2024–25 school year results for the SPSF.

# SPSF aligns Virginia's accountability system with federal requirements; separates out accreditation

- The SPSF is a single system used for state and federal accountability purposes
  - identical indicators, weighting, measurement, and scoring
  - less complex than prior approach, which had two separate systems that measured performance differently
- The SPSF separates school accreditation (i.e., compliance) and accountability (i.e., performance)
- Subject-matter experts and majority of school divisions viewed these changes positively

# Framework is more precise and inclusive than the prior system

- Disaggregates previously combined indicators for more transparent and useful results
- Uses index scoring for academic indicators\* to incentivize focus on students across the entire achievement spectrum, rather than binary pass/fail scoring that incentivizes focusing only on students “on the bubble”
- Lower minimum student threshold (“N-size”) means more students are included in school accountability calculations and reported results

\*Index scoring used for proficiency and growth.

---

## Finding

The SPSF structure is generally sound, but refinements are needed.

# SPSF's general structure should be used for school accountability in Virginia

- Structural elements (e.g., indicators, weighting, measurement, and scoring) are generally consistent with:
  - federal guidance
  - expert opinion or recommendations
  - other states' accountability systems
- Fundamental changes to the system are not needed
  - unnecessarily disruptive
  - administratively burdensome to school divisions
- JLARC is recommending refinements to the framework

