Comments as of 3/10/11 on Implementation of Education Accountability Act of 2009 (S.B. 09-163)

	Comment	CDE Response
General		
1	We are more likely to see improvements in school performance if the state continues to adhere to a consistent accountability system, rather than wiping the slate clean and beginning to construct yet another accountability system now.	The Department will prioritize consistency when evaluating potential tweaks to the school and district frameworks.
2	The use of acronyms in accountability documents and training materials should be minimized to make information more transparent.	Documents will be reviewed to minimize the use of acronyms.
School and District Performance Frameworks		
3	The use of 5- and 6- year on-time graduation rates should be the default for the school performance framework reports, so that schools automatically are evaluated based on whichever rate generates the best score.	The Department intends to report 4-, 5-, and 6-year graduation rates on the school and district performance framework and to award points for whichever rate yields the highest rating on the school and district performance framework. *See CDE Discussion Memo on Graduation Rate Concept.*
4	The weighting of postsecondary and workforce readiness measures is too high and disadvantages districts with diverse student populations. The weighting of growth is too high and disadvantages high-achieving school districts.	The Department has attempted to balance the different perspectives from school districts in Colorado while meeting the statutory requirement to give the greatest consideration to growth and postsecondary and workforce readiness. In addition, the Department has also attempted to reflect the statutory expectation that all students graduate ready for college and career success. We believe this requires substantial weight for postsecondary and workforce readiness measures including graduation rates. The weighting was vetted extensively with the field prior to finalization. The Department believes maintaining the current weighting has advantages for comparability, training, and public awareness building.

5	School and district median growth percentiles are linked to N-size and to characteristics of the student population (e.g., poverty). Large and economically impacted districts/schools are more likely to have very high median adequate growth percentiles and therefore will have to obtain a MGP of 70 or higher to exceed state expectations. Overall, these factors combine to make the frameworks biased in favor of small districts with low percentages of students in poverty.	The Department disagrees that N-size causes a bias in the accreditation categories it assigns to school districts and the plan assignments it makes for schools. The use of multiple performance indicators, three-year averaging, inclusion of adequate growth, weighting of status and postsecondary and workforce readiness, and assignment to only four broad categories of schools and five categories of districts, all mitigate the N-size correlation of median growth percentiles. *See CDE Research & Evaluation Unit response to Poudre Research & Evaluation Unit paper.*
6	Median adequate growth percentile is a concept that is difficult for educators to understand and the MAGP is often an unrealistic target. The frameworks do account for that in the rubric, but it makes the scoring very complex.	The Department agrees that adequate growth can be perceived as a complex concept as built into the SPF and DPF. However, it is an essential concept, namely the amount of growth needed for a student to catch up or keep up. The AGP included in the Performance Frameworks is not a hard target, so even if it seems so high that it is realistically unattainable, reaching the target is not necessary for receiving full points on that sub-indicator or indicator. The Department is developing materials targeted to parents that will simplify the concepts and the Department will make every effort to build awareness and improve training materials.
7	Including adequate growth in the growth indicator rubric re-introduces a status component to what would otherwise be pure growth. This sets a higher standard for low SES schools that typically have lower status achievement.	The Department agrees with this observation, but disagrees that the influence of adequate growth in final growth ratings is too great. A basic tenant of the Education Accountability Act of 2009 and CAP4K is the expectation of all students reaching PWR. Therefore, statute requires that we include the adequacy of growth to reach that destination in our school and district accountability system. Adequate growth currently exerts a minimal influence on the final growth ratings for districts and schools. *See* CDE Research & Evaluation Unit's *Response to Poudre Research & Evaluation Unit Paper*.

8	Should schools and districts be penalized for not meeting their participation rate requirements on the state standardized assessments in more than one subject area? It seems frustrating that if a student is frustrated and quits the test, he is considered a non-participant.	CDE will consider methods for revising SPF and DPF reports to include additional information about reasons for non-participation. Please provide any specific recommendations that you think would be helpful for this addition to the framework reports.
9	A large number of districts that are accredited with distinction are very small and all of them seem to serve mostly white and "non-poverty" students.	The Department disagrees that the distribution of districts by size or demographics constitutes a bias in the performance frameworks. The great majority of Colorado school districts have fewer than 1,000 students, so these will tend to be well-represented in all accreditation categories. It should also be noted that there are many, non-diverse, low-poverty small districts that are not accredited with distinction. Also see response to #6 above and CDE Research & Evaluation Unit's *Response to Poudre Research & Evaluation Unit Paper*.
School and District Performance Framework Reports		
10	It seems reasonable to have the release of school performance frameworks to principals either a week after the superintendents or at same time.	The Department intends to make SPFs available to principals through CEDAR one week following their availability to superintendents. Note that it is the district's local access manager's responsibility, however, to create users and assign them to school-level accounts.
11	Locating the framework reports on the CDE website is somewhat confusing. On SchoolView, information on school district performance can be found in two sections, School Performance and the Learning Center.	The Learning Center will only include training and awareness materials.
12	Parents have indicated that the school and district framework reports are confusing and not easily understood without help. If the state wants parents and other stakeholders to use SPF/DPF reports to make decisions about school choice, it will have to provide some in-depth training resources aimed specifically at lay audiences.	The Department has initiated a communications strategy to gather more input from parents and community members before making the next round of revisions to SchoolView.

Comments as of 3/10/11 on Implementation of Education Accountability Act of 2009 (S.B. 09-163)

Request for Reconsideration		
13	CDE could provide greater clarity about what constitutes a good Request for Reconsideration and could perhaps post more samples of what was not acceptable to CDE for reconsideration, section by section.	The Department will highlight specific examples of successful and unsuccessful Requests for Reconsideration on SchoolView.
Unified Improvement Planning		
14	More training is needed on unified improvement planning, particularly universal level training; training should be provided earlier in the year.	The Department has used the limited funds available to prioritize support for the districts and schools that demonstrated the most need. Additional funding is being sought.
15	It is not clear how the unified improvement plan template is simpler for some rural districts—in some instances, it seems like more information is required now than was required before.	The unified improvement plan template does not require any information from districts that is not already required by federal and/or state law.
16	The unified improvement plan template should include more CDE plan requirements for other units of CDE, but only where it makes sense to do so.	The Department will consider how to expand the use of the Unified Improvement Plans.
17	The department might consider how to develop a ***voluntary*** peer review process for unified improvement plans for those schools/districts whose unified improvement plans are not reviewed by CDE.	The Department supports this recommendation and looks forward to collaborating with the field on making this possible.
18	Review and approval of school plans, and their assigned plan type, should be a district responsibility.	The Department only reviews Priority Improvement and Turnaround plans for schools, and leaves review of all other plans to district. The Department currently is responsible for approving school Turnaround plans, but is seeking a statutory amendment that would leave this responsibility to districts.
19	The unified improvement plan form needs some minor revisions, as it seems to be redundant in some areas and somewhat confusing.	The Department will be making minor revisions to the unified improvement plan template that will make the template less redundant and confusing. Specific recommendations are

		welcomed.
20	Although the unified improvement plan template is complex, districts should have time to become familiar with this format before significant revisions are made that require districts to start a new learning process again.	The Department will attempt to keep the major components of the unified improvement plan template consistent when making revisions to the template.
21	There has been limited use of school or district accountability committees in developing the unified improvement plans. It is difficult for them and parents to understand the template.	The Department has initiated research to gather more input from parents and community members before making the next round of revisions to the unified improvement plan template.
22	The timeline for developing and submitting unified improvement plans does not align with school planning times (i.e., schools generally plan at the end of the year or over the summer and implement in the fall).	The timeline for submission of plans is aligned with federal requirements for plan submissions, in order to allow districts and schools to submit a single plan that addresses both state and federal requirements.
23	The process of developing plans with staff, although difficult, has been very beneficial. This is particularly true of the process of analyzing data and root causes of low performance.	The Department is pleased that this is the case.
24	Districts should be granted more flexibility to submit plans tailored to the district's needs, but which include all required criteria. Additional plan templates could be pre-approved by CDE, or CDE could provide examples of alternative models.	The Department supports district efforts to improve the planning template. However, the intent behind a standardized unified plan includes: • Ability of state to create a relational database that can be searched by educators; • Developing common language and understanding around improvement planning; • Efficiencies in CDE review of plans for federal purposes; and • Efficiencies in posting plans to SchoolView and allowing greater public transparency.
25	Districts with fewer than 1000 students should be permitted to submit a single plan that satisfies the district and school plan requirements.	The Department agrees and is seeking a statutory amendment that will allow this.

