Virginia's K-12 accountability system: key JLARC findings and fixes

JLARC's December 2025 report identifies significant design problems that limit the credibility and usefulness of Virginia's new school accountability system. While the framework includes some useful changes from the previous system, JLARC finds that weighting, scoring rules, and unclear labels prevent the new system from accurately identifying school quality or consistently driving improvement. Most importantly, JLARC's analysis shows the system places very heavy weight on proficiency and relatively little weight on growth, causing summative scores to reflect student demographics more than a school's contribution to learning.

Key JLARC findings

Proficiency-heavy design amplifies demographics in results**:** JLARC's dominance analysis (Table G-4) shows the strongest drivers of summative scores are the share of economically disadvantaged students and school level (high school), followed by percentage of white students (Rank 3) and Black students (Rank 4). JLARC's modeling shows this weighting choice is consequential: equal weighting of proficiency and growth significantly reduces the correlation between scores and economic disadvantage and changes which schools fall in the bottom quartile (about 10 percent under equal weighting and 18 percent under reversed weighting). The report found, "nearly all experts interviewed by JLARC indicated that the SPSF's weighting on growth could be increased."

Growth is downweighted and cannot meaningfully offset proficiency**:** JLARC finds that Virginia places unusually high weight on proficiency relative to growth compared with other states - third highest in the country. As a result, growth plays too small a role to counterbalance proficiency-driven outcomes, making improvement difficult to see in ratings, especially for high-need schools.

Labels do not clearly communicate performance**:** JLARC finds that labels in the new system do not form a clear, continuous performance scale and can imply meanings the system does not measure, undermining public understanding and trust.

School-level comparisons are not fair**:** JLARC finds that elementary, middle, and high schools are measured using different indicators but scored against the same thresholds, a structure that advantages high schools. Two-thirds of high schools are labeled Distinguished, compared to 9 percent of middle schools and 17 percent of elementary schools.

Improvement expectations exceed current capacity**:** JLARC finds the redesigned improvement model is labor-intensive and raises concerns about staffing. Under current levels, each Office of School Improvement staff member would be responsible for about 30 schools (a 36 percent increase over prior workloads), while subject-matter specialists may each support 40–80 schools. JLARC warns that without additional staffing, it is unclear whether the state can deliver the depth of support required for high-need schools.

What JLARC indicates needs to change

- Rebalance proficiency and growth so summative scores better reflect school impact rather than student background
- Fix labels to clearly communicate performance on a coherent scale
- Establish school-level scoring thresholds to ensure fair comparisons
- Add guardrails to scoring so no single indicator dominates results
- Strengthen improvement capacity, including additional state-level staffing to support high-need schools
- Improve public reporting of accountability results by providing clearer indicator-level information, comparisons to similar schools, multi-year trends, and contextual data