This recommendations report was adopted by resolution of the State Board of Education on Thursday, November 15, 2018.

RECOMMENDATIONS FOR REFORMS TO OHIO'S DISTRICT AND SCHOOL REPORT CARDS

Recommendations report submitted to the State Board of Education of Ohio by the Accountability and Continuous Improvement Committee (Expanded).

BACKGROUND

The State Board of Education invited education stakeholders to participate in an expanded series of Accountability and Continuous Improvement Committee meetings, as noted in Ohio's Strategic Plan for Education, to address short-term (2017-18 Report Card) and long-term (next iteration of the Report Card) issues surrounding the Ohio School Report Cards. The group reviewed each element of the report card including the federal ESSA requirements, state Ohio Revised Code requirements, state board authority and previously identified issues and options.

The group recognizes the value of the Report Card as part of the statewide accountability system. At the same time, it shares a belief that the current version needs improvement by means of additional clarity and providing a more complete story for each district and school.

Report Cards are very high profile and generate much interest from stakeholders across the state. Many ongoing discussions are occurring regarding the purpose and future of Ohio School report cards. Multiple legislative proposals have been presented to the General Assembly including work by Representative Mike Duffey (R- Worthington) who has actively participated in the work of this committee. Other groups including the Buckeye Association of School Administrators (BASA), Ohio Association for Gifted Children and the Fordham Institute have made recommendations that informed the work of this committee.

The desired outcome of the group is to collaboratively work on improving the Report Card in order to better communicate the story of Ohio's schools and districts by making recommendations to the State Board of Education's Accountability and Continuous Improvement Committee. These recommendations could include Board actions through their direct authority and/or recommendations for future legislative change.

PURPOSES OF THE REPORT CARD

Ohio School Report Cards are designed to meet multiple purposes. The group has identified these as the most important:

Support the state's interest in gauging its education system's performance: The state has a legitimate interest in knowing how well its education system performs, and the extent to which the students in the system are being prepared for future success. District and school report cards help the state to identify excellence as well as underperformance. In the latter case, report cards identify districts and schools that need support with improvement efforts.

Advance equity: Ensuring equity in the education system is challenging. A well-designed accountability system can help shine light on inequities based on specific student characteristics – socio-economic status, race/ethnicity, disability, English language competency, etc.

Communicate to parents and the community: Report cards can provide communities with information related to certain aspects of the preparation of students for future success. It should answer key questions:

- Are students, generally, learning foundational skills and knowledge?
- Are subgroups of students learning foundational skills and knowledge?
- Is the school or district improving in its fundamental mission to educate students?

Support school and district improvement efforts: Report cards can drive discussions among local boards, teachers and administrators about the causes of underperformance and the strategies and actions that can lead to improvement. The data included demonstrates to educators, school administrators and families where their schools are succeeding as well as areas where they need to improve. The data provided by the report card system, combined with important local data, becomes the basis for a continuous improvement process to build on areas of success and identify targeted plans to address challenges. There are many examples across the state where report card data has stimulated actions to be taken to improve education.

Report Card purpose: Report cards are not meant to replace local data, but instead should complement local data sources. Report Cards are annual, summative snapshots of performance and are not meant to be formative. Report Card data, including the corresponding diagnostic information, should inform ongoing instructional decisions, but are not intended to be the primary source of information used during the school year to make adjustments to instructional activity. Report cards are not intended to be punitive even though some people may use them in this manner.

DESIGN PRINCIPLES

The group's work was guided by these design principles:

- **Fair:** Perhaps the most common complaint about report cards is whether they fairly portray the performance of the school or district. Report cards need to be fair.
- **Honest:** Report cards need to be able to honesty differentiate between schools and districts that are performing well and those that are not. They need to be an honest portrayal of what is happening.
- **Reliable and Valid**: Report cards should provide information that consistently measures the concepts intended to be measured.
- **Clear and Easy to Understand**: While the measures may be complex, the public facing communications should be clear, easy to understand, and simplified.

<u>**RECOMMENDATIONS**</u>

It is in that context that this list of recommendations regarding the state report card is presented.