---

## In this presentation

Background

Overall assessment of new system

**Results from new system**

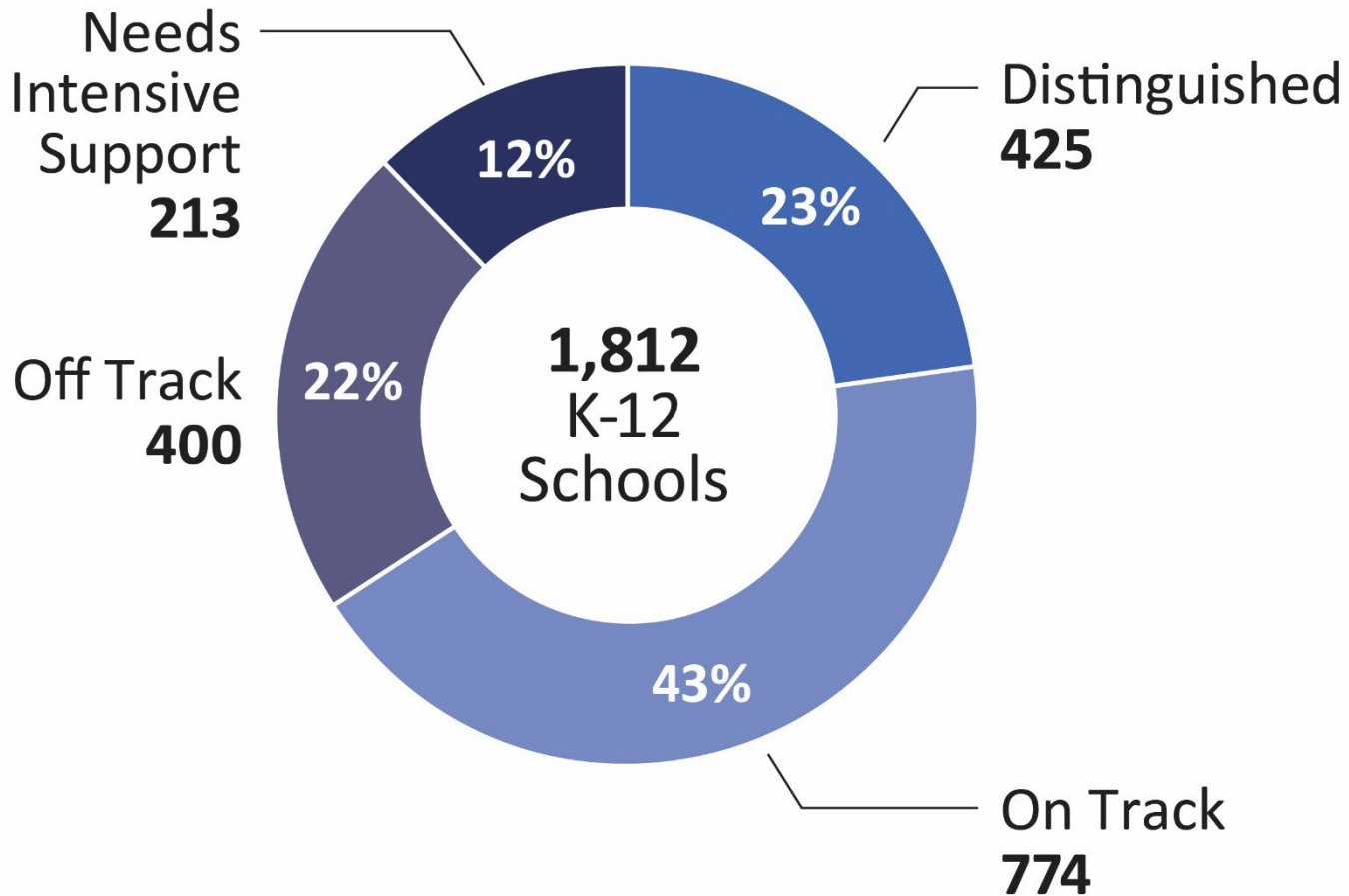
Labeling school performance

Proficiency and growth indicators

Other potential refinements

VDOE's school improvement program

## Majority of schools labeled *Distinguished* or *On Track* in 2024–25



# Finding

Economic disadvantage has the strongest relationship with schools' SPSF results among all student demographic and school characteristics.\*

Schools with more economically disadvantaged students tend to receive lower overall ratings from the SPSF.

\*Additional student and school characteristics examined include students who are English learners, Asian, Black, Hispanic, white, or two or more races, and students with disabilities; provisionally licensed teachers, out-of-field teachers, and teacher vacancies; per-pupil spending; enrollment; population density; VDOE superintendent region; and school level.

