In the context of the proposed Board Resolution 036- 17/18 "Achieving Excellence for All: Establishing a Framework for Continuous Improvement," this informative provides historical background on some of the accountability tools used in the past by the District and recommends criteria for establishing an up-to-date school effectiveness reporting framework. This informative does not evaluate the Resolution, but offers food for thought that is relevant to the Resolution.

A. Summary

In recent years, L.A. Unified has attempted various ways to be transparent and accountable to parents and students regarding the quality of its schools. The Academic Performance Index (API), a statewide indicator of school performance, was used widely from 1999 to 2013. Schools received Report Cards that reported a variety of school, student and teacher indicators from 2008 until 2017. Internally, during the CORE waiver period, schools were given Quality Indicators that guided District support policies. Currently, however, no district-wide system of school information reporting is available for L.A. Unified schools.

In this informative, the IAU reviews past school performance reporting systems that could be useful for comparison as the District develops its new framework for continuous improvement. We also recommend three criteria for a valid and operationally useful framework:

1. **Use specific language referring to instructional goals and education needs, rather than "deficit" definitions that focus on achievement gaps,**

2. **Use multiple measures, and**

3. **Avoid ranking except where necessary for distribution of funds with such a requirement.**

The IAU also suggests that it could assist in a technical advisory role in the development of a new framework.

B. Background

NCLB & API

In 1999, California's Public Schools Accountability Act created the Academic Performance Index (API) to measure the performance of schools and students. The API was produced consistently until 2013, when

the State Board of Education (SBE) suspended it while developing a new measurement system to represent schools' progress toward the state's revised educational goals. The numeric API score ranged from a low of 200 to a high of 1000. The interim statewide API performance target for all schools was 800. A school's growth was measured by how well it moved toward or past that goal.

API scores were calculated for all students in a school as well as numerous API scores for each subgroup at the school (such as by race, English Learner Status, students with disabilities, and socioeconomically disadvantaged pupils). In addition to the API score, each school received an API Statewide Rank score, which ranked a school with all schools in California based on API score, and an API Similar Schools score, which ranked schools relative to 100 other schools in the state with similar demographic profiles (including parent education level, poverty level, student mobility, and student ethnicity). Each rank ranged from 1 to 10, with a score of 10 meaning that the school's API fell into the top 10%.

In 2001, Congress reauthorized the Elementary and Secondary Education Act (ESEA) as the No Child Left Behind Law (NCLB). This law had a large impact on districts and schools across the country. Under NCLB, states were required to test students in reading and math and report back results for both the students as a whole and for various subgroups. Schools tracked their progress and goals using "adequate yearly progress" or AYP. Failure to regularly meet AYP was met with a series of sanctions, including allowing students to transfer out of school and offering free tutoring to students. [1]

API scores, ranking and indicators of AYP were available for more than a decade on the California Department of Education (CDE) website for all schools in L.A. Unified. Schools that consistently failed to meet achievement goals were designated "Program Improvement" schools, and could face sanctions or be provided with additional District support.

School Report Cards

In addition to information on schools available on the CDE website, L.A. Unified produced School Report Cards for every LAUSD school on an annual basis for eight years beginning 2008-09. The Report Cards included a variety of information including API scores, proficiency rates, school climate indicators, student and teacher attendance rates, and results for student and parent experience surveys. However, due to budget cuts, the School Report Cards were discontinued in 2015-16.

CORE Waiver and SQIS

When it became clear that the majority of schools nationwide were not going to reach their proficiency goals by 2013-2014 and Congress failed to act when NCLB was due for reauthorization in 2007, the Obama administration created a series of state waivers for NCLB's requirements. [2] Two years after the NCLB waivers were announced, the U.S. Department of Education approved eight California school districts for a waiver. [3] This group of Districts formed a cooperative, called the California Office to

[1] Editorial Projects in Education Research Center. (2015, April 10). Issues A-Z: No Child Left Behind: An Overview. Education Week. Retrieved from http://www.edweek.org/ew/section/multimedia/no-child-left-behind-overview-definition-summary.html/
[2] U.S. Department of Education. (2011, September 23). Letters from the Education Secretary or Deputy Secretary. Retrieved from https://www2.ed.gov/policy/gen/guid/secletter/110923.html
[3] U.S. Department of Education. (2013, August 6). Obama Administration Approves NCLB Waiver Request for California CORE Districts. Retrieved from https://www.ed.gov/news/press-releases/obama-administration-approves-nclb-waiver-request-california-core-districts

Reform Education (CORE), and committed to a set of deliverables and activities and were accountable to each other as well as to the federal government.[4]

In addition to data sharing, cross-collaboration, and research partnerships, the CORE districts developed an accountability system called the School Quality Improvement System (SQIS). SQIS used ten different indicators to understand how schools were reaching the primary goals and improving upon their prior results. These are described in Table 1.

