

(Note that if there is only limited time to testify, I will not read the sections in [brackets] below.)

Chairman Brenner, Vice Chair Slaby, members of the House Education Committee, thank you for the opportunity to speak today in support of House Bill 591. My name is Dr. Kara Morgan and I am an at-large governor-appointed member of the State Board of Education.

I want to be clear that I am not testifying on behalf of the board, but rather as an individual and a professional with twenty years of experience in and around the federal government in quantitative analysis, evaluation, and performance management.

[You may be aware that the State Superintendent, Paolo DeMaria, formed a workgroup under the Accountability and Continuous Improvement committee of the board this spring to review the report card and make recommendations for improvement. You may also be aware that last month, at the November meeting, the State Board approved advancing the recommendations from the workgroup committee to the legislature. ]

The first point I wanted to make is that one important recommendation from the Board is that Ohio define the purpose of the report card. As work is done to revise the report card, it will be important to have a shared understanding of the purpose so that the end-product is cohesive and consistent with that purpose.

[The workgroup identified three core purposes for state report cards, which I agree with:

- (1) The state report card should provide performance management information for the state, districts, and schools;
- (2) The state report card should provide visibility into equity of educational opportunity across the state; and
- (3) The state report card should communicate to parents and communities about school and district effectiveness in educating children.]

My second point, which is not part of the workgroup recommendations or HB 591, is that a strong continuous improvement process should be implemented to include on-going user-focused studies with each of the audiences identified in the purpose statement. This is necessary to ensure that the report card meets their needs now and continues to meet their needs over time. Findings from these studies should inform future revisions and updates to the report card.

[Now, with these three purposes in mind, there are a few differences between the workgroup report's recommendation and HB 591, and I'm going to focus today on the one that I believe is most significant.]

The main point of my testimony is on how to report results from standardized tests. As you know, under current law, the component of the report card that reports passage rates for standardized tests is called "Achievement," and there is also a component that uses the standardized test scores a different way, called "Progress" and also referred to as "Value-added."

[The Achievement component includes a section on "performance index" and a section on "indicators met." Both HB 591 and the workgroup recommend changes to the "Achievement" component. HB 591 renames as it "Standardized Testing" with the idea that this is more intuitive in terms of public understanding. Whether this is understood well can be tested with the on-going user studies that I mentioned earlier.

Both HB 591 and the workgroup recommend taking out the “indicators met” section as a graded component. Its formulation violates a basic principle of good performance measures by using an arbitrary level to separate “acceptable” from “unacceptable,” so I agree with this recommendation.]

Where the workgroup and HB 591 differ is whether to keep the “Performance Index” part of the Achievement component. I agree with HB 591 that Performance Index does not provide meaningful information about the performance of the school or of the teacher, and therefore is not useful as a performance measure or a tool to evaluate equity.

Let me explain. When evaluating the impact of a program (like a year of education), as an evaluator, the most important piece of information to start with is a baseline. A baseline is where the person is starting from on that measure at the beginning of the program. Most of my work has been in the public health sector, so I’m going to use an example from there. Let’s say, for example, I am evaluating a program to reduce obesity rates. I would start with a baseline measure of obesity in each person who is starting the program – and I will use BMI or Body Mass Index as a measure of obesity. Then, the program would begin – maybe it contains training on nutrition, exercise classes, use of a fitness tracker, counseling sessions, work-out buddies – any types of interventions that have been shown to impact obesity might be included. Let’s say it’s a 6-month program. At the end of 6 months, I would measure BMI again. I would say the program was effective if the difference between the beginning and the end showed a decrease in average BMI.

[To take this example further, say you had two programs to reduce obesity, one in an YMCA and one in a health clinic, and you want to compare these two programs to see which one is better in terms of reducing BMI. At the end of a 6-month program, if you are showing data like that which is currently used for the performance index component of the report card, you would show a graph like the one in Figure 1.