ACHIEVEMENT

The Indicators Met measure within the Achievement Component has inherent weaknesses (such as not differentiating between schools that are close to meeting or far from meeting a target).

1) *Legislative recommendation*: Therefore, the Achievement component should rely solely on the performance index. The **Indicators Met measure should be eliminated as a graded measure**. Data about the percentage of students performing proficient or better on state assessments

should continue to be reported. For comparison purposes, reporting should also include similar districts and state level data.

K-3 LITERACY

The Committee has determined that the current K-3 Literacy component is misleading. Report card users think it is a measure literacy performance for all K-3 students when in fact it is a complicated portrayal of efforts to improve outcomes for struggling readers. Some schools may have a small number of students struggling with literacy, while the vast majority of students are succeeding – but the current measure only reflects the struggling students. Making sense of this measure is very challenging.

1) *Legislative recommendation*: It is recommended that the **K-3 Literacy measure be eliminated. If an early literacy measure continues to be included, it should be the Promotion Rate** which measures the percentage of students meeting literacy requirements to be promoted to the fourth grade. This should include comparisons to similar districts and the state average.

2) *Additional consideration*: If the current measure is maintained, it should be renamed to more accurately reflect its focus on struggling readers (NOTE: This consideration was implemented for the September 2018 Report Card and the measure was re-named "Improving At-Risk K-3 Readers"); and the label of "Not Rated" should be reconsidered for clarity.

PREPARED FOR SUCCESS

The committee believes the Prepared for Success measure has promise. Its current structure does not appropriately value different accomplishments. Its tiered structure adds confusion and makes debatable differentiations between various accomplishments. The group discussed several options to improve the Prepared for Success measure.

1) *Legislative recommendation*: The **Prepared for Success measure should be refined to include additional measures of college, career and life preparedness** (for example: military enlistment, ASVAB, CLEP, CTAG, career prep program credentials, Ohio Means Jobs Readiness Seal, etc.).

2) *Board Recommendation:* The Committee also recommends that the **dual tier structure of Prepared for Success be restructured into a single tier** that provides similar credit for all measures (for example, AP and College Credit Plus would have the same weight as remediation free status).

3) *Board Recommendation:* The above recommendations should apply to the Career Technical Planning District Report Card as well.

VALUE-ADDED

The Committee recognizes the importance of growth measures in understanding the progress of students and supports its use as an important equity consideration. At the same time, measuring growth is complex and Ohio's current system has many challenges including how the measure is

communicated, translated into a letter grade, and interrelated with other policies and systems (such as formative assessments).

The Committee identified initial themes (See Appendix A) and reconvened in October 2018 to further discuss these themes, as well as recommendations made by the Value-Added Technical Advisory Group *(See Appendix B).*

1) *Board Recommendation*: Include the Value-Added Technical Advisory Group report (with legislative recommendations on technical details such as number of years of data, interpretation of the gain index, and revising the grade scale) as an appendix to the Committee's final report to be included in potential State Board legislative recommendations (See Appendix B).

2) *Legislative Recommendation:* The committee recommends the use of one-year value-added for accountability purposes while reporting multi-year value-added for the additional benefit of viewing larger trends.

3) *Legislative Recommendation:* The committee recommends the elimination of the subgroup demotion.

4) *Legislative Recommendation:* The committee recommends the elimination of the Value-Added Rankings required in Ohio Revised Code 3302.21(A)(2).

A-F LETTER GRADES

The Committee spent much time discussing the A-F letter grade system, which is the current system of meaningful differentiation of school and district performance required by state law and used to meet federal ESSA requirements.

1) *Legislative Recommendation*: The committee recommends **eliminating all A-F letter grades** for the entire report card; and adopting an ESSA-compliant dashboard while still maintaining high expectations and aspirational goals.

DESIGN and COMMUNICATIONS

The committee extensively considered how the "report card" is presented. To some, the report card is the landing page (first screen) that appears on a computer screen when a school or district is selected on the Department's report card web page. Others consider the report card to include all pages of the report card PDF – in many cases in excess of 30 pages. Ultimately users need to be able to access both

high level information as well as the background detail. However, the most important consideration is what appears on the first page. In all actions taken to improve the report card, the goal is for the first page to provide clarity of content and be understandable to parents, caregivers, and the community.