---

[4] Note that there were two more districts in CORE that did not submit a waiver application, Clovis Unified and Garden Grove. Sacramento City and Sanger later did not renew their waiver applications when the other six districts did. Kudson, J. and Garibalid, M. (2015). None of Us Are As Good As All of Us: Early Lessons From the CORE Districts. *AIR.* Retrieved from http://coredistricts.org/wp-content/uploads/2017/08/AIR-Report-August-2015.pdf

Table 1. School Quality Improvement System Improvement Measures for the Index[5]

Indicator	Metrics	Description
Student Academic Growth	Student Growth Percentile (SGP), scale of 1 to 99 with a point estimate and 95% CI	Examines the extent schools help students move from point A to point B relative to their peers[6]
Student Social-Emotional Learning	Self-reporting qualitative survey by students in grades 5-12; scale of 1 to 99	Four competencies examined: growth mindset, self-efficacy, self-management, & social awareness
Indicator	**Metrics**	**Description**
High School Readiness	Students have the following in 8th grade: Min GPA of 2.5; min attendance 96%; no Ds or Fs in ELA or Math; never suspended	Shows the % of 8th graders who meet criteria that indicates likelihood of graduating on time
School Culture/Climate	Qualitative surveys by students in grades 5-12, teachers and staff, and parents and guardians	Teaching & learning, interpersonal relationships, safety, & school-community engagement
College and Career Readiness	N/A	Under development with Linked Learning Analytics
Graduation Rates	Students placed in a 4-year cohort. The # of graduates 4 years later/ students in the cohorts = the 4-year cohort grad rate (same for 5- and 6-year rates).	Examines 4-year, 5-year, and 6-year cohort graduation rates
Suspension Rates	(1) % of individual students suspended/ expelled at a school during the year (i.e., unduplicated counts of student suspensions/ students enrolled). (2) # suspensions per 100 students at school. Reported to schools only.	The percent of students suspended and/or expelled, reported in two ways
English Learner Progress	# of students who re-designate in a year divided by those students re-designated during the year plus any non-re-designated ELs with five years or more of instruction[7]	Includes a re-designation metric specifically to attend to the reclassification of ELs as FEP before students become Long-Term ELs
Chronic Absenteeism	# days attended/ # days student enrolled. # of chronically absent students is aggregated to the school level to find the # and % of chronic absences. Students must be enrolled min 45 days. If a student is enrolled in multiple schools, data are aggregated across all schools attended. Aggregate determines if student was chronically absent and is used for the student's last school.	Chronic absence is defined as missing 10% or more of school over the course of the school year for any reason, including excused and unexcused absences
Academic Performance	The percentage of students who meet grade level standards in ELA and Math as measured by state tests	Uses the scale score on statewide tests for students in grades 3-8, similar to the CA school dashboard
Disproportionality in SPED	Formula: $E = A + \{3*\sqrt{[A*(100-A)/N]}\}$ E = Max % of the total SPED enrollment in a school allowed for a specific ethnic group (3 SE above expected value) A = % of the same ethnic group in general education in the school (expected value) N = The school total SPED enrollment	Identifies cases where a subgroup is over-represented in SPED with statistical significance (99% CI). Data sent to districts to decide how to share info, the school a student attends may not be the school in which s/he was identified for SPED.