Average BMI after 6-month program

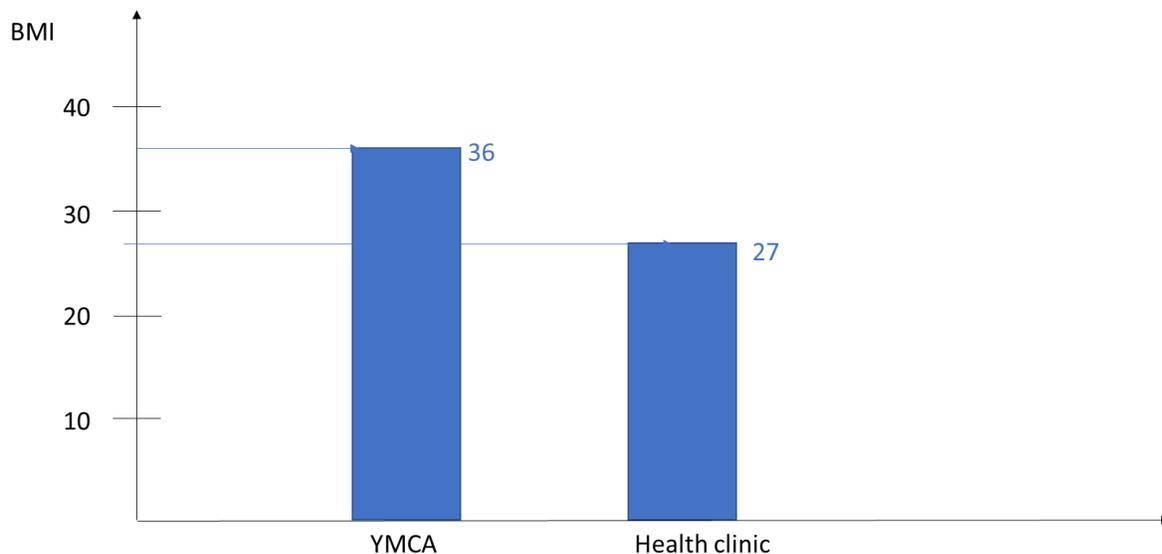
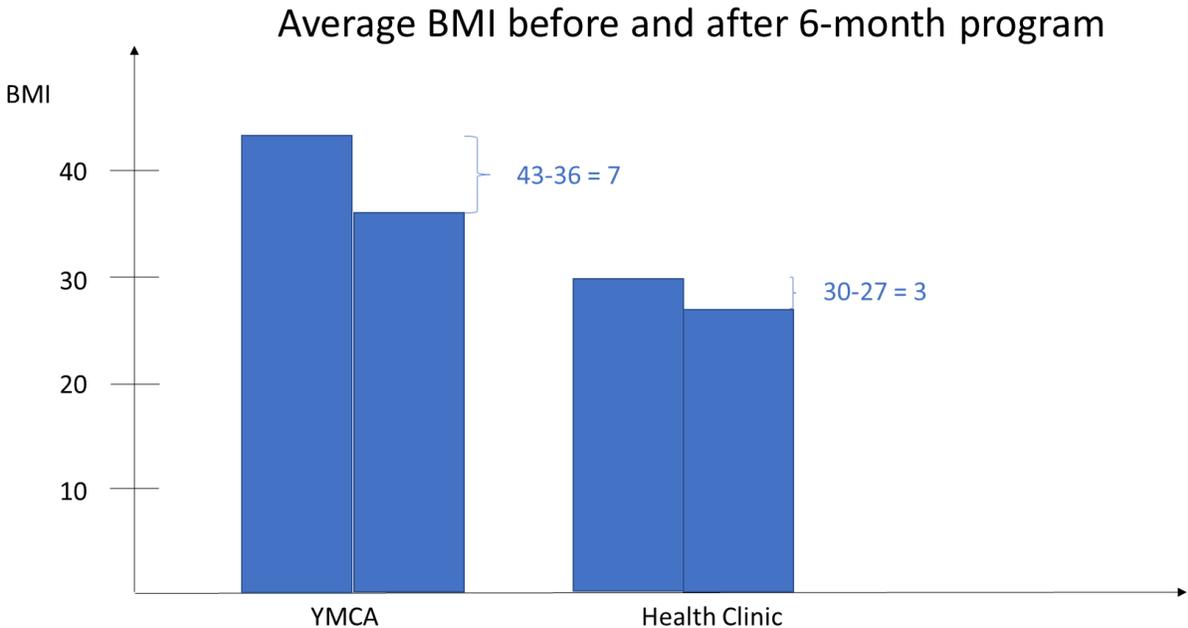


