



RESPONSE TO REQUEST FOR INFORMATION

**Arizona School Accountability
Components & System
RFI #ADED16-0002**

Submitted to
Arizona Department of Education

Submitted on the date of
July 15, 2016



temboinc.com • Philadelphia, PA • Portland, OR

SECTION 0

Table of Contents



Section 1

Demonstration of Values	3
1.1 Our Guiding Principles	3
1.2 Our Guiding Questions	3
1.3 Our Approach	5

Section 2

Our Organization & Background	6
2.1 Our Vision	6
2.2 Our Clients	6
2.3 Our Accountability Product Suite	7
2.4 Our Team	9

Section 3

System Design	10
3.1 Metric Selection	10
3.2 Alignment to College /Career Expectations	10
3.3 Differentiated Weights and Metrics	10

Section 4

Measuring Student Growth	11
4.1 Measuring Student Growth	11
4.2 Experience	11
4.3 Expediting Growth Scores	11
4.4 Interpretation	12





Demonstration of Values

1.1 Our Guiding Principles

There is a common misconception in public education that a high stakes accountability system should define what makes a particular school “excellent” or “failing”. This is not the case.

Instead the definition of educational excellence begins with the administration. Before delving into details of school accountability, state education agencies (SEAs) must have clear policy goals, both for their school leaders and for the students in their stead. The administration’s vision for progress and improvement in public education sets the standard; an accountability system serves to measure schools against that standard. The administration’s policy goals are the compass that guides each and every decision in the development of a fair and transparent accountability system.

Put another way, there is no one-size fits all accountability solution. Rather, accountability systems must be tightly aligned to the administration’s larger vision for academic success. Without alignment to these higher order goals, any accountability system, no matter how sophisticated or statistically sound, will ultimately be ineffective. This is because accountability systems are incentive systems. They communicate clear incentives to district and school leaders on what to value, e.g. improve your proficiency rate on this assessment; improve your 4-year graduation rate for these students. If the system is not tied to policy goals, then the incentives of the system will diverge from the incentives of the administration, leading to conflict, confusion, and inefficiency.

When accountability systems are aligned to broader SEA policy goals, the systems instantly become more fair and more transparent. School and district leaders feel that the system is fair because the system is consistent with the expectations that have been set from above. Similarly, the public has more transparency into the system because it reinforces the state-level goals, rather than introducing new information and new incentives.

1.2 Our Guiding Questions

Using this principle of alignment as a guide, there are several key decisions required to build an accountability system that is robust, coherent, and effective in improving school quality. In particular, when considering specific measures for inclusion, one should ask:

Is this measure meaningful?

In Tembo’s experience, whenever an educational agency begins developing a new accountability framework, there is no shortage of stakeholders arguing for the inclusion of this measure or that. The first criteria any measure must pass is simply: *Is this measure meaningful?* Does it align with the overall policy goals? Is there any empirical basis for believing that the measure will help achieve those policy goals?



Does this measure create the right incentives?

As mentioned, an accountability system is a system of incentives. If a measure is included in the system, school leaders have an incentive to improve their performance on that measure. Unfortunately, there is often a difference between the letter of the law and the spirit of the law. Take, for instance, measures of school climate and safety. The designers of accountability systems may wish to include a meaningful measure of school safety like suspension or expulsion rates, believing that schools with lower rates have better climates. However, once suspension and expulsion rates are included in the system, school leaders have an incentive to underreport actual incidents. The actual quality of school climate would remain unchanged (or perhaps even decrease), yet accountability scores would increase.

Do we have the data to report this measure?

Even when measures incentivize the right behavior, accountability system builders must ensure that those measures are supported by high quality data. The quality of the system is directly proportional to the quality of the underlying data. Variability in data quality and completeness between schools weakens the overall system by casting doubt on the reliability and validity of reported scores. A fair and transparent system will only include measures that are widely collected, using a controlled collection system, ideally with consistency year over year.

How do we differentiate high and low performers on this measure?

If a school has a 4-year graduation rate of 94.3%, is that good, bad, or somewhere in between? How do you know? Once we decide to include a measure in the accountability system, we must still decide how to translate a school's value into a measure of quality. Generally speaking, there are two ways to do this: 1) by relative comparison to peers, and 2) by absolute comparison to a pre-defined benchmark. The choice between relative and absolute comparison is one that should be carefully considered, especially with respect to the system's alignment with the higher policy goals of the administration.