DISCUSSION MEMO ON GRADUATION RATE CONCEPT

From: Marie Huchton; Somoh Supharukchinda

To: Executive Team

Date: March 7, 2011

RE: Incorporating the new on-time graduation rate into the District and School Performance Frameworks

Background

Currently, the District and School Performance Framework (DPF, SPF) reports include the old, adjusted graduation rate calculation as a part of the Postsecondary and Workforce Readiness (PWR) indicator. (The other two postsecondary measures are drop-out rate and composite ACT score.) The Postsecondary indicator accounts for 35% of a district or high school's overall framework points. Districts receive the following ratings based on their adjusted graduation rate:

Graduation Rate	Greater than or equal to 90%	**Exceeds**
	Greater than or equal to 80% and less than 90%	**Meets**
	Greater than or equal to 65% and less than 80%	**Approaching**
	Less than 65%	**Does Not Meet**

To comply with No Child Left Behind requirements and State Board rules, Colorado must now shift to the new graduation rate calculation using the formula and methodology set by the National Governors Association "Graduation Counts Compact." The introduction of this new, on-time graduation rate calculation has the potential to impact several districts' and schools' ratings, so it is critical that CDE is cognizant of its use in the DPF and SPF and its impact on state accountability. Simply replacing the old calculation with the new calculation could result in unintended consequences for those districts where the new, on-time 4-year graduation rate is lower than the adjusted rate. Some districts' 4-year graduation rates could be lower – and may no longer meet the 80% minimum state expectation – as a result of practices CDE supports (e.g., re-engagement efforts, concurrent enrollment, time-variable promotion).

"In some cases, the new formula would appear to penalize districts that are making a concerted effort to keep students in school," said Deputy Commissioner Diana Sirko. "If a district runs a strong concurrent enrollment program, for instance, they could be doing a terrific job of keeping students engaged in school but the new on-time rate makes it look as if their overall performance has dropped. The new formula is not designed to send a message about the pros or cons of efforts to provide safety nets or genuine alternatives for students. The new formula provides a common definition nationwide for comparability's sake—and that's all."

Old versus new calculation of graduation rates

The new four-year formula defines "on time" as only those students who graduate from high school four years after entering ninth grade. Under this four-year "on-time" formula, a student is assigned a graduating class when they enter ninth grade. The graduating class is assigned by adding four years to the year the student enters ninth grade. In other words, the formula anticipates that a student entering ninth grade in fall 2010 will graduate with the Class of 2014.

With the old system, students who took longer than four years to graduate were factored into the formula.

Recommendation

CDE should incorporate 4-year, 5-year and 6-year graduation rate calculations (under the new formula) into the DPF and SPF, and give districts and schools credit for whichever rate is highest. While the 4-year graduation rate from the most recent cohort provides the most current information about performance, the 5-year and 6-year rates are better indicators for those districts and schools making a concerted effort to keep students in school (to prevent drop-out, better prepare students for postsecondary and workforce readiness, etc.).

"We want to give districts credit for whatever rate is highest," said Associate Commissioner Richard Wenning. "This reinforces the principle of allowing time to become a variable given Colorado's expectation that all students will leave high school prepared for postsecondary education or the workforce."

CDE should publish all the available graduation rates for the three most recent cohorts. This means that up to six graduation rates could be published on the DPF and SPF:

- The 4-year graduation rate for the 2010 cohort
- The 4-year and 5-year graduation rate for the 2009 cohort
- The 4-year, 5-year and 6-year graduation rate for the 2008 cohort

For the most recent 2010 cohort, those students who started ninth grade in 2006 and had 2010 as their anticipated year of graduation, only the 4-year graduation rate is available. Their 5-year graduation rate will become available at the end of the current school year, and their 6-year rate will become available the next year. For the 2009 cohort, those students who started ninth grade in 2005 and had 2009 as their anticipated year of graduation, the 4-year rate and the 5-year rate are available, but their 6-year rate will only become available next year. For the 2008 cohort, those students who started ninth grade in 2004 and had 2008 as their anticipated year of graduation, the 4-, 5-, and 6-year rates are available. This is because we already know what percentage of the 2008 cohort graduated in 2008 (4-year rate), 2009 (5-year rate) and 2010 (6-year rate).

The table below gives a visual representation of all the graduation data available from the prior three years. Note that rates are demonstrative only and that rates where the graduation base counts are below 16 will not be reported or included in accountability decisions.

		4-year	5-year	6-year
Anticipated Year of Graduation	**2008**	89.7	91.6	***92.8***
	2009	86.7	88.5	
	2010	89.6		

1-year DPF and SPF calculation

For the 1-year DPF and SPF, points will be awarded based on the highest value among the following: 2010 4-year graduation rate, 2009 5-year graduation rate and 2008 6-year graduation rate (the shaded cells in the table above). By taking the best result from among these three options (shown in bold italics/red), districts and schools can be rewarded for their graduation successes, whether or not that is reflected in their "on-time" 4-year rate or a more delayed 5- or 6-year rate.

3-year DPF and SPF calculation

For the 3-year DPF and SPF, points will be awarded based on the highest value among the following: aggregated 2008, 2009 and 2010 4-year graduation rate, aggregated 2008 and 2009 5-year graduation rate, or aggregated 2008 6-year graduation rate (the shaded cells in the table below). For each of these, the aggregation is the result of adding the graduation totals for all available years and dividing by the sum of the graduation bases across all

available years. Note that rates are demonstrative only and that rates where the graduation base counts are below 16 will not be reported or included in accountability decisions.

		4-year	5-year	6-year
Anticipated Year of Graduation	**2008**	89.7	91.6	92.8
	2009	86.7	88.5	
	2010	89.6		
	Aggregate	88.7	90.1	***92.8***

Alignment with NCLB Accountability/AYP calculations

Colorado has submitted an NCLB Accountability amendment request for AYP calculations that follows similar requirements to the recommendation presented above. The state's proposal to the US Department of Education allows for graduation rate targets to be met by any of the following options:

- 2010 4-year on-time rate;
- An increase from the 2009 4-year rate to the 2010 4-year rate;
- 2009 5-year rate; or
- 2008 6-year rate.

Anticipated Impact

The anticipated impact of this change in the graduation component of the Postsecondary and Workforce Readiness Indicator of the performance frameworks is displayed in the attached worksheet. The worksheet shows that in the most recent 1-year district performance framework results, when using the best of the 4-, 5-, and 6-year on-time graduation rate recommended in this paper, only four districts end up with a different accreditation category than they would have when using the 2010 adjusted graduation rate. Three districts improve from Accredited with Improvement Plan to Accredited with Performance Plan as a result of higher PWR rating, and one district drops from Accredited with Performance Plan to Accredited with Improvement Plan as a result of a lower PWR rating.

A more comprehensive review of specific case studies, including example scenarios of schools that receive higher 4-year rates and schools that receive higher 6-year rates, is also included in the attached analytical brief.

Using the most recent 1-year DPF results, two different graduation rates were used to generate the Postsecondary and Workforce Readiness indicator and overall DPF plan types. The graduation rates used were: 1) the 2010 old adjusted grad rate vs. 2) the best of the new 4, 5, and 6 year grad rates. The chart below shows the four districts that had different plan types using the new graduation rate (Plan Type Change) and the fifteen districts that had different PWR Ratings using the new graduation rate (PWR Change).