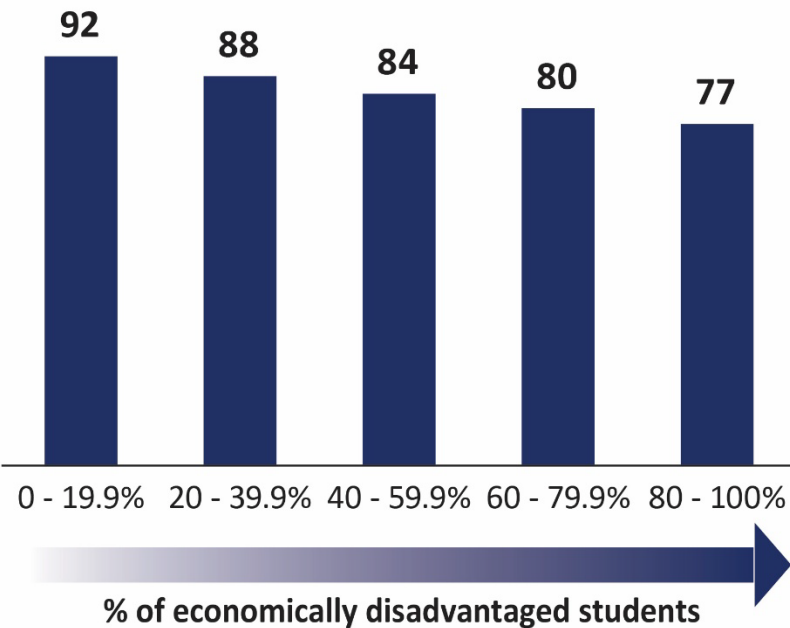


## Accountability results often reflect demographics; lack of consensus about whether appropriate

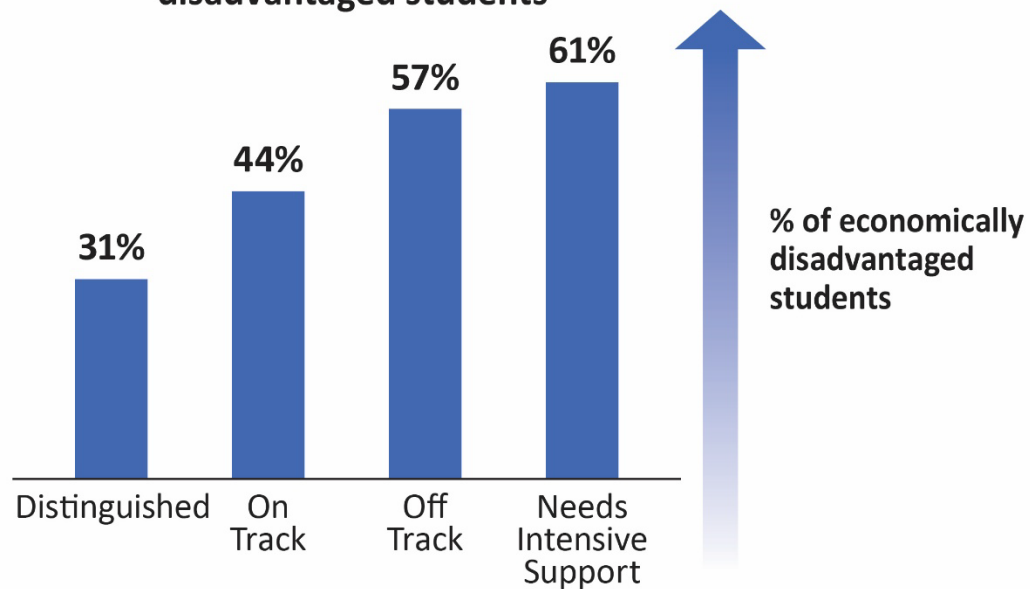
- Typical for school accountability because many commonly used indicators are correlated with demographics
- Lack of consensus among experts about how much results should reflect student demographics
  - *Point:* Results that strongly reflect student demographics maintain the same expectations for all students and can help to target support to address disparities
  - *Counterpoint:* Results that are overly reflective of student demographics evaluate schools on factors outside of their control rather than on their effectiveness

# Schools with more economically disadvantaged students tend to receive lower ratings (2024–25)

Average SPSF summative score



Average % of economically disadvantaged students



---

## In this presentation

Background

Overall assessment of new system

Results from new system

### **Labeling school performance**

Proficiency and growth indicators

Other potential refinements

VDOE's school improvement program

# School performance ratings are highly visible and carry implications for schools

- Accountability results are a primary way of communicating school performance
- States often aggregate accountability results into single overall rating, though they are not required to
- Various formats for school ratings, e.g., letter grades, stars, colors, descriptive labels (i.e., terms)
  - No evidence one format better than another
- Ratings can shape public perceptions; affect school staff morale, recruiting, and retention; and influence state and local policy, such as funding

# SPSF uses four descriptive labels to communicate results, which can be challenging to develop

- SPSF uses descriptive labels that reflect how well each school meets the state's performance expectations
  - *Distinguished, On Track, Off Track, Needs Intensive Support*
- According to VBOE and VDOE leadership, the labels use plain language that stakeholders understand, avoid jargon, and are not ambiguous
- Developing clear, widely understood, and broadly accepted descriptive labels can be challenging

---

## Finding

Some SPSF performance labels do not clearly communicate school performance or function as part of a continuous scale.

# SPSF labels do not clearly communicate school performance

- *On Track* and *Off Track* imply trend of schools' performance (i.e., improving or declining)
  - labels not intended to communicate trend, and trend may be inaccurate (*Off Track* school could be improving)
- *Distinguished* overstates some schools' performance
  - *Distinguished* school could underperform on some indicators or have weaknesses outside of accountability system
- Many division and school staff were particularly concerned with the *Off Track* label; viewed as “too negative”

## SPSF labels do not clearly relate to each other as part of an intuitive, continuous scale

- No intuitive progression through four labels, and labels use disconnected terms
- Other states use descriptive labels with continuous scales

### Virginia:

- *Distinguished*
- *On Track*
- *Off Track*
- *Needs Intensive Support*

### West Virginia:

- *Exceeds Standard*
- *Meets Standard*
- *Partially Meets Standard*
- *Does Not Meet Standard*



---

## Recommendation

VBOE should develop new terminology for the SPSF performance labels that more clearly communicates school performance as part of a continuous scale.