What is the relative importance of this measure over any other?

Once a handful of measures have been selected for inclusion in the system, we must decide the relative importance of each. Again, this is a time where alignment to pre-existing policy goals should drive the conversation. Which is more important as a measure of college and career readiness: 4-year graduation rate, 5-year graduation rate, or performance on SAT/ACT? Should we value academic growth more highly, less highly, or equally to academic performance? Answers to these questions depend on a deep knowledge of the existing educational environment, as well as clear guidelines from the administration regarding their goals for public education outcomes state-wide.

Is this measure understandable?

Finally, when constructing an accountability system, system builders should continually ask themselves whether the measures they include make sense to the stakeholders involved. Externally, this often means balancing statistical rigor with public understanding and public perception. Building the most "fair" system may require a



trade-off from building the most “transparent” system, as more intricate measures built on elaborate statistical models are often more precise instruments, but may appear more opaque to external audiences. Internally, system builders can improve transparency and understanding through detailed and precise documentation of every step in calculating overall accountability scores. This could include publishing raw data sources, detailing the inclusion and exclusion rules for calculation, and providing detailed guides for calculation to the schools, districts, and the public.

1.3 Our Approach

Traditionally, when Tembo begins an accountability-focused engagement with a large district or state, our team plans and facilitates a series of discovery meetings to articulate the fundamental requirements for the system and reporting needs for various audiences and stakeholders. We have included outlines from some of these meetings below as a way to illustrate our approach:

Discovery Meeting A: Defining school quality

- What work is underway to articulate school quality and to what standards are schools/operators held?
- What is the purpose of the existing reporting methods (dashboards, report cards), and how are it intended to drive program accountability and/or continuous improvement?
- Which audiences will use the existing reporting methods (dashboards, report cards) and for what varying purposes?

Discovery Meeting B: Refining performance metrics to communicate program quality

- Which data sources are currently used to measure and communicate quality?
- How do these measures perform in terms of transparency, precision, and comprehensiveness (i.e., are they clear, well-measured and generalizable across schools/operators)?
- What additional data could be used to measure activities and outcomes (e.g. additional metrics in future years)? To what extent are these proposals transparent, precise, and comprehensive?

Discovery Meeting C: Operationalizing metrics

- Which current and proposed metrics rate highly in terms of transparency, precision, and comprehensiveness?
- How should these metrics be operationalized through data?
- What is each metric’s working definition, student/teacher universe, exclusion rules and calculation criteria?
- For any model-based metrics, what is the appropriate estimation strategy?



Our Organization & Background



2.1 Our Vision

Tembo was founded as a small business in 2010 in response to a need for better data management, analysis, and visualization services in the education sector. Tembo has since become a trusted partner and analytic leader in K-12 education, and provides a wide range of accountability and assessment-related data services to state departments of education, school districts, charter management organizations, schools, and philanthropists. Our work primarily focuses on the development of thoughtful accountability systems and reporting tools for those systems, with a goal to support SEAs and LEAs as they:



Define quality

Tembo applies smart science & best practices to develop good metrics of school & program quality



Communicate quality

Tembo designs & builds reports & interactive tools that provide clear, transparent information to school leadership, educators, and families



Drive quality

Tembo creates innovative evaluations & diagnostic analyses that drive conversations around quality & equity

Tembo believes that the combination of smart analysis and intuitive design can transform numbers into accessible, actionable information for educators, administrators, and the public. Our team years of experience with education data reinforce that thoughtful analyses and engaging reporting can drive more informed decisions. Tembo is unique in its broad expertise with educational data and deep concern for how results are communicated and used.

2.2 Our Clients

Our team has worked closely with a number of states and large urban districts in the design of accountability systems and tools to report out accountability results. These clients include, most recently, the Delaware State Department of Education, the Tennessee State Board of Education, the Washington DC State Superintendent of Education, the School District of Philadelphia, the Washington DC Public Charter School Board, New York City Department of Education, Sacramento City Unified School District, and Shelby County Schools, among others.