					OLD ADJUSTED	NEW 4,5,6 YR			OLD ADJUSTED	NEW 4,5,6 YR	
emh	dist_nam	ACHIEVEMENT_ RATING	GROWTH_ RATING	GROWTH_GAPS_R ATING	PWR_RATING	PWR_RATING	PWR Change	TEST_PARTICIPATION_RATING	TOTAL_DPF_PLAN_TYPE	TOTAL_DPF_PLAN_TYPE	Plan Type Change
A	CENTENNIAL R-1	Does Not Meet	Meets	Meets	Approaching	Meets	***	95% Participation Rate Met	Improvement	Performance	***
A	HOEHNE REORGANIZED 3	Meets	Approaching	Approaching	Approaching	Meets	***	95% Participation Rate Met	Improvement	Performance	***
A	JOHNSTOWN-MILLIKEN RE-5J	Meets	Approaching	Approaching	Meets	Approaching	***	95% Participation Rate Met	Performance	Improvement	***
A	KEENESBURG RE-3(J)	Meets	Approaching	Approaching	Approaching	Meets	***	Does Not Meet 95% Participation Rate	Improvement	Performance	***
A	BRUSH RE-2(J)	Approaching	Meets	Meets	Approaching	Meets	***	95% Participation Rate Met	Performance	Performance	
A	CRIPPLE CREEK-VICTOR RE-1	Approaching	Approaching	Approaching	Approaching	Meets	***	95% Participation Rate Met	Improvement	Improvement	
A	GARFIELD 16	Approaching	Meets	Meets	Approaching	Meets	***	95% Participation Rate Met	Performance	Performance	
A	HUERFANO RE-1	Approaching	Approaching	Does Not Meet	Does Not Meet	Approaching	***	95% Participation Rate Met	Priority Improvement	Priority Improvement	
A	LAMAR RE-2	Approaching	Approaching	Approaching	Approaching	Meets	***	95% Participation Rate Met	Improvement	Improvement	
A	PARK COUNTY RE-2	Meets	Meets	Meets	Exceeds	Meets	***	95% Participation Rate Met	Performance	Performance	
A	RANGELY RE-4	Approaching	Approaching	Approaching	Approaching	Meets	***	95% Participation Rate Met	Improvement	Improvement	
A	ROARING FORK RE-1	Approaching	Meets	Meets	Approaching	Meets	***	95% Participation Rate Met	Performance	Performance	
A	SHERIDAN 2	Does Not Meet	Approaching	Approaching	Approaching	Does Not Meet	***	Does Not Meet 95% Participation Rate	Turnaround	Turnaround	
A	SOUTH CONEJOS RE-10	Does Not Meet	Does Not Meet	Does Not Meet	Approaching	Meets	***	95% Participation Rate Met	Priority Improvement	Priority Improvement	
A	WELD COUNTY RE-1	Approaching	Approaching	Approaching	Meets	Approaching	***	Does Not Meet 95% Participation Rate	Improvement	Improvement	
A	ACADEMY 20	Exceeds	Meets	Meets	Exceeds	Exceeds		95% Participation Rate Met	Performance	Performance	
A	ADAMS 12 FIVE STAR SCHOOLS	Approaching	Meets	Approaching	Approaching	Approaching		Does Not Meet 95% Participation Rate	Improvement	Improvement	
A	ADAMS COUNTY 14	Does Not Meet	Approaching	Approaching	Does Not Meet	Does Not Meet		Does Not Meet 95% Participation Rate	Turnaround	Turnaround	
A	ADAMS-ARAPAHOE 28J	Does Not Meet	Approaching	Approaching	Does Not Meet	Does Not Meet		Does Not Meet 95% Participation Rate	Turnaround	Turnaround	
A	AGATE 300				Exceeds	Exceeds		95% Participation Rate Met	Performance	Performance	
A	AGUILAR REORGANIZED 6	Approaching	Does Not Meet		Meets	Meets		95% Participation Rate Met	Improvement	Improvement	
A	AKRON R-1	Meets	Approaching	Approaching	Exceeds	Exceeds		95% Participation Rate Met	Performance	Performance	
A	ALAMOSA RE-11J	Approaching	Approaching	Approaching	Approaching	Approaching		95% Participation Rate Met	Improvement	Improvement	
A	ARCHULETA COUNTY 50 JT	Meets	Meets	Meets	Approaching	Approaching		95% Participation Rate Met	Performance	Performance	
A	ARICKAREE R-2	Exceeds	Approaching		Approaching	Approaching		95% Participation Rate Met	Improvement	Improvement	
A	ARRIBA-FLAGLER C-20	Meets	Does Not Meet	Does Not Meet	Meets	Exceeds		95% Participation Rate Met	Improvement	Improvement	
A	ASPEN 1	Exceeds	Meets	Meets	Exceeds	Exceeds		95% Participation Rate Met	Performance	Performance	
A	AULT-HIGHLAND RE-9	Approaching	Approaching	Approaching	Meets	Meets		95% Participation Rate Met	Improvement	Improvement	
A	BAYFIELD 10 JT-R	Meets	Meets	Approaching	Meets	Meets		95% Participation Rate Met	Performance	Performance	
A	BENNETT 29J	Meets	Approaching	Approaching	Meets	Meets		95% Participation Rate Met	Performance	Performance	
A	BETHUNE R-5	Approaching	Approaching		Meets	Meets		95% Participation Rate Met	Performance	Performance	
A	BIG SANDY 100J	Meets	Approaching	Approaching	Meets	Meets		95% Participation Rate Met	Performance	Performance	
A	BOULDER VALLEY RE 2	Meets	Meets	Approaching	Meets	Meets		95% Participation Rate Met	Performance	Performance	
A	BRANSON REORGANIZED 82	Meets	Approaching	Approaching	Approaching	Approaching		95% Participation Rate Met	Improvement	Improvement	
A	BRIGGSDALE RE-10	Meets	Meets		Exceeds	Exceeds		95% Participation Rate Met	Performance	Performance	
A	BRIGHTON 27J	Approaching	Approaching	Approaching	Approaching	Approaching		95% Participation Rate Met	Improvement	Improvement	
A	BUENA VISTA R-31	Meets	Meets	Approaching	Meets	Meets		95% Participation Rate Met	Performance	Performance	
A	BUFFALO RE-4	Meets	Approaching	Approaching	Exceeds	Exceeds		Does Not Meet 95% Participation Rate	Performance	Performance	
A	BURLINGTON RE-6J	Approaching	Approaching	Approaching	Meets	Meets		95% Participation Rate Met	Improvement	Improvement	
A	BYERS 32J	Meets	Meets	Approaching	Meets	Meets		95% Participation Rate Met	Performance	Performance	
A	CALHAN RJ-1	Meets	Meets	Approaching	Meets	Meets		95% Participation Rate Met	Performance	Performance	
A	CAMPO RE-6				Meets	Meets		95% Participation Rate Met	Performance	Performance	
A	CANON CITY RE-1	Meets	Approaching	Approaching	Approaching	Approaching		95% Participation Rate Met	Improvement	Improvement	
A	CENTER 26 JT	Does Not Meet	Approaching	Approaching	Approaching	Approaching		95% Participation Rate Met	Improvement	Improvement	
A	CHARTER SCHOOL INSTITUTE	Approaching	Meets	Approaching	Approaching	Approaching		95% Participation Rate Met	Improvement	Improvement	
A	CHERAW 31	Approaching	Approaching	Approaching	Exceeds	Exceeds		95% Participation Rate Met	Performance	Performance	
A	CHERRY CREEK 5	Meets	Meets	Approaching	Meets	Meets		95% Participation Rate Met	Performance	Performance	
A	CHEYENNE COUNTY RE-5	Approaching	Meets		Exceeds	Exceeds		95% Participation Rate Met	Performance	Performance	
A	CHEYENNE MOUNTAIN 12	Exceeds	Meets	Meets	Exceeds	Exceeds		95% Participation Rate Met	Performance	Performance	
A	CLEAR CREEK RE-1	Meets	Meets	Meets	Meets	Meets		95% Participation Rate Met	Performance	Performance	
A	Colorado School for the Deaf and	Does Not Meet	Approaching	Approaching				95% Participation Rate Met	Priority Improvement	Priority Improvement	
A	COLORADO SPRINGS 11	Meets	Approaching	Approaching	Approaching	Approaching		95% Participation Rate Met	Improvement	Improvement	
A	COTOPAXI RE-3	Meets	Approaching		Meets	Meets		95% Participation Rate Met	Performance	Performance	
A	CREEDE SCHOOL DISTRICT	Exceeds			Exceeds	Exceeds		95% Participation Rate Met	Performance	Performance	