1) *Department recommendations:* The design could be improved by:
   - Adding more descriptive narrative on the purpose of the report card to the landing page (i.e. homepage) (NOTE: the narrative was updated on the landing page for the September 2018 release based on these recommendations);
   - Reviewing language to improve clarity; and ensure clear definitions and descriptions of measures are accessible up front;
   - Relocating the "District Profile" link to the Report Card overview for increased prominence (NOTE: The District Profile link was moved to the Report Card overview for the September 2018 release based on these recommendations);
   - Adding additional clarifying language regarding the graduation rate cohorts (NOTE: the report card was updated to explain which class the graduation rate is measuring and provides the year in which students started high school and the graduation year).

ADDITIONAL COMMITTEE CONSIDERATIONS

1) **Additional Profile information**. The committee expressed great interest for the report card to include additional profile information that focuses on supporting our students in alignment with Each Child, Our Future – Ohio's Strategic Plan for Education. This information would help tell a more complete story about Ohio's schools (e.g. AP courses offered, Art courses, additional teacher information, etc.)

2) **National Comparisons/Assessments.** While there was not consensus, the committee expressed interest in national comparisons of student performance (e.g. national assessments) while continually looking at ways to reduce testing.

We, the members of the Accountability and Continuous Improvement Report Card Workgroup, appreciate the opportunity to be part of this process to make a meaningful contribution to addressing the present challenges of the Ohio School Report Card.

COMMITTEE MEMBERS
Nancy Hollister, **Chair**
Cathye Flory, **Vice Chair**
Lisa Woods
Pat Bruns
Laura Kohler
Antoinette Miranda
Eric Poklar
Charles Froehlich

EXTERNAL MEMBERS
Randy Smith, OSBA
Stephanie Starcher, BASA
Scott Emery, OAESA
Tyler Keener, OASSA
Margie Toy Ma, OPTA
Donna O'Connor, OEA
Brad Dillman, OFT
Jamey Palma, Career Tech
Jan Osborn, ESC

While clear recommendations have not yet emerged, several key themes have been identified for future discussion when the Committee reconvenes.

1) **Testing structure**. The Committee understands that the Value-Added system is exclusively dependent on the underlying assessments used. The Committee discussed the differences in intent and practice of formative assessment systems (such as MAP and STAR) and state assessments. In many cases, formative systems provide useful information that the current state system is not intended or designed to provide. At the same time, multiple testing structures lead to concerns about over-testing and incoherent feedback from the data. The committee is interested in exploring innovative approaches to formative assessments or state testing that may address these concerns. This could include working with formative assessment vendors to address state concerns on issues such as alignment with state standards and, in particular, the depth of knowledge required to meet state standards.

2) **Formally studying the relationship between state and vendor test results**. A related point is that state data and formative vendor data do not always produce consistent results, even though they are both supposedly aligned to state standards. The committee discussed possible reasons for this (breadth and depth, above grade level testing, etc.). However, it would be beneficial to more formally study and understand these relationships.

3) **Distribution of results**. While the committee discussed a general preference to eliminate all A-F letter grades (including Value-Added), concerns were also raised about the distribution of letter grades in the current system. Specifically, there are concerns regarding the "W" shaped distribution of results for Value-Added, that is, significant numbers of A's and F's, very few B's and D's, and a moderate amount of C's. This issue was also raised during ESSA stakeholder feedback and reiterated by staff. This phenomenon is solely a function of where/how the letter grade cut lines are established – a policy that is prescribed in state law, but for which recommendations to adjust could be made.

4) **Number of years of data**. A related point, and one that had been raised during ESSA stakeholder engagement (particularly from urban districts) is the statutorily required use of three years of data. The Value-Added grade is essentially a three-year average, which means that results from previous years influence current and future grades. Districts with poor results a few years ago are still connected to those results even if improvements have since occurred. This three-year approach was implemented to add more stability to the measure, but conversely means the measure is not necessarily reflective of the most recent year.