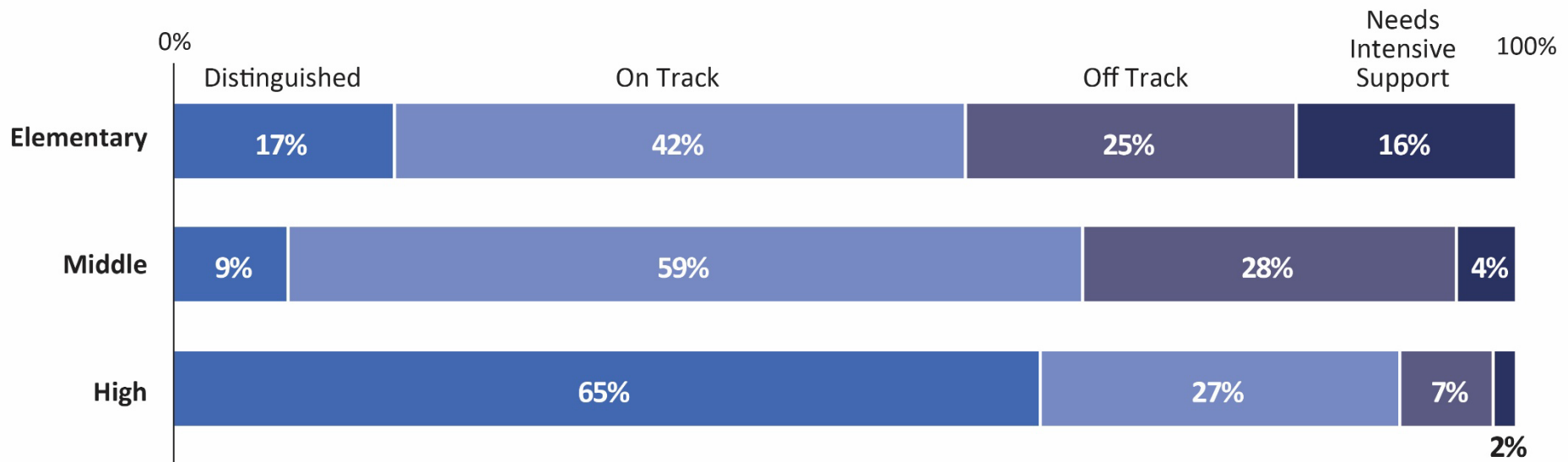
---

## Finding

The SPSF scoring thresholds ignore differences in how school levels are measured, significantly benefiting high schools.

# High schools significantly more likely than other schools to have higher ratings

- Elementary, middle, and high schools are evaluated on different indicators with different weighting
- However, SPSF uses same scoring thresholds for each level, which reduces comparability and fairness of results



---

## Recommendation

VBOE should develop new SPSF scoring thresholds, unique to each school level, to result in comparable proportions of schools at each level receiving each label.

# States must use their accountability systems to identify low-performing schools

- ESSA requires states to identify lowest-performing Title I schools
  - Comprehensive Support and Improvement (CSI)
- ESSA also requires states to identify schools with underperforming student subgroups\*
  - Targeted Support and Improvement (TSI)
  - Additional Targeted Support and Improvement (ATSI)
- These designations are in addition to SPSF labels

\*Student subgroups include “major” racial and ethnic subgroups, economically disadvantaged students, English learners, and students with disabilities. In Virginia, the “major” racial and ethnic subgroups include Asian, Black, Hispanic, and white students and students of multiple races.

---

## Finding

The SPSF should ensure that Comprehensive Support and Improvement schools (the most serious federal designation) receive the lowest state performance label.