					OLD ADJUSTED	NEW 4,5,6 YR			OLD ADJUSTED	NEW 4,5,6 YR	
emh	dist_nam	ACHIEVEMENT_ RATING	GROWTH_ RATING	GROWTH_GAPS_R ATING	PWR_RATING	PWR_RATING	PWR Change	TEST_PARTICIPATION_RATING	TOTAL_DPF_PLAN_TYPE	TOTAL_DPF_PLAN_TYPE	Plan Type Change
A	CROWLEY COUNTY RE-1-J	Approaching	Approaching	Approaching	Meets	Meets		95% Participation Rate Met	Improvement	Improvement	
A	CUSTER COUNTY SCHOOL DISTRICT C-1	Meets	Approaching	Approaching	Meets	Meets		95% Participation Rate Met	Performance	Performance	
A	DE BEQUE 49JT	Meets	Meets		Does Not Meet	Does Not Meet		95% Participation Rate Met	Improvement	Improvement	
A	DEER TRAIL 26J	Meets	Approaching	Approaching	Exceeds	Exceeds		95% Participation Rate Met	Performance	Performance	
A	DEL NORTE C-7	Approaching	Approaching	Approaching	Meets	Meets		95% Participation Rate Met	Performance	Performance	
A	DELTA COUNTY 50(J)	Meets	Meets	Meets	Meets	Meets		95% Participation Rate Met	Performance	Performance	
A	DENVER COUNTY 1	Does Not Meet	Approaching	Approaching	Approaching	Approaching		95% Participation Rate Met	Improvement	Improvement	
A	DOLORES COUNTY RE NO.2	Approaching	Approaching	Does Not Meet	Meets	Meets		95% Participation Rate Met	Improvement	Improvement	
A	DOLORES RE-4A	Meets	Meets	Approaching	Meets	Meets		95% Participation Rate Met	Performance	Performance	
A	DOUGLAS COUNTY RE 1	Meets	Meets	Approaching	Meets	Meets		95% Participation Rate Met	Performance	Performance	
A	DURANGO 9-R	Meets	Meets	Approaching	Meets	Meets		95% Participation Rate Met	Performance	Performance	
A	EADS RE-1	Meets	Approaching		Exceeds	Exceeds		95% Participation Rate Met	Performance	Performance	
A	EAGLE COUNTY RE 50	Meets	Meets	Meets	Approaching	Approaching		95% Participation Rate Met	Performance	Performance	
A	EAST GRAND 2	Meets	Meets	Approaching	Meets	Meets		95% Participation Rate Met	Performance	Performance	
A	EAST OTERO R-1	Approaching	Meets	Meets	Approaching	Approaching		95% Participation Rate Met	Performance	Performance	
A	EATON RE-2	Meets	Meets	Approaching	Meets	Meets		95% Participation Rate Met	Performance	Performance	
A	EDISON 54 JT	Meets	Meets		Meets	Meets		95% Participation Rate Met	Performance	Performance	
A	ELBERT 200	Meets	Meets	Meets	Exceeds	Exceeds		95% Participation Rate Met	Performance	Performance	
A	ELIZABETH C-1	Meets	Meets	Approaching	Meets	Meets		95% Participation Rate Met	Performance	Performance	
A	ELLICOTT 22	Approaching	Approaching	Approaching	Meets	Meets		95% Participation Rate Met	Performance	Performance	
A	ENGLEWOOD 1	Does Not Meet	Approaching	Approaching	Approaching	Approaching		Does Not Meet 95% Participation Rate	Turnaround	Turnaround	
A	EXPEDITIONARY BOCES	Meets	Meets	Meets	Exceeds	Exceeds		95% Participation Rate Met	Performance	Performance	
A	FALCON 49	Meets	Meets	Approaching	Meets	Meets		95% Participation Rate Met	Performance	Performance	
A	FLORENCE RE-2	Approaching	Approaching	Approaching	Meets	Meets		95% Participation Rate Met	Improvement	Improvement	
A	FORT MORGAN RE-3	Does Not Meet	Approaching	Approaching	Approaching	Approaching		95% Participation Rate Met	Improvement	Improvement	
A	FOUNTAIN 8	Meets	Approaching	Approaching	Meets	Meets		95% Participation Rate Met	Improvement	Improvement	
A	FOWLER R-4J	Meets	Meets	Meets	Meets	Meets		95% Participation Rate Met	Performance	Performance	
A	FRENCHMAN RE-3	Meets	Approaching	Does Not Meet	Exceeds	Exceeds		95% Participation Rate Met	Performance	Performance	
A	GARFIELD RE-2	Approaching	Approaching	Approaching	Approaching	Approaching		95% Participation Rate Met	Improvement	Improvement	
A	GENOA-HUGO C113	Approaching	Approaching		Exceeds	Exceeds		95% Participation Rate Met	Performance	Performance	
A	GILPIN COUNTY RE-1	Meets	Meets	Approaching	Meets	Meets		95% Participation Rate Met	Performance	Performance	
A	GRANADA RE-1	Approaching	Approaching	Approaching	Exceeds	Exceeds		95% Participation Rate Met	Performance	Performance	
A	GREELEY 6	Approaching	Approaching	Approaching	Approaching	Approaching		95% Participation Rate Met	Improvement	Improvement	
A	GUNNISON WATERSHED RE1J	Meets	Meets	Approaching	Meets	Meets		95% Participation Rate Met	Performance	Performance	
A	HANOVER 28	Approaching	Approaching	Does Not Meet	Exceeds	Exceeds		Does Not Meet 95% Participation Rate	Improvement	Improvement	
A	HARRISON 2	Approaching	Meets	Approaching	Approaching	Approaching		95% Participation Rate Met	Improvement	Improvement	
A	HAXTUN RE-2J	Approaching	Meets	Meets	Exceeds	Exceeds		95% Participation Rate Met	Performance	Performance	
A	HAYDEN RE-1	Approaching	Approaching	Approaching	Meets	Meets		95% Participation Rate Met	Improvement	Improvement	
A	HINSDALE COUNTY RE 1	Exceeds			Exceeds	Exceeds		95% Participation Rate Met	Performance	Performance	
A	HI-PLAINS R-23	Approaching	Approaching		Exceeds	Exceeds		95% Participation Rate Met	Performance	Performance	
A	HOLLY RE-3	Approaching	Meets	Meets	Meets	Meets		95% Participation Rate Met	Performance	Performance	
A	HOLYOKE RE-1J	Approaching	Approaching	Approaching	Meets	Meets		95% Participation Rate Met	Improvement	Improvement	
A	IDALIA RJ-3	Approaching	Approaching		Meets	Meets		95% Participation Rate Met	Performance	Performance	
A	IGNACIO 11 JT	Does Not Meet	Does Not Meet	Does Not Meet	Approaching	Approaching		95% Participation Rate Met	Priority Improvement	Priority Improvement	
A	JEFFERSON COUNTY R-1	Meets	Meets	Approaching	Meets	Meets		95% Participation Rate Met	Performance	Performance	
A	JULESBURG RE-1	Meets	Meets	Approaching	Does Not Meet	Does Not Meet		Does Not Meet 95% Participation Rate	Improvement	Improvement	
A	KARVAL RE-23	Does Not Meet	Approaching	Approaching	Approaching	Approaching		95% Participation Rate Met	Priority Improvement	Priority Improvement	
A	KIM REORGANIZED 88				Exceeds	Exceeds		95% Participation Rate Met	Performance	Performance	
A	KIOWA C-2	Meets	Meets	Meets	Exceeds	Exceeds		95% Participation Rate Met	Performance	Performance	
A	KIT CARSON R-1	Meets	Does Not Meet		Exceeds	Exceeds		95% Participation Rate Met	Performance	Performance	
A	LA VETA RE-2	Meets	Approaching	Approaching	Exceeds	Exceeds		95% Participation Rate Met	Performance	Performance	
A	LAKE COUNTY R-1	Does Not Meet	Approaching	Approaching	Approaching	Approaching		95% Participation Rate Met	Priority Improvement	Priority Improvement	
A	LAS ANIMAS RE-1	Approaching	Approaching	Approaching	Meets	Meets		Does Not Meet 95% Participation Rate	Improvement	Improvement	
A	LEWIS-PALMER 38	Exceeds	Meets	Meets	Exceeds	Exceeds		95% Participation Rate Met	Performance	Performance	
A	LIBERTY J-4	Approaching			Exceeds	Exceeds		95% Participation Rate Met	Performance	Performance	
A	LIMON RE-4J	Meets	Meets	Meets	Meets	Meets		95% Participation Rate Met	Performance	Performance	