5) **Relative weight of growth measure.** Many measures, especially achievement measures, are correlated with socio-economic status. All students, regardless of their starting point, can show growth in Ohio's system and the Value-Added measures are designed to measure that growth – which is an important tool with which to evaluate the equity of educational outcomes. Many stakeholders have suggested increasing the relative weight of growth measures. Currently, it is equal to achievement (by state law), and 20% of the overall grade (by administrative rule).

6) **Technical fixes.** There are some technical options that could be considered including the following:

   a. How to communicate grades (ratings) when a school's achievement improves, but does not meet growth expectations.
   b. The current subgroup demotion when calculating the component grade. In state law, schools cannot receive an "A" for the Progress Component if any of the subgroup grades are lower than a "B".
   c. The interpretation of the Value-Added gain index, which is currently based on growth *and* a measure of statistical strength.
   d. The availability of a predictive model to support the system properly accounting for gifted students (e.g. how do middle school students count when they accelerate over a grade into Algebra I?) and assisting with acceleration decisions.

7) **Communications**. Measuring growth is inherently complex and there are known challenges to effectively communicating Value-Added measures. These range from branding, to interpretation, to understanding the formula. The communication challenges vary between different audiences – how value-added should be communicated to parents is different than how it should be communicated to Building Leadership Teams (BLTs).

8) **Training and Professional Learning**. Emphasis should also be placed on education and training on Value-Added data and measures. This could build on the current structure of Regional Data Leads (RDLs).

The Value-Added Advisory (VA Advisory) consists of a broad range of experts with both technical and policy capacity to provide the State Board of Education and the Department meaningful feedback on the topic of value-added measures and data. The group currently includes school and district practitioners (superintendents, principals, curriculum and assessment personnel) from a wide variety of settings (large urban, rural, suburban, community schools), higher education partners and Educational Service Center partners (data and school improvement personnel) serving as Regional Data Leads.

As the State Board of Education directed the Accountability and Continuous Improvement Committee to review the report card and provide recommendations to the Board, the Value-Added Advisory has taken this opportunity to address several of the themes identified by the committee and provide recommendations and/or considerations for the committee. Most of the themes identified by the stakeholder workgroup are issues that the VA Advisory has been considering over the past few years.

The following considerations and recommendations are organized according to the themes identified by the Accountability and Continuous Improvement's Extended Report Card Stakeholder workgroup initial report that was presented to the State Board of Education at the July 2018 business meeting.

OVERARCHING CONSIDERATIONS

1. **Letter grades.** The VA Advisory members agree and acknowledge that letter grade assignments complicate the interpretation of the data.
2. **Timing of data**. There is great interest for the data to be provided as soon as possible to make it more actionable and useful for educators.
   - *Recommendations*:
     i. **Publish one-year data.** The one-year data is valuable and should be easily accessible. (*Note:* Other considerations focus on using the one-year as the basis of the determinations on the report card).
     ii. **Advanced reports.** More detailed reports, including drill-down information on one, two, and three years of data across grades and subjects should be made available in the advanced reports – ideally at the same time as release of the report card.
     iii. **Consider EVAAS changes.** The SAS EVAAS® system is the main portal for accessing detailed diagnostic information. However, it is usually not available until early October due to logistical considerations. The Department should work with SAS to prioritize ways to make data available as soon as possible.
3. **Need for professional development**. The group emphasized the need for large scale professional learning opportunities to increase the understanding and use of data and tools

that provide meaningful feedback to educators to help students. This includes resources to enhance the use of existing tools such as EVAAS®.

- *Recommendation:* Advocate for state budget resources for statewide supports for professional learning opportunities that build on the current Regional Data Lead Network to ensure that all educators have access to supports and resources.