# SPSF does not ensure CSI schools receive the lowest state label, sending inconsistent message

- CSI is the most serious federal designation for low performance
- Many CSI schools were labeled *Off Track* in 2024–25, rather than *Needs Intensive Support*
- VBOE intended to label CSI schools *Needs Intensive Support* automatically, but policy not implemented
- Experts advise CSI schools should receive state's lowest rating to ensure:
  - federal and state labels send consistent message
  - lowest-performing schools are prioritized for support

---

## Recommendation

VBOE should amend regulations so that any school with a federal Comprehensive Support and Improvement designation is labeled in the lowest state performance category.



---

## Finding

The SPSF lowers all Targeted Support and Improvement and Additional Targeted Support and Improvement schools' performance labels, which is imprecise and overly punitive.

## SPSF lowers all TSI & ATSI schools' labels, which is imprecise and overly punitive

- SPSF lowers TSI and ATSI schools' performance labels so that subgroup performance is factored into schools' accountability results
  - e.g., TSI school scoring in *On Track* range lowered to *Off Track*
- Lowering all TSI and ATSI schools' performance labels
  - does not account for difference between the two federal designations
  - is unnecessary because many of these schools are already labeled as low-performing (e.g., *Off Track* or *Needs Intensive Support*)

---

# Recommendation to change regulations for how TSI & ATSI designations affect performance labels

- Regulations can be changed to:
  - maintain VBOE's emphasis on subgroup performance
  - better align with intent of federal designations
  - avoid being unnecessarily punitive

---

## In this presentation

Background

Overall assessment of new system

Results from new system

Labeling school performance

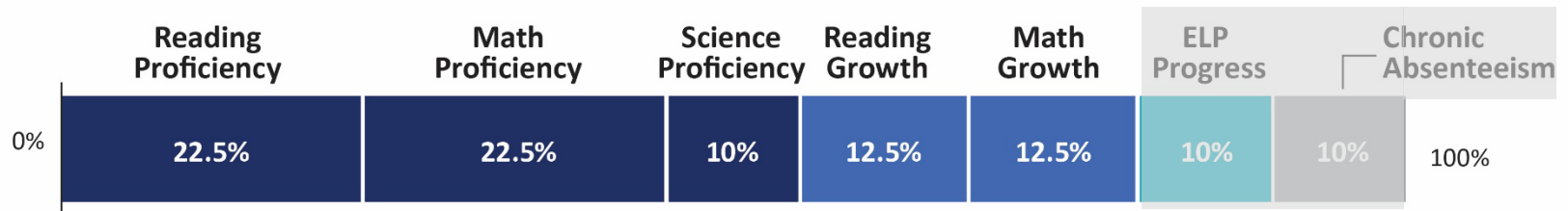
**Proficiency and growth indicators**

Other potential refinements

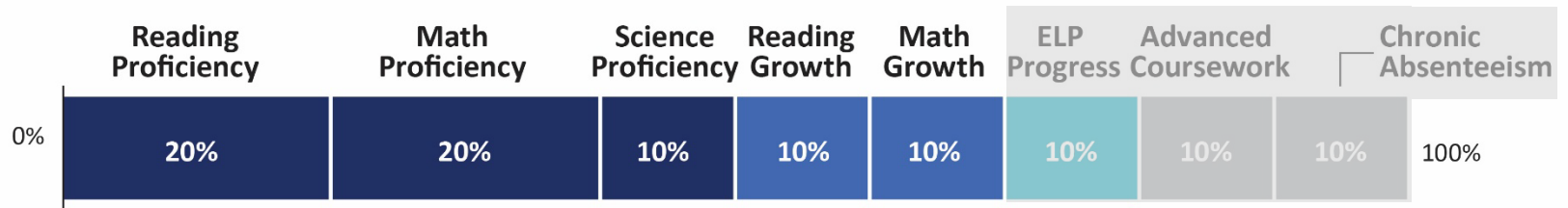
VDOE's school improvement program

# Proficiency and growth account for a majority of rating for elementary and middle schools

## ELEMENTARY SCHOOLS



## MIDDLE SCHOOLS



SPSF weighting for the (2024–25) school year.