					OLD ADJUSTED	NEW 4,5,6 YR			OLD ADJUSTED	NEW 4,5,6 YR	
emh	dist_nam	ACHIEVEMENT_ RATING	GROWTH_ RATING	GROWTH_GAPS_R ATING	PWR_RATING	PWR_RATING	PWR Change	TEST_PARTICIPATION_RATING	TOTAL_DPF_PLAN_TYPE	TOTAL_DPF_PLAN_TYPE	Plan Type Change
A	LITTLETON 6	Meets	Meets	Approaching	Exceeds	Exceeds		95% Participation Rate Met	Performance	Performance	
A	LONE STAR 101	Meets	Approaching		Exceeds	Exceeds		95% Participation Rate Met	Performance	Performance	
A	MANCOS RE-6	Meets	Meets	Approaching	Meets	Meets		95% Participation Rate Met	Performance	Performance	
A	MANITOU SPRINGS 14	Meets	Approaching	Approaching	Exceeds	Exceeds		95% Participation Rate Met	Performance	Performance	
A	MANZANOLA 3J	Approaching	Approaching	Approaching	Meets	Meets		Does Not Meet 95% Participation Rate	Improvement	Improvement	
A	MAPLETON 1	Does Not Meet	Meets	Approaching	Approaching	Approaching		95% Participation Rate Met	Improvement	Improvement	
A	MC CLAVE RE-2	Meets	Meets	Approaching	Meets	Meets		95% Participation Rate Met	Performance	Performance	
A	MEEKER RE1	Meets	Meets	Meets	Meets	Meets		95% Participation Rate Met	Performance	Performance	
A	MESA COUNTY VALLEY 51	Meets	Meets	Approaching	Approaching	Approaching		95% Participation Rate Met	Performance	Performance	
A	MIAMI/YODER 60 JT	Approaching	Approaching	Approaching	Meets	Meets		95% Participation Rate Met	Performance	Performance	
A	MOFFAT 2	Approaching	Approaching	Approaching	Exceeds	Exceeds		95% Participation Rate Met	Performance	Performance	
A	MOFFAT COUNTY RE:NO 1	Approaching	Approaching	Approaching	Meets	Meets		95% Participation Rate Met	Improvement	Improvement	
A	MONTE VISTA C-8	Approaching	Approaching	Approaching	Approaching	Approaching		95% Participation Rate Met	Improvement	Improvement	
A	MONTEZUMA-CORTEZ RE-1	Approaching	Meets	Approaching	Does Not Meet	Does Not Meet		95% Participation Rate Met	Improvement	Improvement	
A	MONTROSE COUNTY RE-1J	Approaching	Approaching	Approaching	Approaching	Approaching		95% Participation Rate Met	Improvement	Improvement	
A	MOUNTAIN BOCES	Does Not Meet	Does Not Meet	Approaching	Does Not Meet	Does Not Meet		Does Not Meet 95% Participation Rate	Turnaround	Turnaround	
A	MOUNTAIN VALLEY RE 1	Approaching			Exceeds	Exceeds		95% Participation Rate Met	Performance	Performance	
A	NORTH CONEJOS RE-1J	Meets	Approaching	Approaching	Meets	Meets		95% Participation Rate Met	Performance	Performance	
A	NORTH PARK R-1	Meets	Meets		Meets	Meets		95% Participation Rate Met	Performance	Performance	
A	NORTHWEST COLO BOCES				Exceeds	Exceeds		95% Participation Rate Met	Performance	Performance	
A	NORWOOD R-2J	Meets	Approaching	Meets	Meets	Meets		95% Participation Rate Met	Performance	Performance	
A	OTIS R-3	Meets	Meets		Exceeds	Exceeds		95% Participation Rate Met	Performance	Performance	
A	OURAY R-1	Meets	Exceeds		Meets	Meets		95% Participation Rate Met	Performance	Performance	
A	PARK (ESTES PARK) R-3	Meets	Meets	Meets	Meets	Meets		95% Participation Rate Met	Performance	Performance	
A	PAWNEE RE-12	Meets			Exceeds	Exceeds		95% Participation Rate Met	Performance	Performance	
A	PEYTON 23 JT	Meets	Meets	Approaching	Exceeds	Exceeds		95% Participation Rate Met	Performance	Performance	
A	PLAINVIEW RE-2	Approaching			Exceeds	Exceeds		95% Participation Rate Met	Performance	Performance	
A	PLATEAU RE-5	Exceeds	Approaching		Exceeds	Exceeds		95% Participation Rate Met	Performance	Performance	
A	PLATEAU VALLEY 50	Approaching	Meets	Meets	Does Not Meet	Does Not Meet		95% Participation Rate Met	Improvement	Improvement	
A	PLATTE CANYON 1	Exceeds	Meets	Meets	Meets	Meets		95% Participation Rate Met	Performance	Performance	
A	PLATTE VALLEY RE-3	Meets	Approaching	Approaching	Exceeds	Exceeds		95% Participation Rate Met	Performance	Performance	
A	PLATTE VALLEY RE-7	Approaching	Meets	Meets	Meets	Meets		95% Participation Rate Met	Performance	Performance	
A	POUDRE R-1	Meets	Meets	Approaching	Meets	Meets		95% Participation Rate Met	Performance	Performance	
A	PRAIRIE RE-11	Meets	Meets		Meets	Meets		95% Participation Rate Met	Performance	Performance	
A	PRIMERO REORGANIZED 2	Approaching	Meets		Meets	Meets		95% Participation Rate Met	Performance	Performance	
A	PRITCHETT RE-3				Meets	Meets		95% Participation Rate Met	Performance	Performance	
A	PUEBLO CITY 60	Approaching	Approaching	Does Not Meet	Approaching	Approaching		95% Participation Rate Met	Priority Improvement	Priority Improvement	
A	PUEBLO COUNTY 70	Meets	Approaching	Approaching	Approaching	Approaching		95% Participation Rate Met	Improvement	Improvement	
A	RIDGWAY R-2	Exceeds	Meets	Meets	Exceeds	Exceeds		95% Participation Rate Met	Performance	Performance	
A	ROCKY FORD R-2	Does Not Meet	Approaching	Does Not Meet	Approaching	Approaching		95% Participation Rate Met	Priority Improvement	Priority Improvement	
A	SALIDA R-32	Meets	Approaching	Approaching	Meets	Meets		95% Participation Rate Met	Performance	Performance	
A	SANFORD 6J	Meets	Meets	Approaching	Meets	Meets		95% Participation Rate Met	Performance	Performance	
A	SANGRE DE CRISTO RE-22J	Approaching	Approaching	Approaching	Meets	Exceeds		95% Participation Rate Met	Performance	Performance	
A	SARGENT RE-33J	Meets	Approaching	Meets	Exceeds	Exceeds		95% Participation Rate Met	Performance	Performance	
A	SIERRA GRANDE R-30	Does Not Meet	Approaching	Approaching	Exceeds	Exceeds		95% Participation Rate Met	Performance	Performance	
A	SILVERTON 1							95% Participation Rate Met	NA	NA	
A	SOUTH ROUTT RE 3	Meets	Approaching	Approaching	Meets	Meets		95% Participation Rate Met	Improvement	Improvement	
A	SPRINGFIELD RE-4	Meets	Approaching	Approaching	Exceeds	Exceeds		95% Participation Rate Met	Performance	Performance	
A	ST VRAIN VALLEY RE 1J	Meets	Meets	Approaching	Meets	Meets		95% Participation Rate Met	Performance	Performance	
A	STEAMBOAT SPRINGS RE-2	Exceeds	Meets	Meets	Exceeds	Exceeds		95% Participation Rate Met	Performance	Performance	
A	STRASBURG 31J	Meets	Meets	Meets	Meets	Meets		95% Participation Rate Met	Performance	Performance	
A	STRATTON R-4	Approaching	Approaching	Approaching	Meets	Meets		95% Participation Rate Met	Performance	Performance	
A	SUMMIT RE-1	Meets	Approaching	Approaching	Meets	Meets		95% Participation Rate Met	Performance	Performance	
A	SWINK 33	Meets	Meets	Meets	Exceeds	Exceeds		95% Participation Rate Met	Performance	Performance	
A	TELLURIDE R-1	Exceeds	Exceeds	Meets	Exceeds	Exceeds		95% Participation Rate Met	Performance	Performance	
A	THOMPSON R-2J	Meets	Meets	Approaching	Meets	Meets		95% Participation Rate Met	Performance	Performance	