NUMBER OF YEARS OF DATA (ACI REPORT THEME #4)

***Theme identified by the ACI Report Card Workgroup:** A related point, and one that had been raised during ESSA stakeholder engagement (particularly from urban districts) is the statutorily required use of three years of data. The Value-Added grade is essentially a three-year average, which means that results from previous years influence current and future grades. Districts with poor results a few years ago are still connected to those results even if improvements have since occurred. This three-year approach was implemented to add more stability to the measure, but conversely means the measure is not necessarily reflective of the most recent year.*

The advisory members discussed at length the impact of using up-to three years of data versus using a single year of data in the graded value-added measures. Using three years of data creates more stability in the measure, while using one year of data is more responsive to the progress happening year to year. One-year results are much more statistically aligned to one-year changes in performance and are more consistent with accountability structures which tend to look at most recent performance. Additionally, one-year results are proximally aligned with the most current efforts of schools.

Recommendations regarding years of data included in the measure must be made to the General Assembly as Ohio Revised Code 3302.03 (C)(1)(e) states the Department shall use up to three years of value-added data as available.

While the advisory members did not arrive at consensus with a single recommendation, several considerations did emerge.

One-year data. Using one year of data in the value-added calculation is responsive to measuring progress and activities within districts and schools and decreases the complexity of interpreting and communicating results. However, the trade-off is that one-year measures are inherently less stable (results may move from positive to negative or vice versa). This is not necessarily a weakness as the measure is designed to be sensitive to changes in growth. However, it does create a communications challenge. In any given year, there are schools and districts that move from an "A" to an "F" or "F" to an "A". Special note should be made that during tests transitions, there are considerably more of these shifts.

Consideration: Consider making a legislative recommendation to transition to a one-year measure for accountability purposes. Special consideration might be needed if Ohio were to adopt major changes in its testing program.

Recommendation: If transitioning to a one-year measure for accountability purposes, the VA Advisory recommends reporting the three-year data as the trend data is valuable.

Up to three years of data. If the graded measures and Progress Component continue to use up to three years of data as available, there may some additional actions to address previously identified concerns:

Consideration: **'Higher Of' option.** Consider making a legislative recommendation to allow for the "higher of" the three year or one-year measure to determine the accountability determination. For example, if the three-year data translates to a "D", but the most recent one-year of data shows great improvement and is an "A", then the grade would be an "A". This would allow for the stability of the three-year measure, but still recognize important recent improvements. This idea was viewed by some members as problematic and could be interpreted as using inconsistent measures that are difficult to compare.

Consideration: Weight years within the measure. A weighting structure could be applied that would place stronger emphasis on the most recent year of data to include some responsiveness while still maintaining the stability of using multiple years. For example – when combining the three years of data, Year 1 (3 years ago) could be weighted at 20 percent, and Year 2 (2 years ago) and Year 3 (most recent year) could be weighted at 40 percent. This would give more weight to the recent year but would add considerably more complication to the measure, and potentially appear arbitrary.

- *Recommendation.* If maintaining the three-year data in the measure for accountability purposes, the VA Advisory recommends reporting the one-year data as the responsive information it provides is valuable.

INTRPRETATION OF THE GAIN INDEX (ACI REPORT THEME #6)

Theme identified by the ACI Report Card Workgroup: There are some technical options that could be considered including the following: (c.) The interpretation of the Value-Added gain index, which is currently based on growth and a measure of statistical strength.

One of the known challenges of the value-added measure is that the resulting output is a "Gain Index" which is a measure of statistical strength rather than an amount of growth. An Index of 5 means that there is more statistical confidence that growth expectations were exceeded than an index of 4. It does not necessarily mean that a 5 has more growth than 4, even though it is commonly expressed and misunderstood that way.

This could be addressed by factoring in "effect size" in the index. This commonly used statistical measure would provide a correct interpretation of more/less growth. While it could be viewed as adding a level of complexity, it would allow the measure to be more easily interpreted and align with how many users already interpret it.

- *Recommendation*: Refine the business rules in determining the gain index to account for effect size in conjunction with related changes to grade scale.