2.3 Our Accountability Product Suite

In the course of our extensive experience working to define school quality with states and local education agencies, we have created a series of tools that we are now releasing as a configurable product suite. These products support accountability system design, development, and reporting from end-to-end.



BUILD

A tool for creating a custom performance framework based on what matters most.

The Tembo Builder is a flexible online tool that allows state users to upload school-level data and create a custom performance rubric. The tool guides the user through key decisions like organizing metrics within domains; deciding on detailed scoring decisions (e.g. weights, floors, and targets); and handling missing data. As you build your framework, you can preview results in real-time and even compare the results of alternative frameworks to better understand trade-offs between each.



FORECAST

A user-friendly dashboard that enables district & school leaders to understand their results and plan for improvement.

Good accountability frameworks help school leaders to narrow their focus and work towards concrete and measurable outcomes. Unfortunately, accountability frameworks are often indecipherable in complexity, leaving school leaders unsure what specific changes they can make to improve their scores. The Tembo Forecaster is a secure online tool that allows schools leaders to connect the dots between specific measures and their overall accountability score. For example, using the Forecaster, a school leader could easily see the impact of a 5% increase on attendance in their overall accountability rank. This level of detail can help drive informed strategic planning decisions and improve alignment to the accountability system as a whole.





FIND

An intuitive way for families to search for schools based on what they value most

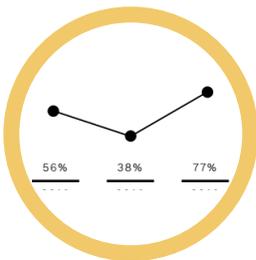
The Tembo Finder is a configurable web tool that allows families to explore schools in their area based on their personal preferences. Using the tool, parents can sort and filter schools by performance, programmatic offerings, and/or location, exploring individual schools in more detail via our School Reports. Tembo's goal is to provide the community with a clear, accessible way to find information about the schools available to them – free of complicated controls, advertisements, and unnecessary clutter.



EXPLORE

School and District Reports that detail performance and personality in a clean, consistent way

Individual school and district profiles, scorecards, and reports equip families with valuable information about school and district quality. However, existing reports are often one-size fits all. With Tembo Explore, state and local education agencies can flexibly configure personalized public reports based on their specific needs. Once configured, these these tools allow families and school leaders to explore details about schools' performance and programs, as well as how schools stack up against one another in an accessible and transparent way.



ANALYZE

A portfolio-level tool that allows analysts to answer key questions and chart broad trends

Our team created the Tembo Analyzer to provide state and district leaders with the power to quickly answer ad hoc questions about their schools. Are high performing schools also demonstrating high growth? How does this relationship differ between charter schools and traditional public schools? With Tembo Analyzer, users can quickly plot any two metrics on a scatterplot, filters results, compare groups, and view trend lines. The result is a quick, one-stop tool for charting key performance data, noting trends, and creating visuals for dissemination, all with a few clicks.



2.4 Our Team

Our proposed project team has an extensive portfolio of experience providing thoughtful guidance in the development of school accountability systems, and the creation of tools to engage portfolio managers, analysts, school leaders, and the public in the system results. Below, we provide biographical details about our proposed team's leadership; these individuals would serve an active role sponsoring the Arizona work.

David Stewart | CEO & Founder

David has been working with school accountability and student assessment data for more than 20 years. Before founding Tembo, David led the school accountability portfolio in the New York City Department of Education as the Executive Director of Evaluation and Performance Reporting. Prior to NYC DOE, David served as Vice President of Product Development at Grow Network/McGraw-Hill. David has deep experience in the design of school accountability systems, and has most recently worked with the Office of the State Superintendent of Education in Washington, DC, Shelby County Schools, Sacramento City Unified School District, the School District of Philadelphia, and the Public Charter School Board in Washington, DC.

Michael Moore | Chief Operating Officer

Michael is an expert in the visual display of quantitative information. He has completed his PhD coursework at the Graduate School of Education at the University of Pennsylvania. Prior to joining Tembo, Michael was an IES Research Fellow at the Center for Policy Research in Education, and served as the Editor-in-Chief of *Perspectives on Urban Education*. Michael has driven school accountability system development and related reporting projects in the School District of Philadelphia, Delaware Department of Education, Baltimore City Public Schools, and the DC Office of the State Superintendent of Education.



