**Architects**

* David Stewart,Founder and CEO, Tembo Inc**.**
* Joe Siedlecki, Portfolio Director, Michael & Susan Dell Foundation

**A Note on Intended Uses and Limitations of Accountability System**

* The school is the unit of change and measurement. Rating schools, rather than districts or teachers, creates a building wide incentive to work as a team to improve education outcomes for students.
* This is a school accountability system, not a holistic evaluation of all things a school does. As such, we’d recommend that cities provide parents with supplementary information to inform their school decision making.

**Design objective(s)**

* Our annual accountability system evaluates the ability of an elementary school to prepare all students for academic success in middle school. This will create an incentive for schools to ensure student mastery of on the basic academic skills and knowledge necessary for middle grades success, therefore optimizing probability of student success in middle grades.
* The system:
  + Evaluates ability of an elementary school to prepare students for academic success in the next grade span (middle school)
  + Measures achievement and growth, but is designed in a way that prioritizes high achievement for students who have met college and career ready standards (CCR) and prioritizes high growth for students who have yet to meet CCR standards.
  + Creates incentives to serve all students well, and a to close gaps in growth

**Clear explanation of the proposed accountability system**

1. **Indicators of academic achievement**

* Includes a “proficiency”/CCR measure in reading, math, and science (if tested)
* Overall, reading and math would be counted equally and science would be counted 1/3 as much as either reading or math
  1. The logic being that reading and math are absolutely necessary basic skills for success in many other subject areas
* Include an “advanced” measure in both reading and math that counts 1/3 as much as the proficiency measure.
  1. This creates extra incentive to attain better than basic/proficient achievement
* 3rd grade reading and 5th grade math would be double counted
  1. Research has shown early grades literacy and 5th grade math to be important education milestones. This feature creates incentives for schools to focus on early grade literacy and then pivot to a focus on later grades mastery of math concepts.

1. **Indicators of student growth or an alternative**

* While there are myriad methods of calculating growth we use Median Growth Percentile (MGP). We do this because we don’t believe that the model should control for anything other than prior academic performance and because MGP models are easier for the public to understand.
* Equal weight would be given to reading and math, no other subjects would be included
* All students would be counted equally, so the median would be the median student in the group within the school, not the average of the median student in each grade
  1. This creates incentive to serve all students well
* We would also include MGP for all students and by subgroup in math and reading with comparison to MGPs of each subgroup in the district and state (metric would be difference between MGP of each subgroup with N>25 and the highest performing subgroup in the state)
  1. Calculating and showing the MGPs of each subgroup accomplishes two things: 1) it shows gaps in growth for students with the same abilities (prior scores) and 2) it encourages schools to do everything they can to ensure that no group is systematically learning less than their peers. Scoring for subgroup MGPs is described in the extra credit section.
* Change in probability of success (proficiency/CCR or above) on 8th grade math and reading assessments (This could only be calculated once there are enough years to build the model. It could start with probability of success on 6th grade math and reading assessments and then increase each year as more data become available.) This metric would also be calculated by subgroup. This could be both a performance (one year probability) and a growth measure (change in probability).
  1. This creates incentive to prepare students for middle grades. It also builds in a bit of a continuous feedback/improvement loop into the accountability system itself.

1. **Indicators of progress toward English language proficiency**

* Percent of students progressing by at least one level in a given year
  1. Note that we recommend a sliding scale of significance for this metric. See the scoring section for more information.