---

[5] CORE Districts. (2017). School Quality Improvement System Improvement Measures. Retrieved from http://coredistricts.org/our-data-research/improvement-measures/

[6] Controlling for prior achievement and the following demographics: SPED, EL, FRLP, homeless, and foster youth

[7] The calculation method excludes EL students identified for SPED who cannot take the CELDT with reasonable accommodation

The inclusion of Student Academic Growth in the improvement metrics using the Student Growth Percentile (SGP) brought CORE into conflict with the teachers unions.[8] The use of student growth as part of a teacher evaluation system led nine teachers unions, including California Teachers Association and United Teachers of Los Angeles (UTLA), to sign a public letter opposing the waiver proposal.[9] Out of the CORE districts, only United Educators of San Francisco officially supported the waiver.[10]

The accountability system instituted by CORE used the above SQIS indicators to designate schools into a detailed series of five categories, all but one of which had subcategories, for a total of eight designations. This categorization of schools was the backbone of a school support structure that provided different levels of support to different categories of need.[11] The lowest four categories were identified for after-school intervention, professional development support and other services.[12]

In December 2015, the Every Student Succeeds Act (ESSA) was signed into law as the reauthorization of the Elementary and Secondary Education Act (ESEA). With this, CORE's system was no longer the basis for school accountability. While the CORE accountability system was seemingly abandoned by the 2016-2017 school year, the assignments of particular school categories remains.

Given that the school designations are still used, it is possible that the metrics for SQIS accountability are still being tracked. However, the District states that it does not maintain an "annual list" of lowest performing schools, and that it has not yet compiled a bottom 5% list as the state Board of Education has not yet established the performance criteria.

C. Recommendations

1. **Use specific language referring to instructional goals and education needs, rather than "deficit" definitions that focus on achievement gaps.**

   When conversations about school performance take place in centralized and public settings, the discussions sometimes dangerously veer away from specifics and into the realm of generalities. To characterize a school as "underperforming" overall invalidates areas where the school may be succeeding and prevents emphasis from being placed on specific problem areas. It also creates a climate of ambiguity and uncertainty.

   In theory, labeling schools that fail to meet academic expectations as *underperforming* is one of the first steps in the process towards school improvement. Once identified as underperforming, schools are more likely to receive greater attention and deeper analysis to find and understand trends that may be negatively affecting student and school achievement. In practice however, public conversations about underperformance often lack specificity while the deeper analysis is conducted out of the public eye. Publicly highlighting a school's underperformance without

---

[8] Aron, H. (2013, July 19). Teacher Unions Oppose NCLB Waiver that the 'CA 9' Want. *LA School Report*. Retrieved from http://laschoolreport.com/teachers-unions-oppose-federal-waiver/
[9] California Teachers Association. (2013, June 12). NCLB Waiver CORE Union Letter. Retrieved from https://www.documentcloud.org/documents/726825-nclb-waiver-core-unionletter061213.html
[10] Kudson, J. and Garibalid, M. (2015).
[11] Note that non-Title I schools are not classified and not all Title-I schools received a classification.
[12] L.A. Unified. (2013, December 10). CORE Waiver FAQs. Retrieved from file:///C:/Users/sydney.ganon/Downloads/CORE_Corner_FAQs_121913.pdf

simultaneously contextualizing its specific indicators can be a demoralizing force that negatively affects staff and student performance.

When a school is considered underperforming, the notion of who is failing to meet expectations is ambiguous. It can sound like the students are underperforming. Are the students, teachers, school site administration, or district administration to blame for the underachievement of a given school? Researchers claim that the lack of clarity around ownership over school performance negatively effects the psyche of students in these schools. Students are likely to heed generic critiques of school performance as a form of personal criticism and place the blame on themselves rather than on the institutions expected to serve them. This can create a self-fulfilling prophecy in which students interpret generic descriptions of underperformance as a reflection of their own ability. Such an interpretation can lead students to conclude that they cannot meet academic expectations, thereby perpetuating underperformance.[13]

To continue increasing transparency, accountability, and school performance across L.A. Unified, the District would be well served to be specific about the areas in which a school is underachieving and the potential root causes. The Independent Analysis Unit (IAU) recommends taking into account at least the following contextual factors when evaluating school performance:

- school type;
- Board District and Local District;
- total school enrollment;
- enrollment as a percent of capacity;
- specialized program options;
- overall school budget;
- student demographic and socioeconomic information;
- neighborhood safety;
- and basic services and conditions.

This list may need to be shortened for efficiency, but a handful of these factors--if consistently applied--will go a long way toward clarifying the meaning of language about school performance. This contextual data would be in addition to school-level academic performance indicators such as SBAC English language arts and mathematics results, course passing rates by grade level, attendance rates, graduation rates, and reclassification rates.