					OLD ADJUSTED	NEW 4,5,6 YR			OLD ADJUSTED	NEW 4,5,6 YR	
emh	dist_nam	ACHIEVEMENT_ RATING	GROWTH_ RATING	GROWTH_GAPS_R ATING	PWR_RATING	PWR_RATING	PWR Change	TEST_PARTICIPATION_RATING	**TOTAL_DPF_PLAN_TYPE**	**TOTAL_DPF_PLAN_TYPE**	Plan Type Change
A	TRINIDAD 1	Approaching	Approaching	Approaching	Meets	Meets		95% Participation Rate Met	**Improvement**	**Improvement**	
A	VALLEY RE-1	Meets	Approaching	Approaching	Approaching	Approaching		95% Participation Rate Met	**Improvement**	**Improvement**	
A	VILAS RE-5	Approaching	Does Not Meet	Does Not Meet	Does Not Meet	Does Not Meet		Does Not Meet 95% Participation Rate	**Turnaround**	**Turnaround**	
A	WALSH RE-1	Meets	Approaching		Meets	Meets		95% Participation Rate Met	**Performance**	**Performance**	
A	WELD COUNTY S/D RE-8	Does Not Meet	Meets	Approaching	Approaching	Approaching		95% Participation Rate Met	**Improvement**	**Improvement**	
A	WELDON VALLEY RE-20(J)	Meets	Meets		Exceeds	Exceeds		95% Participation Rate Met	**Performance**	**Performance**	
A	WEST END RE-2	Approaching	Approaching	Approaching	Meets	Meets		95% Participation Rate Met	**Improvement**	**Improvement**	
A	WEST GRAND 1-JT.	Meets	Meets	Meets	Meets	Meets		95% Participation Rate Met	**Performance**	**Performance**	
A	WESTMINSTER 50	Does Not Meet	Does Not Meet	Does Not Meet	Approaching	Approaching		95% Participation Rate Met	**Priority Improvement**	**Priority Improvement**	
A	WIDEFIELD 3	Approaching	Approaching	Approaching	Meets	Meets		95% Participation Rate Met	**Improvement**	**Improvement**	
A	WIGGINS RE-50(J)	Meets	Meets	Approaching	Meets	Meets		95% Participation Rate Met	**Performance**	**Performance**	
A	WILEY RE-13 JT	Meets	Meets	Approaching	Exceeds	Exceeds		95% Participation Rate Met	**Performance**	**Performance**	
A	WINDSOR RE-4	Meets	Meets	Approaching	Meets	Meets		95% Participation Rate Met	**Performance**	**Performance**	
A	WOODLAND PARK RE-2	Meets	Approaching	Approaching	Meets	Meets		Does Not Meet 95% Participation Rate	**Performance**	**Performance**	
A	WOODLIN R-104	Approaching			Exceeds	Exceeds		95% Participation Rate Met	**Performance**	**Performance**	
A	WRAY RD-2	Meets	Meets	Approaching	Meets	Meets		95% Participation Rate Met	**Performance**	**Performance**	
A	YUMA 1	Approaching	Meets	Meets	Meets	Meets		95% Participation Rate Met	**Performance**	**Performance**	

Addressing Misconceptions of Bias in the Performance Frameworks

March 10, 2011

Colorado Dept. of Education
Research and Evaluation Unit

William Bonk
Marie Huchton

Addressing Misconceptions of Bias in the Performance Frameworks

In some recent papers and presentations, the Poudre School District's Research and Evaluation Unit has maintained that Colorado's District and School Performance Frameworks are biased against large schools and districts and that the state's accountability system is therefore inconsistent and unfair.

As a key component of its implementation of the Education Accountability Act of 2009, the Colorado Department of Education created the state Performance Frameworks. Through a comprehensive and evenhanded public process, the department solicited and incorporated input from the field during all phases of the design and preliminary reporting. CDE contends that the Performance Frameworks resulting from this substantial process are fair, equitable, and that above all, they represent a successful collective effort to create the tension needed in the system to drive insight and improvement, while balancing the many legitimate concerns of stakeholders across the state. This response will address the concerns raised by Poudre School District, showing that their claims of bias are based on false premises, a flawed statistical argument, and an incorrect understanding of how the CDE's measure of adequate growth influences overall framework scores.

There is broad agreement across the state that students' academic growth needs to be a part of an effective educational accountability system. It is also clear that, if college and career readiness is to be the goal of the public school system, both students' level of proficiency as well as the level of growth needed to bring them to and keep them at proficiency must be examined. It is with these principles in mind that the Performance Frameworks were designed - and adjusted - in consultation with many stakeholder groups. The department does not feel that Poudre's Research and Evaluation Unit has adequately provided any solid evidence to suggest that that the use of median growth percentiles, or the inclusion of adequate growth in the frameworks, in any way discriminates against large schools or districts.

Henry Braun, in discussing the consequential validity of accountability systems, wrote the following wise words: "Assessment practices and systems of accountability are consequentially valid if they generate useful information and constructive responses that support one or more policy goals (Access, Quality, Equity, Efficiency) within an education system, without causing undue deterioration with respect to other goals." Does Colorado's new accountability system adhere to this idea? Does it produce results that promote the desired kinds of conversations around performance management and improvement? Does the Colorado Growth Model promote the coherence and alignment that we need? The answers that we have heard to these questions have been a resounding, "Yes!" across the state.

Using the student growth percentile method entails making some choices; although the Colorado Growth Model may not answer every question, it does give us answers to more of the right questions than any other method. While we appreciate the critical eye with which Poudre School District's Research and Evaluation Unit has looked at the output of the performance frameworks in 2010 in the interest of helping the state create an equitable accountability system, the department respectfully disagrees with their conclusions and invite them to reconsider their position.

Claim: Median Growth Percentiles are biased

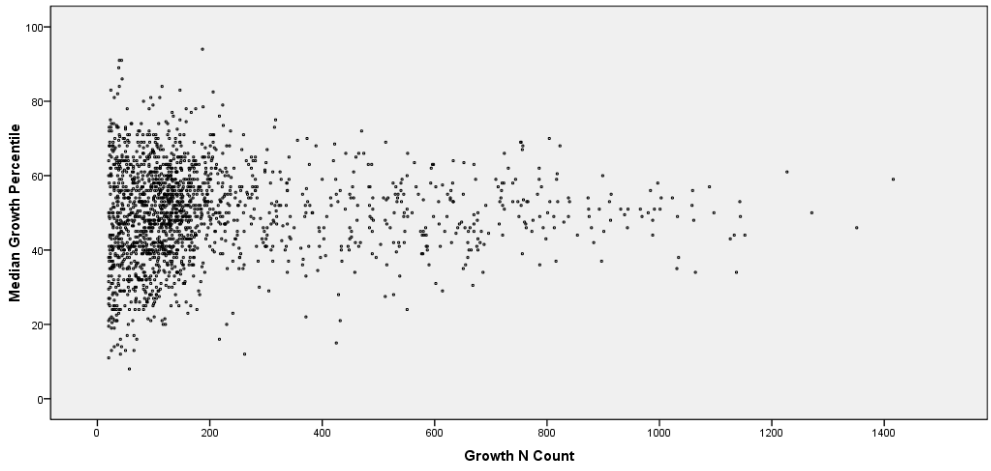
Does the Colorado Growth Model make it nearly impossible for large schools to get high growth? Poudre School District's claim is that the Performance Frameworks are biased against large districts and schools because, regardless of their actual level of effectiveness, high median growth percentiles (MGPs) are virtually impossible for large districts and schools to attain. This idea of a size-based bias is demonstrably incorrect in several ways.