---

## Finding

The SPSF's proficiency indicators are generally well designed.

The SPSF's growth indicators are an improvement over the prior system but still have shortcomings and will likely be changed.

# SPSF proficiency indicators are generally well designed

- Standalone indicators for reading, math, and science
  - Improvement over prior system's combined "achievement" indicators
  - Improves transparency into which factors drive performance
- Use of index scoring
  - Schools receive full credit for students who score proficient, partial credit for students who score below proficient, and extra credit for those scoring advanced
  - Encourages focus on students at all achievement levels, more so than prior system's pass rates

# SPSF's growth indicator is an improvement, but scoring raises concerns and will likely change

- The SPSF's growth indicators are an improvement over the prior system:
  - standalone growth indicators
  - measure growth for all students
- Growth model\* used for 2024–25 is generally appropriate
- Framework's *scoring* of growth indicators may not have accurately and fairly reflected growth, but VBOE will likely change growth measure for the 2025–26 school year

\*SPSF measured growth in reading and math for students in grades four–eight using a value-added model, Virginia's Visualization and Analytics Solution (VVAAS).



---

## Finding

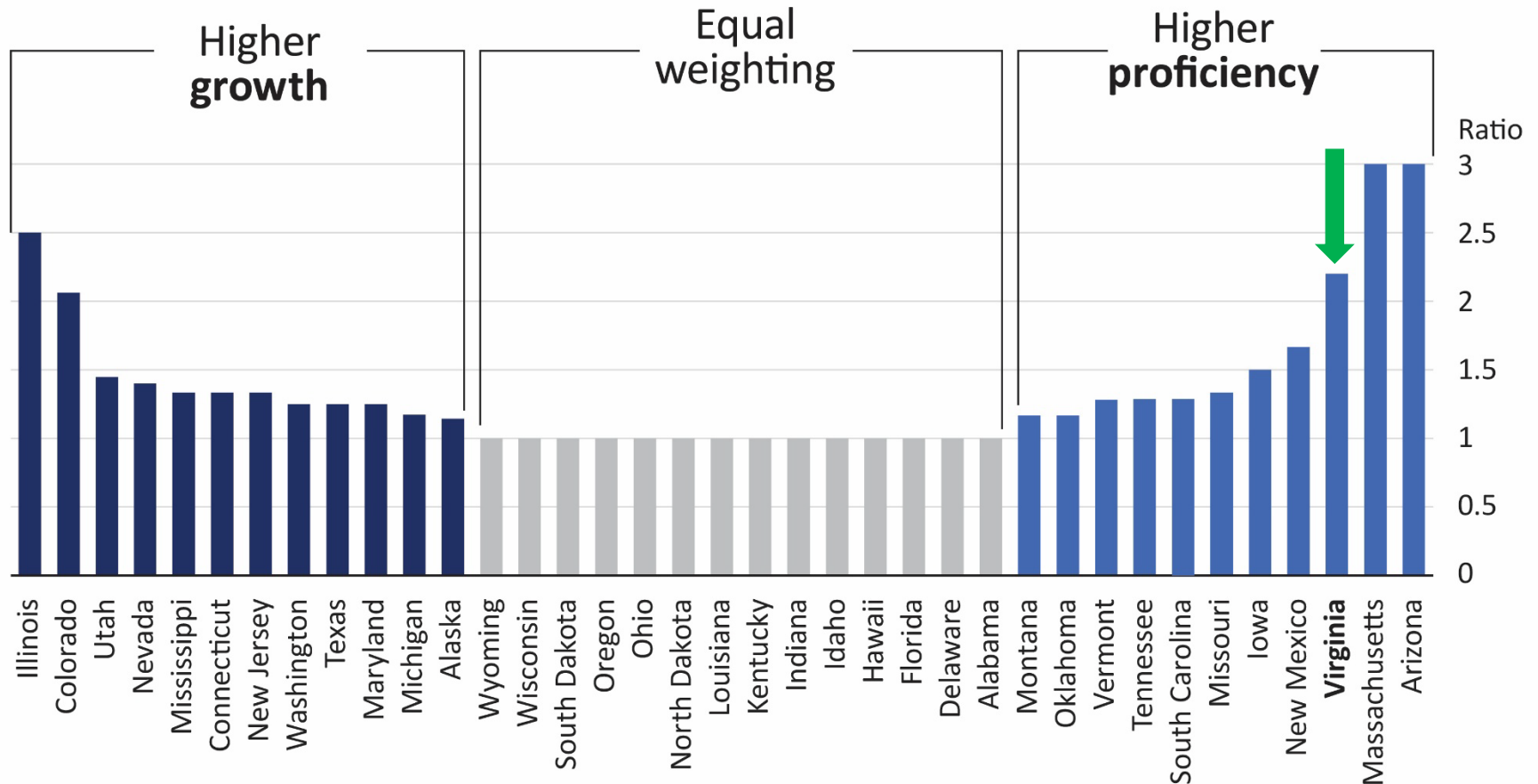
The SPSF's much heavier weighting of proficiency over growth does not fully capture schools' contribution to student learning and increases the impact of student demographics on schools' scores.