2. Use multiple measures

Schools are tasked with instruction in a variety of academic subjects, across anywhere from three to twelve grade levels and for students with a wide variability of needs. It stands to reason that no one measure can capture the varied outcomes these efforts produce. In addition, even if a single indicator of student achievement is reflective of school leadership, organizational function, school climate, and teacher quality, if that measure stands alone, there is no way to determine which of the many features of schools supports or inhibits success. Both input and output must be measured and distinguished from background noise. Behavioral, cognitive and non-cognitive indicators are all useful in different ways.

---

[13] http://ices.library.ubc.ca/index.php/criticaled/article/viewFile/183936/184137

Even specialized measures intended to indicate progress in single subject areas and for particular groups of students can mislead. For instance, many local education agencies, including L.A. Unified, commonly use *reclassification rates* as a way of evaluating programs for English learners (ELs), with higher reclassification rates typically interpreted as signifying higher program quality. Certainly there are advantages to including reclassification rates in school accountability plans. For example, research shows that it is harmful to keep ELs in EL programs for longer than necessary as it may inhibit full access to rigorous content and reduce opportunities to interact with fluent English speakers.[14]

However, there are also disadvantages associated with an overemphasis on reclassification rates, such as the premature removal of specialized language and academic supports. A narrow focus on reclassification rates may pressure schools, fearing corrective action, to reclassify ELs before they are ready to exit the program. Such incentives may also lead schools to neglect students they deem too far away from reclassification in one year.

For analytical purposes, another disadvantage of too much focus on reclassification rates is a possible failure to notice long-term reclassification patterns. Dual language programs, for instance, have lower cumulative reclassification rates in the elementary grades, but experience *higher* cumulative reclassification rates in the long run compared to English immersion programs. Focusing on cumulative reclassification rates in 5th grade, for example, could result in the erroneous assumption that dual language programs are less effective than English immersion programs. However, by looking at longer-term outcomes such as cumulative reclassification rates in 11th grade, researchers show that ELs in dual language programs have higher overall reclassification rates compared to ELs in English immersion programs. In this case, remaining an EL longer is *not* necessarily associated with poorer academic outcomes.[15]

This does not invalidate reclassification rates! They provide important information about a school, and encourage the effort to move EL students into the full curriculum. However, for analytical purposes, a complementary, more objective measure such as growth in CELDT scores will provide a more nuanced picture of a school's success in helping all of its ELs make progress toward reclassification.

In general, education researchers recognize that any one measure is necessarily incomplete and may be unreliable. Combining multiple measures fills in the missing parts of the puzzle and contributes to increased reliability.

3. **Avoid ranking except where necessary for distribution of funds that require such ranking**

- Ranks based on one or a few measures risk being reductionist, when not only the reality of the school but the preferences of parents are multi-dimensional. Simply put, one family's most preferred high school because of its high mathematics achievement may be much lower on another family's list because of its lack of arts instruction. It is difficult to measure the relative quality of schools due to differences in student composition, neighborhood context, historical challenges, and programmatic variety

---

[14] Umansky, I. M., & Reardon, S. F. (2014). Reclassification patterns among Latino English learner students in bilingual, dual immersion, and English immersion classrooms. *American Educational Research Journal*, *51*(5), 879-912.
[15] *ibid.*

- Ranks based on multiple measures require that those measures be weighted, usually arbitrarily. For example, suppose percent of students in poverty is one indicator, and prior average achievement scores in Math and ELA are two more indicators. To obtain a numeric ranking order, those three data points must be combined to equal one number. This is worse than apples and oranges. Does one take 2.5 times achievement in Math plus 2.0 times achievement in ELA minus 3.5 times the percent in poverty? Or some other weighting? For an accountability system to be credible, it is necessary to identify where those weights come from.

- Since the factors and weights that make up the ranking should be public for the sake of transparency, it is sometimes possible to figure out how to game the ranking, i.e., obtain a higher rank with less actual effort. Even without a public formula, people respond to hidden incentives; over time, they might notice which of several factors gets rewarded, and shift their behavior toward that factor--perhaps in a way contrary to the intention of the index designers.

While a single ranking order may be necessary at times, especially when distributing limited resources, all other purposes benefit more from a profile of carefully chosen indicators. To be more reliable and robust, the profile should include many indicators. However, to be understandable to the lay public and parents, it should be as simple as possible. These are competing demands that a group designing such a system must weigh. There will always be stakeholders who want a single ranking index order, but it is important to recognize that their preferences might not match the preferences of others, and yielding on this point compromises the fairness of the system to other stakeholders.