Figure 1. Average BMI after 6 month program

In this graph, it looks like the health clinic program is better because the ending BMI is lower.



However, if you include the starting point for each program in terms of the BMI of the participants, the story looks different. Figure 2 shows the before and after BMI for each program. It shows that the YMCA program started off with a higher BMI and had a greater decrease over the 6 months. It shows that the change in BMI over the six months was 7 for the YMCA program and only 3 for the Health Clinic program. So this shows that the health clinic program was more effective. The point is, you cannot tell which program is more effective by just looking at the measurement at the end. You also need to consider the starting point.]

The current report card's chief flaw, addressed in HB 591, is that most of its components – all except for the Progress measure - base the scores and corresponding letter grades primarily or exclusively on the final result, without any recognition of the starting point.

[We know from research that achievement gaps exist based on socio-economic factors and that these factors are not randomly distributed across schools and districts. Since students come in with different levels and abilities and that baseline is not accounted for in the measure, the performance index measure is not a meaningful measure of school or teacher performance. Rather, it is more a reflection of the skills and abilities that the students are bringing into the school.

HB 591's choice of converting "Achievement" into a "Standardized Testing" component and simply reporting the raw scores is a way to keep that component but not use it as a comparator. This allows useful insights about how students are performing, trends in the district, trends in comparison to other districts, etc. ]

As you know, the "Progress" component of the report card DOES take into account a baseline, and measures the improvement over the year. This is a meaningful performance measure, although with flaws that do need to be addressed, as laid out in the bill. Primarily, the flaws are that as the measure is currently formulated, it is difficult for teachers, principals, and superintendents to understand and

explain, which means it is not being used as effectively it could be as a performance measure to drive continuous improvement. With improvements, it would be an effective performance measure.

Value-Add in general, as utilized nationally in other states and in Ohio, provides schools with useful information about student growth and therefore is a valuable tool for understanding what areas students are improving on, and which may need more attention at the school.

When compared to Achievement, value-add provides a closer-to-equal opportunity for schools to demonstrate their impact without being confounded by factors that are known to impact performance level. It also does not provide a biased advantage to those districts who, by way of demographics, have higher achieving students, as the performance index currently does. Instead, it measures whether all students show progress year over year, and therefore helps incentivize improvement across the state.

Both the workgroup report and HB 591 recommend revising “Progress” (that is, the Value-Add measure) to make it more transparent, informative and useful to teachers and schools. I strongly agree with this recommendation.

In sum, moving away from grading schools on demographic-correlated outcomes without baseline comparison is one step towards a system that will be more fair, meaningful and capable of driving the behavioral changes desired.]

In summary, my testimony is intended to make three main points:

- First, the legislature should strongly consider defining the purpose of the report card, to ensure everyone is on the same page and working toward the same goal;
- Second, it will be important to implement a continuous improvement model to ensure that the purpose is being met for all the identified audiences.
- Finally, using test scores to evaluate how well districts or schools are performing without taking into account how well the students were performing already is inconsistent with evaluation practice and is not useful as a performance measure. Instead, change over the year (also called “growth” or “progress”) which is currently captured in the value-added measure, does provide useful feedback. So, the focus should be on that type of measure, with improvements made in the calculations and presentation of the measure results to increase the understandability and usability of the analysis for schools and teachers.

Thank you for allowing me to testify today. I would be happy to answer questions.