DISTRIBUTION OF RESULTS (ACI REPORT THEME #3)

> ***Theme identified by the ACI Report Card Workgroup:*** *While the committee discussed a general preference to eliminate all A-F letter grades (including Value-Added), concerns were also raised about the distribution of letter grades in the current system. Specifically, there are concerns regarding the "W" shaped distribution of results for Value-Added, that is, significant numbers of A's and F's, very few B's and D's, and a moderate amount of C's. This issue was also raised during ESSA stakeholder feedback and reiterated by staff. This phenomenon is solely a function of where/how the letter grade cut lines are established – a policy that is prescribed in state law, but for which recommendations to adjust could be made.*

One of the concerns stakeholders have raised regarding Value-Added is the distribution of results, which often look like a "W" shape – with many As and Fs and Cs, but few Bs and Ds. This distribution is further driven by three years of data which has the tendency to move letter grades towards the A or F range (more data leads to more statistical certainty). These results are a function of statutory definitions that specify how a value-added measure translates into a grade, not an inherent weakness in the methodology. The VA Advisory group has spent a significant amount of time examining the distribution of results over the past few years.

The Committee is opposed to a pre-determined distribution of results (i.e. forced amount of each letter grade). There is a perception that exists of 'winners' and 'losers' – because of this perception, it is important that this conversation is framed with a focus on meaningful rating assignments and not about a preset distribution of results.

The letter grade assignments based on the gain index are established in Ohio Revised Code 3302.03 (A)(1)(e). The VA Advisory members reviewed several frameworks for grading that could better communicate schools and districts meeting growth expectations. Any changes to this framework would need to be considered in light of the previous recommendations.

- *Legislative Recommendation*: The Revised Code could be changed to provide flexibility in the determination of the grade scale which is currently prescribed in state law.
  - i. Any such changes should be considered in the context of the other recommendations in this document (such as the one-year vs three-year of data)

ii. The workgroup raised some concerns regarding the relative nature of the distribution (so in theory, every schools could meet expectations, but not every school could get an "A).

iii. The committee also discussed the interpretation of the "C" grade which meets expectations.

iv. Any changes to the grade scale could address these concerns. (for example, a statutory change could give authority to the State Board to determine an updated grade scale in administrative code).

TECHNICAL FIXES (ACI REPORT THEME #6)

__Theme identified by the ACI Report Card Workgroup:__ There are some technical options that could be considered including the following: a.) How to communicate grades (ratings) when a school's achievement improves but does not meet growth expectations; b.) The current subgroup demotion when calculating the component grade. In state law, schools cannot receive an "A" for the Progress Component if any of the subgroup grades are lower than a "B"; c.) The interpretation of the Value-Added gain index, which is currently based on growth and a measure of statistical strength; and d.) The availability of a predictive model to support the system properly accounting for gifted students (e.g. how do middle school students count when they accelerate over a grade into Algebra I?) and assisting with acceleration decisions.

Component Demotion. Regarding the current subgroup demotion, law states that schools cannot receive an "A" for the Progress Component if any of the subgroup grades are lower than a "B". This statutory requirement was an equity provision to ensure that subgroups are being addressed. For example, a school could be doing very well overall but not meeting growth expectations for students with disabilities. The challenge is that a "C" in value-added means that growth expectations are being met. A school could have an "A" Overall, "A" with gifted students, "A" with the Lowest 20%, and a "C" with students with disabilities. This should result in an "A" for the component grade, but the demotion means the component is a "B" even though growth expectations were met.

- *Legislative Recommendation*: The advisory group acknowledges the intention of the demotion but finds it counterintuitive to the fact that a "C" is meeting expectation, and therefore recommends eliminating the subgroup demotion. **If not eliminated, any demotion should be limited to application based only on subgroup grades of "D" or "F".**

Value-Added Gain Index Rankings. The current measure is based on growth and statistical strength, and therefore when ranked numerically, does not provide valuable information or the correct interpretation of the data. State laws requires a separate reporting of value-added rankings which are often mischaracterized.

- *Legislative Recommendation*: The advisory group recommends eliminating the Value-Added Rankings required in Ohio Revised Code 3302.21(A)(2).

TRAINING AND PROFESSIONAL LEARNING (ACI REPORT THEME #8)

Statewide Opportunities. There is a strong desire by the VA Advisory group, and a need across the state, to scale professional development on data literacy and the use of value-added data.

- *Recommendation*: Advocate for state budget resources for statewide supports for professional learning opportunities that build on the current Regional Data Lead Network to ensure that all educators have access to supports and resources.