---

# Both proficiency and growth measure important aspects of student academic achievement

- Proficiency:
  - measures how students perform against state standards
  - identifies schools with students not meeting standards
- Growth:
  - measures student performance from one year to the next
  - identifies schools whose students are not making enough progress

# SPSF weights proficiency much higher than growth compared with most other states



Ratio of combined proficiency indicator weighting / combined growth indicator weighting for elementary schools.

# Weighting is policy decision, but growth better captures school contributions to student performance

- VBOE's reasons for high proficiency weight include:
  - meeting grade-level standards is a key outcome
  - maintaining high expectations for all students
  - some stakeholders care most about student proficiency
- Proficiency results tend to be highly correlated to student demographics, especially economic disadvantage
- In contrast, student growth results:
  - are less correlated with student demographics
  - better capture *schools'* contribution to student performance

---

## Policy option

VBOE could increase the weight assigned to growth indicators to more fully account for schools' contribution to student learning.

---

## In this presentation

Background

Overall assessment of new system

Results from new system

Labeling school performance

Proficiency and growth indicators

**Other potential refinements**

VDOE's school improvement program

---

# Finding

Scoring used for the SPSF's college, career, and military readiness indicator (the “3Es”) is problematic.

“3Es” = Enrollment, Employment, and Enlistment.

# Students can earn points for “3Es” achievements across multiple pathways

- High schools earn points for student achievements\*, and one student can score points across multiple pathways
- Intended to incentivize schools to make opportunities available and to encourage broad participation

**Example:** A single student with achievements in all 3 pathways



\*Examples of 3Es achievements include completion of career & technical education, work-based learning, AP and dual-enrollment, and the Armed Services Vocational Aptitude Battery (ASVAB).



## High schools can score much higher on 3Es, which distorts schools' overall scores

- Schools can potentially score much higher on 3Es than on other indicators
  - Overall score driven more by 3Es than intended by its weight
  - Distortion of schools' overall score, which is misleading

	Intended weight (SPSF 3Es weighting)	Effective weight (Proportion of summative score comprised of 3Es)
Schools with an unweighted 3Es indicator score >125	25%	34%

---

## Recommendation

VBOE should change the 3Es indicator scoring by either capping the total number of points (i) a school can be awarded for an individual student or (ii) a school can receive in the 3Es category.

---

## Finding

SPSF's rules for reweighting English Language Proficiency (ELP) progress indicator for elementary and middle schools with too few English learners reduces comparability and fairness of results.