First, when analyzing data, more data always lead to a sounder analysis. By definition, growth in larger schools is based on more observations, so their MGPs are highly accurate descriptions of their students' overall academic progress. The school's MGPs are also less affected by other variables than those based on a smaller number of observations. For example, consider the effect of a number of students receiving private tutoring outside of school and, as a result, getting higher scores on state tests. This outcome could have a significant effect on a small school's numbers, but would not be as noticeable in those of a large school. From a purely statistical point of view, there is no reason that more observations should lead to a less accurate MGP value. On the contrary, more data always lead to a more precise summary statistic like the MGP. Large schools are not penalized just because the n-count is higher.

Growth percentiles are, however, a normative metric, with a restricted range of values. When using growth percentiles, do large schools necessarily get the "average" because they make up so much of the total that they themselves *define* that average? Actually, the largest school in the data plot presented by Poudre School District in their Figure 1 had 1,699 student growth records, out of a total of 100,491 high school students receiving math growth percentiles in that year. This school represents a mere 1.7 percent of the state's total n-count, so the idea that the largest school is big enough to define the average by itself lacks credibility.

The Poudre district's paper also implies that small schools get median growth percentiles they don't really deserve, asserting that small schools' MGPs are driven by a random statistical process rather than what happens between its teachers and students. In support of this argument, the paper points out the greater amount of school-to-school variability among schools with small n-counts. This justification is misleading, however, because Figure 1 in Poudre's position paper includes medians for schools with n<20, and it does not calculate medians separately for elementary, middle and high schools (as is the practice in public reporting and in the state's accountability system). When properly plotted, the pattern of variability by n-count is not nearly as extreme as in Poudre's plot, and is shown in Figure 1:

Figure 1. Math MGPs for Schools by EMH Academic Year 2006-2007.



In Figure 1 there is indeed a diminishing of the school-to-school variability in MGPs as n-count increases beyond 200 or so. (There is also a very sharp decline in the number of schools with greater than 250

to 300 students, so it is not easy to tell how stable this pattern truly is. Poudre's paper does not address how sampling theory might apply in the case of this sample of schools and their growth data.) As those who study educational data know, school size is correlated with a large number of other variables unmeasured in this model - such as urban setting and poverty - that tend to be associated with lower academic performance. There is no straightforward way of isolating out the pure effect of n-count on growth given the effect of these other variables and the small number of schools with large n counts. Thus, it should not be assumed that the existence of a few data points near the 50th percentile for schools or districts of a given size constitute a proof that this is as high or as low as they will ever get. Poudre School District's paper seems to suggest that its readers glance at the plot and make a judgment about what is going on, when in reality the purpose of statistics is to keep us from jumping to conclusions on the basis of mere perception. Is student n-count the foremost characteristic determining observed growth at the two largest schools in the plot above, Cherry Creek High School and Prairie Middle School? Is the primary driver of their MGPs just a statistical lottery based on their size?

This way of thinking also has implications for the case of small schools. Peetz Junior-Senior High School in Plateau Re-5 School District earned a reading median of 95, with a total three-year n-count of 59, on their 2010 School Performance Framework results. Poudre's reasoning would suggest that Peetz may not have deserved this high median, that it had been "driven by a statistical phenomenon rather than some characteristics of educational pedagogy, adopted curriculum or other subject-specific quality." By how much would Poudre School District suggest reducing Peetz's MGP? Similarly, what would Poudre suggest be done about the school with a three-year n-count of 23 and observed MGP of 12 in reading? Would Poudre recommend adjusting this school's MGP, because as a small school it probably didn't deserve such a low one? Such a practice is indefensible for a functional accountability system. The fact is that the MGP is the best estimate we have of the students' academic growth in these schools because it is based on all the pertinent observed data**.**

There is undoubtedly an influence of other, unmeasured variables likely contributing to the wide spread of values observed in MGPs with small n-counts (the pattern of variability that Poudre's paper bases some of its argument on). As datasets get smaller, the effects of unmeasured variables tend to not cancel themselves out as readily. However, it is simply not appropriate to apply an n-based correction when we are not sure to what extent any particular school's data *was actually subject* to the influence of those variables. What is fair is to use pertinent, available data and build a body of evidence, rather than apply some statistical fix of questionable appropriateness.

The Poudre paper also argues that the shape of the MGP distribution for schools of various sizes (see Figure 1) is a statistical phenomenon, writing that "we are sampling from a student population when we are calculating median student growth percentiles and then making the claim that the estimate produced represents the typical level of normative growth attained by the students of a specific school or district" (page 4). We disagree wholeheartedly with this assertion. Although the district is correct in stating that other (unmeasured) variables' effects will tend to not cancel each other out as n-counts get smaller, this still does not mean that we are taking random samples of student growth percentiles from schools or districts. The MGPs are descriptive measures of the growth that was observed among *all* the students belonging to a particular group such as a grade within a school (the vast majority of students with growth percentiles in most cases). Creating an MGP is not an attempt to estimate the "true" value for some population from which we have taken a flawed, random sample, like in opinion polling – **this already *is* the population**. We are not sampling, and the standard practices around sampling do not apply.

When describing data, care must be taken to use language carefully. The Poudre report stated that "smaller schools are much more likely to obtain median student growth percentiles that exceed 60 as well as fall below 40." This should not be taken to suggest that "likelihood" implies probability in the conventional sense of the term. Schools are not rolling the dice for their MGPs, and MGPs are not assigned randomly with a magnitude dependent on school size. MGPs are based on the observed test scores and growth percentiles from the students within a school.

There is unmistakably *something* happening in a school or district (instructional practice, among other things) that creates meaningful variation in growth percentiles among students, and that *something* is worth measuring. The size of the influence of unmeasured variables on MGPs cannot be estimated by sampling theory alone, because the assumptions underlying it are not met in these data. It is not prudent to assume that n-count is the primary driver for differences in MGPs among schools and districts, and that sampling theory can "fix" the problem. The typical approach would be the creation of standard errors, essentially a guess: "This is how low and how high we think this school's MGP would likely have come out if we had drawn repeated random samples of a given size." But the use of MGPs based on the set of actual observations is more defensible than some guess based on an imaginary super-population, using an n-count-based adjustment.

As the theory and practice around these growth percentiles continues, an improved approach to understanding the relationship between MGPs and n-count may indeed emerge. However, there is no evidence that the growth percentile metric is systematically biased for or against any schools, large or small. What has been presented is an argument resting on incorrect assumptions: it presupposes the estimation of a population parameter from a sample when there is none; it assumes random sampling when there is none; it assumes that there should be equal variation in growth among schools of different sizes when there are readily understood reasons for that not to be the case.

Claim: School size must be taken into account in growth calculations[1]

Does state law require that school size be taken into account? The idea of considering school size in growth calculations (cited in Poudre's paper on page 3) is a part of the rules for giving the Governor's Distinguished Improvement Awards, not for the accountability system in general. The text of the law[2] is presented below

(1) The state board shall annually present financial awards to the public schools in the state demonstrating the highest rates of student longitudinal growth as measured by the Colorado growth model. The technical advisory panel convened pursuant to section 22-11-202 shall recommend to the state board, and the state board shall establish by rule, the method by which to identify schools that demonstrate the highest rate of student longitudinal growth in a school year, as measured by the Colorado growth model. The technical advisory panel shall take school size into account in preparing its recommendations.
(2) Of the moneys available for awards pursuant to this part 6, two thirds shall be awarded pursuant to this section.
(3) An award issued pursuant to this section shall be known as a "Governor's Distinguished Improvement Award".