1. **Indicators of student success or school quality**

* We recommend that the “school quality” metric should be derived from the University of Chicago 5 Essentials Survey. The 5 Essentials Survey measures whether or not a school is organized for success and is predictive of future academic outcomes. Schools scoring strongly on 3 or more of the 5 essentials are 10 times more likely to improve student learning than schools that are weak.
* More specifically, we suggest measuring the Effective Leaders (teachers report that the principal works with teachers to implement a clear and strategic vision for the school) and Ambitious Instruction (students report that classes are academically demanding and engaging) domains every year. We suggest the schools take the full 5 essentials survey (Effective Leaders, Collaborative Teachers Involved Families, Supportive Environment, and Ambitious Instruction), once every three years.
* This method of measuring school quality does two things: First, it helps provide schools with a blueprint for the basic activities that a school must be good at to achieve strong academic gains for their students. Second, it is general enough that it applies to all types of school models. Third, it is backed by 20 years of evidence by the Consortium on Chicago School Research. Fourth, it has been administered at the state level (all schools in IL are surveyed), so has shown an ability to scale.
* We suspect some contestants will recommend using measures such as grit, growth mindset, and social and emotional learning data. We posit that such measures are too nascent in their development and application to be meaningfully included in accountability systems.

1. **Calculating summative grades**

* All metrics would be scored on a continuous scale using appropriate floors and targets. The floors and targets would partially be based on historical performance such as 10th and 90th percentiles and partially based on state goals/values (e.g. you may want to set a goal of 100% proficiency even if the 90th percentile score is lower than that)
* We recommend that achievement and growth (we include ELL growth in the growth categorization) account for 90% of a school score and that “school quality” metrics account for 10%.
* Achievement and Growth would be weighted 20/60, 40/40, or 60/20 depending on the results.
  1. High growth (75%+ of points earned in growth category) schools would have a higher growth weight.
  2. High achievement (75%+ of points earned in achievement category) schools would have a higher achievement weight.
  3. If a school earns more than 75% of points in both achievement and growth, it would have a higher weight in the domain with the higher score.
  4. All other schools would have the same weight for both domains.
* **Extra credit option 1:** The entire score could be boosted by 10% or lowered by 10% if all subgroup MGPs exceed the highest performing subgroup statewide or if all of the subgroup MGPs are lower than the highest performing subgroup statewide.
* The English Proficiency indicator and the school quality/student success indicator would each count as 10% of the overall score as long as at least 10% of the student population is an English Language Learner. If not, the percentages would be adjusted to 5 and 15 respectively.
* **Extra credit option 2:** *At some future time (after the US DOE has taken in to account our feedback (see below section))* we would include an “equity” section. We would measure student attrition rates for the school and for the low income sub group. Schools with student attrition rates higher than their city average would lose points. This creates an incentive for schools to continue to serve all students that have enrolled. As it stands, this data is usually collected at the LEA level and rarely reported to states. We would also recommend, in the intervening years, that schools with high level (X% of students with 15 or more absences) of chronic absenteeism (overall and in subgroups) be penalized.

1. **What about schools with low-performing sub-groups**

* All the same metrics would be calculated for the low performing subgroup as an additional category. The metrics would have their own floors and targets that encouraged exceptional growth. This category would take 20 points from the achievement/growth domains.

1. **School grades or ratings**

* Overall results would be described with four levels: High performing (blue), performing (green), approaching performing (yellow), and low performing (red). Within the low performing category we would further identify the lowest 5% (overall in state) as “Unaccredited” schools. These labels would also be applied to the domain scores based on similar percentage correct-based cut scores.
* Cut scores for these levels and all metrics weights, floors, and targets would be established using an accountability simulation tool. The state may start using this tool with an assumption that the distribution should be 10%, 40%, 40%, 10%, but the actual cut scores wouldn’t be determined until the state evaluated the actual distribution of scores. Every effort would be made to identify more natural cut points.
* We would also propose that school ratings and scores be displayed individually as well as against a peer group of similar schools (like demographically). This would help provide readers (policy makers, regulators, parents, and school administrators) with additional information on performance relative to other schools.

**Ideas for US Department of Education**

* We would recommend that the Department consider making collection and reporting of student attrition/mobility rates a requirement. These data are important to understand who a school is serving and if schools are providing students (and sub groups) with equitable access to an education.