COLORADO
Department of Education

MEMO

TO: State Board of Education
FROM: Rhonda Haniford, Associate Commissioner, School Quality and Support Division and Lisa Medler, Executive Director, Accountability and Continuous Improvement Unit
RE: 2022 Accountability and Data Availability
DATE: December 2, 2021

In exploring options for the 2022 state accountability process, the State Board of Education requested two follow up studies after the October Board Retreat: (1) a summary of the state board's authority to make adjustments to the current accountability system and (2) an impact study on plan type assignments considering current data elements. The first request – summary of the state board's authority – is being developed by the Attorney General's office and is expected to be available in mid-December. This memo focuses on the second request – impact study on plan type assignments considering current data element availability. Attached to this memo is a PowerPoint to provide further details and visualizations of the findings. The department will produce a more in-depth report on the impact study for release later in December.

Summary of Findings

In looking ahead to calculating 2022 performance frameworks:

- The 2020 state assessment pause and the 2021 modified state assessment schedule have decreased the full dataset needed to run 2022 performance frameworks. In calculating overall ratings, the greatest implications are for the academic growth indicator and a subset of schools and districts that will not have enough data to calculate a rating. In particular, elementary and middle schools will have fewer data points in 2022.
- CDE's recent impact study reveals that by comparing the complete 2019 state accountability dataset with a Mock 2022 dataset (using the 2019 data and removing the missing data elements anticipated in 2022):
    - Just under 90% of schools and districts (including BOCES that receive frameworks) received the same overall framework rating. About 5% increased their rating; about 5% decreased their rating. For every ten schools, the missing data was a factor in getting the same results for at least one of those schools.
    - Very strong correlations were found for both districts (0.991) and schools (0.976) when looking at framework rating assignments -- meaning there is a strong relationship between framework assignments in 2019 and in the Mock 2022 dataset. Correlations for schools and districts by individual rating categories were slightly lower, particularly for Priority Improvement and Improvement plan types. Correlations by school level (i.e., elementary, middle, high school) were strong across all levels, but were strongest for high schools.
    - Approximately 80% of schools and districts received the same growth indicator rating.
    - Schools with less than 75% test participation were about 10% more likely to experience an overall framework rating decrease than schools with higher levels of participation. The interaction of participation with this analysis is important. When looking at 2021 assessment results, participation was considerably lower than in past years. Participation rates for 2022 are not yet known.

- Given the complexity of analyzing participation data, it may be helpful to continue to investigate the potential impact of low participation on any possible framework ratings for fall 2022.