System Design



3.1 Metric Selection

Tembo would not want to propose specific metrics for an Arizona-specific framework until meeting with local stakeholders to learn more about your values and policies. That said, we believe that the selected metrics should encourage schools to focus on both achievement and growth for all students, regardless of their prior year scores.

3.2 Alignment to College /Career Expectations

We recommend developing a hierarchical linear model to better understand the relationship among a wide range of student performance measures and selected college- and career-related outcome measures. A better understanding of these relationships can help you understand the degree to which you may want to focus on one metric over another.

3.3 Differentiated Weights and Metrics

One way to think about how to compare the performance of all student subgroups in a school is to calculate the entire accountability index for each subgroup separately and to come up with a separate grade for each subgroup. Each school's overall rating could be based on the pattern of subgroup grades. This would clearly shine a bright light on any within school achievement gaps and would likely lead to a greater focus on each school's lower performing subgroups.

For parents, the question isn't "Which school is best?", it's "Which school is best for my child?" A school that receives an A because it has a high proficiency rate may not be the best school for a low-performing student if it doesn't also have high student growth scores. To improve a parent's ability to select an appropriate school, they must be able to determine which school would best improve their child's opportunity for success. This can be done with an intuitive, but flexible School Finder tool that allows users to better understand the relative strengths and weaknesses of each school.

To understand the improvement with certain types of schools across the state, state-level administrators need a sophisticated portfolio management tool that allows them to filter and group schools while looking at performance on one or more measures. For example, you may want to look at all schools on a scatterplot where the x-axis is performance and the y-axis is growth. If you could then color code the schools by some grouping variable (e.g. school type, grade level, zip code, city, etc.), you'd be able to quickly see the results of each group of schools. If you could further filter the results by some other school characteristic (e.g. % free and reduced lunch) and see changes over time, you'd be able to quickly determine which schools were improving over time.



Measuring Student Growth

4.1 Measuring Student Growth

Student growth measures all provide a good relative indication of how well one school is improving the scores of its students after controlling for prior test scores over another school. However, student growth measures don't tell us whether the growth for an individual student is enough to improve his or her likelihood of graduating from high school or participating in some kind of post-secondary education or career opportunity. In general, we prefer median growth percentiles to value-added growth scores because we prefer not to control for student characteristics. If a certain subgroup consistently grows less than another subgroup, controlling for that difference in growth isn't going to do anything to help students in the subgroup that have historically shown less growth. It may create a more "fair" accountability system for the school leader, but it may also just be a codification of lowered expectations.

To truly understand how much growth is enough, we recommend developing a model to understand each student's probability of achieving some future educational outcome and then monitoring the changes to these probabilities over time. For example, for a 7th grade student, you may want to know the probability that they'll graduate from high school in five years. If that probability is only 3%, then even if they earn a student growth percentile (SGP) of 65 when they take the 8th grade exam in that subject, it may not move the needle much in terms of their probability of graduating from high school.

The bottom line is that different students need to grow by different amounts if all students are to have an equal shot at post-secondary success. In an ideal world, a state-level accountability system would take this into account.

4.2 Experience

Tembo has significant experience in calculating growth scores for students, teachers, schools, and teacher preparation programs. We have run MGPs for the Office of the State Superintendent of Education in Washington, DC for the past several years and we recently developed a value-added model to better understand the degree to which high schools improve their students' chances of graduating on time after controlling for a wide range of middle school performance indicators and student characteristics. We have an in-house team of data analysts and we partner with expert researchers on an as needed basis.

4.3 Expediting Growth Scores

Tembo could help ADE develop the data formats and the code necessary to generate growth scores immediately following the receipt of validated scores. The scores could then be provided to the state's test vendor(s) in whatever format they require or Tembo could design and produce score reports that incorporate the growth scores.



4.4 Interpretation

As mentioned previously, simply knowing a student's growth score isn't enough. However, coupled with either a probability for some future outcome or the growth score that would be needed to have a particular probability for a future outcome, a student, parent, educator, or school would have a better indication of the meaning of the growth score.