## SPSF's reweighting rules reduce comparability and fairness of results

- 42% of schools have too few (< 15) English learners for the ELP progress indicator to count toward their overall score
- ELP progress indicator's 10% weight is redistributed:
  - proportionally across all other indicators for high schools
  - to growth indicators for elementary and middle schools
- Redistributing weight to growth indicators:
  - alters how a substantial portion of elementary and middle schools are evaluated compared to one another
  - reduces the comparability and fairness of results

---

## Recommendation

VBOE should modify the business rules in the SPSF so that the English Language Proficiency (ELP) progress indicator's weight is proportionally redistributed for elementary and middle schools with too few English learners to receive a score for ELP progress.

## Several additional refinements identified in the report could help strengthen the SPSF

- Changing the score awarded by the proficiency scoring index for students scoring “fail” on tests with three performance levels
- Changing scoring of the enlistment category of the 3Es indicator to be more commensurate with other categories
- Transitioning from “Exception A” to “Exception B” for determining when recently arrived English learners test scores are counted in proficiency and growth calculations
- Establishing a scoring index for the ELP progress indicator

---

## In this presentation

Background

Overall assessment of new system

Results from new system

Labeling school performance

Proficiency and growth indicators

Other potential refinements

**VDOE's school improvement program**

# Key purpose of accountability is to identify schools that need improvement and provide support

- VDOE provides school improvement services to schools identified as low performing
  - Training, technical support, materials, and sharing of best practices
  - Distribution of federal school improvement grants and other federal and state grants
- Low performance in 2024–25:
  - 434 schools received a federal designation (including all *Needs Intensive Support* schools)
  - 191 additional schools labeled as *Off Track* but received no federal designation



---

## Findings

VDOE is redesigning its school improvement program; appears to be an improvement over prior approaches.

Lack of written plan for program and staffing workloads may hinder implementation.

A full evaluation cannot occur until services have been provided.

# VDOE is redesigning and implementing a new school improvement program in conjunction with SPSF

- A 2020 JLARC evaluation found problems with VDOE's school improvement services; changes in VDOE leadership and philosophy have stagnated efforts to improve its effectiveness
- Current VDOE leadership has prioritized redesigning agency's school improvement program
- Redesigned program has positive features and appears to be an improvement over prior approaches
  - Tailored support delivered by subject-matter specialists
  - Prioritizing low-performing schools for existing state grants

## Lack of written plan for program and staffing workloads may hinder implementation

- Lack of a written plan for the large-scale, complex redesign of VDOE's school improvement program could impede successful implementation
- New school improvement program's emphasis on tailored support will be labor intensive and may be difficult to effectively execute with planned staff
  - 14 staff in the Office of School Improvement
  - 6 regional student success specialists
  - 20 subject-matter specialists in reading, math, and science (who also provide support and technical assistance beyond school improvement)

# Effectiveness of the new program cannot yet be determined

- VDOE only recently began to provide support services under its redesigned program, so it cannot be evaluated
- A full evaluation of program effectiveness would entail
  - reviewing VDOE staff experience, qualifications, and workload
  - interviewing staff at schools receiving support
  - examining whether support is sufficiently evidence based
  - analyzing data to determine whether schools improved

---

## School improvement needs detailed planning, sustained attention, and evaluation

- VDOE needs to develop and submit a detailed plan for the newly redesigned school improvement program
- Providing improvement services to lower-performing schools is a critical state responsibility and requires
  - sustained effort for success
  - necessary and sustained attention from future superintendents of public instruction
  - regular monitoring by VBOE to ensure effectiveness of agency's school improvement efforts

# Recommendations

The General Assembly may wish to consider directing:

- VDOE to (i) develop and submit a one-time, detailed plan for the state's school improvement program and (ii) annually develop and submit a status report.
- the superintendent of public instruction to consistently and effectively implement the state's school improvement program.
- VBOE to ensure the state administers an effective school improvement program.

---

## Recommendations (cont.)

The Virginia Board of Education should contract with a qualified expert to conduct an independent evaluation of the state's school improvement program and use the evaluation results to direct changes to the program as needed.

---

## JLARC staff for this report

Justin Brown, Senior Associate Director

Joe McMahon, Project Leader

Kate Hopkins, Senior Legislative Analyst

Sam Lesemann, Senior Legislative Analyst

Laura White, Senior Associate Legislative Analyst