More Detailed Explanation of Findings

- *Current Context* (slide 3): Given current statute, 2022 state assessments and accountability (state and federal) are slated to resume without any adjustments. Assuming the state assessment schedule proceeds as is, the data needed to calculate performance frameworks are mostly available (e.g., academic achievement, post-secondary and workforce readiness indicators). Academic growth is the one indicator that is most impacted by the adjusted 2021 assessment schedule.
- *Missing Data* (slides 4-6): There are two main areas where the state is missing data: (1) Alternating 2021 grade/content CMAS data to calculate growth for all elementary and middle schools and (2) Three-years of aggregate data to calculate frameworks for smaller schools and districts due to assessment pause in 2020.
    - *Elementary and Middle School Growth* (slide 4)*:* Because the state needs two years of data to run growth, the 2021 alternating grade/content assessment schedule means that there is half as much data as in a typical year for elementary (2 out of 4 possible grades/content areas) and middle schools (3 out of 6 areas). Growth calculations for districts may be calculated with two thirds as many data points as in a typical year (10 out of 15 grades/content areas). High schools are not affected as PSAT/SAT was offered in full in 2021 and the 8th grade ELA CMAS to 9th grade EBRW PSAT growth calculation has not been run historically because of construct alignment issues.
    - *Three-Year Frameworks* (slides 5-6): CDE projected plan types that would need an Insufficient State Data (ISD) rating in 2022 because 3-year frameworks cannot run or because 2021 participation was low enough that n-count thresholds cannot be met (i.e., 20 students for growth). These scenarios affect at least 187 elementary schools (compared to 30 in 2019), 128 middle schools (compared to 130 in 2019) and 177 high schools (compared to 12 in 2019). These are conservative estimates, as 2022 assessment results are not yet available and may likely increase the number of schools and districts with ISD plan type assignments.
- *2022 Mock Data Impact Study* (slides 7-28)
    - *Description of Study* (slides 7-10): Using the most recent complete school and district performance framework dataset (2019), CDE mocked up the anticipated dataset for 2022 including the missing data described above. The 2019 complete dataset (e.g., growth, overall framework ratings) was compared with the Mock 2022 dataset. Calculations for the Mock 2022 ratings were conducted consistently with 2019, with the exception of WIDA ACCESS On Track Growth and the READ bonus point. It should also be noted that the study could not adjust for participation rates since 2022 rates are not yet known. The study models the impact of the missing data and is *not* predictive of results in 2022.
    - *Rating Assignment* (slides 11-15): The distribution of ratings statewide across all categories stayed relatively consistent across all rating types (i.e., Turnaround, Priority Improvement, Improvement, Performance, Distinction for districts). When looking at changes in plan type between the 2019 and 2022 Mock dataset, 88% of districts and 89% of schools received the same plan type. Approximately 7% of districts received a higher rating and 5% received a lower rating on the 2022 Mock dataset than in 2019. For schools, 5% received a higher plan type and 6% received a lower plan than in 2019. When examining rating assignments by school level (i.e., elementary, middle, high school), the comparison between the two datasets is still close - within 1-2% for each rating for each school level. For the accountability clock, districts (93%) and schools (88%) kept their plan type in the comparison.
    - *Correlations for Total Percent of Points* (slides 16-17): The correlations on total percent of points earned on framework rating categories between 2019 and the Mock 2022 data were very high

for districts (0.991) and schools (0.976). Correlations within individual rating categories were slightly lower but still fairly strong. The lowest correlations were for schools with Improvement (0.875) and Priority Improvement (0.821) plan types. Note: The threshold for the accountability clock rests between these two categories. Correlations stayed strong when examining by school level ranging from 0.766 (Improvement for elementary schools) to 0.999 (all high schools).

- *Growth Indicator* (slides 18-21): There were some slight variations in the growth indicator ratings (i.e., exceeds, meets, approaching, does not meet, no rating) for schools and districts when comparing 2019 with the Mock 2022 datasets. Correlations between 2019 and the Mock 2022 datasets were fairly strong for both districts (0.850) and schools (0.849) when looking at all growth rating categories combined. However, correlations within individual growth indicator ratings were lower and varied considerably for schools. For example, the exceeds category had the lowest correlation (0.383). Around 80% of districts and schools kept the same growth rating when using the limited mock 2022 data set when compared with 2019. Approximately equal proportions (9-12%) of districts and schools increased or decreased one or more rating categories.
- *Participation Bands* (slides 22-26): To study the impact of lower test participation, CDE categorized schools and districts by participation bands (i.e., at or above 85%, 75-84.99%, below 75%). 85% participation is considered generally sufficient for ensuring representativeness of the student population; 75% is less rigorous but still reasonable**.** More than 90% of districts received the same rating in the Mock 2022 data set, regardless of the student participation rate. The majority of schools received the same rating with the 2022 mock data set. However, schools with less than 75% participation had an increased (10%) likelihood of receiving a lower framework rating on the 2022 Mock dataset. It may be helpful to continue investigating the impacts of lower test participation rates on producing representative framework ratings using the 2021 and potentially simulated and/or imputed data sets.
- Note: While the correlation of results between the full 2019 data set and the limited Mock data set are strong, it is important to note that on-going reduced assessments (like those given in 2021) would limit the ability to calculate growth information. Growth calculations require two years of scores for each student in the same content area.

Next Steps

- CDE staff are available to answer questions for board members. Please contact us or work through Angela Maramba to set up one-on-ones. We can also be available to present at the January board meeting.
- CDE staff will produce a report that provides more detail on the studies referenced in this memo.