Because the Governor's Distinguished Improvement Awards were intended to be financial, school size was an important consideration so that the awards would be provided on a per pupil basis, rather than as a set sum of money regardless of how many students would benefit from it. There was no intention to make school size a consideration for growth an explicit part of the accountability system, either in the original law (HB 07-1048) or in the language from the law that was re-enabled in the Education Accountability Act of 2009 (SB 09-163). Thus, the idea that it is "necessary to take school and district size into account when making judgments about how unusual student growth is" (page 3) is not a part of the state law that guides accountability.

Ultimately, creating a Performance Framework that fairly deals with all of Colorado's districts and schools, in their great diversity of size and makeup, is a challenge. The way that the system should be judged is on whether or not it helps focus attention on performance in key academic areas and fosters useful conversations around improvement and best practice. The department contends that the evidence is clear from the distribution of category assignments that an equitable distribution has been achieved, which has in turn been borne out by feedback received from districts about the accreditation category and school plan type assignments based on these frameworks.

[1] This claim appears differently in the latest version of Poudre's document. Nevertheless, the response provided here applies to either a strong or weak version of the claim.

[2] http://www.michie.com/colorado/lpext.dll?f=FifLink&t=document-frame.htm&l=query&iid=3998343b.453b7660.0.0&q=%5BGroup%20%2722-11-603%27%5D

Claim: Putting adequate growth into the Performance Frameworks pollutes the "pure growth" calculations and gets us back to status measures.

Does using different cut scores for growth based on adequate growth unfairly penalize large, diverse schools and districts? The reason for including median adequate growth percentiles (AGPs) in the Performance Frameworks was to look at the MGPs with respect to the state's proficiency goals for its students. It is common knowledge that many students are not at the state's required level of proficiency – but is there any evidence that they are getting there? Adequate growth identifies the level of growth needed, on average, for a given school's students to be reaching and maintaining the minimal level of proficiency in that specific content area—in other words "catching up" and "keeping up." The Colorado Growth Model helps calculate these numbers, but they are entirely different from the growth calculations themselves. The purpose of an AGP is to measure growth to a standard – was the observed growth (low, moderate or high) sufficient to get students to the destination of proficiency? How much growth was made, and was that growth enough?

Assessment data can be viewed in several ways: (1) the familiar status calculation – by group, what is the percentage of students that has already achieved proficiency? (2) using growth only – how was these students' academic progress relative to that of their academic peers? (3) using both growth and status information – how much growth do students need to demonstrate that they are on track to attaining or maintaining proficiency? For example, a student scoring in the Unsatisfactory level and making only 10th percentile growth in the last year is clearly not on track to make up the lost ground and get to proficiency in the coming years. The Adequate Growth calculations tell us about all the students in the state, not just the obvious cases like this one.

If you're only interested in question (2), then you don't need to ask the other two. The implication of this way of thinking is that getting students to their final destination of proficiency is not important, as long as they are growing. However, while it is encouraging to see high growth numbers for a group that is on average very behind in terms of proficiency, it is important to know if that level of growth is adequate for getting them to the target level of performance. The Education Accountability Act of 2009 (SB 09-163) *requires* educators to ask the question, "Was that level of growth enough to get them to proficiency and, ultimately, college and career readiness?" (It also requires that the first two questions also be considered.)

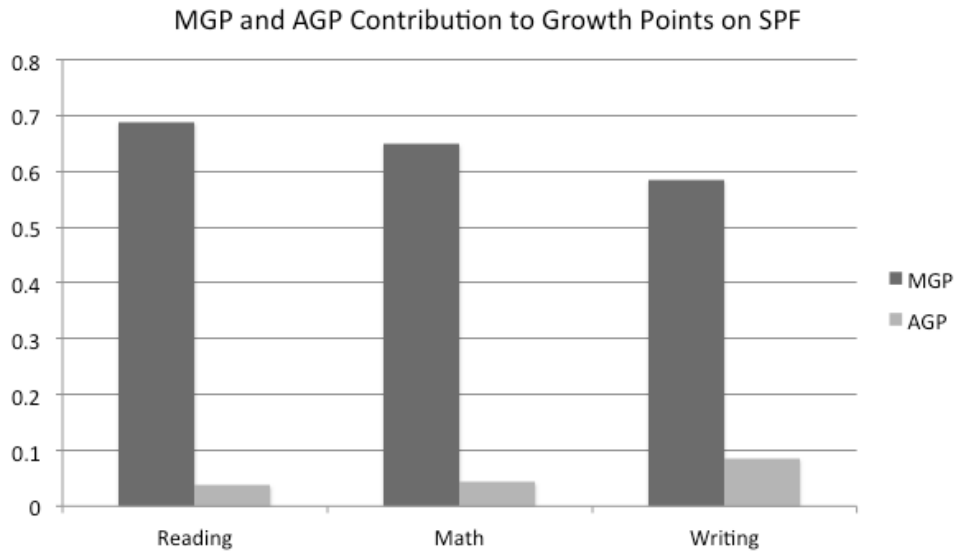
It is true that AGPs are highly correlated with status measures. By definition, the further a group is from the targeted level of achievement, the more growth will be needed to get them there. But does using the AGPs in the Performance Frameworks essentially convert the growth and growth gaps indicators to status measures? What is the end result of this use of both the MGPs and AGPs in the distribution of points on the Performance Frameworks? Has growth been corrupted and overwhelmed by the status information built into the AGPs?

To answer these questions, CDE performed a multinomial logistic regression on the academic growth points earned by schools on the 2010 School Performance Framework[1]. Essentially CDE used a statistical model to predict whether a school level would earn 1, 2, 3 or 4 points on the growth indicator for a given content area based only on that school's MGP and AGP[2] in that content area. Of all the state's schools, 1,895 elementary, middle and high schools had enough data for a point total on this SPF indicator. If status had indeed "taken over" the indicator, the result of the analysis would show that AGP is a better predictor of growth points earned than MGP.

However, the results demonstrated that the opposite is true; the MGP variable is between seven and 18 times better at predicting growth points than AGP, as shown in Figure 2. Status has not therefore taken over the awarding of growth points and accounts for only a very small percentage of variance. **Growth points on the Performance Frameworks are almost completely determined by MGP, not AGP.**

[1] The complete dataset and output of this analysis is available upon request. The data are already published on SchoolVIEW at http://schoolview.org/documents/CDESchoolPerformanceFramework_flatfile_121710.xlsx

[2] Using AGP is similar using the percent proficient/advanced "status" measure, because they are typically correlated at about .90.



Note: Contributions estimated using McFadden's pseudo R^2 statistic.

Figure 2. Multinomial logistic regression results showing MGP and AGP contributions to growth indicator points in the School Performance Framework.

AGP provides a useful perspective on the level of growth needed – it is an unvarnished description of what the situation is. So it is not surprising that AGP and status are correlated with greater academic risk factors at the district and school level (as shown in Figure 9 in Poudre School District's position paper showing the positive relationship between free and reduced lunch percentage and AGP). If students have lower achievement and therefore more ground to make up, they will need to make higher growth to get on track. The recognition of this reality is critical to fulfilling the mission of Colorado's K-12 school system: to get all students to postsecondary and workforce readiness. The state's developing accountability system allows for the identification of where this mission has not been achieved. A year's growth in a year's time (50th percentile growth) is not adequate for students who need to catch up. Even 80th percentile growth may not be enough – but let us look at the situation with open eyes, rather than simply reward high growth but ignore the question of where those students will likely end up in terms of proficiency and postsecondary readiness. Ultimately, the goal is not just growth, not growth to partial proficiency, but growth to proficiency.

The different cut scores on the growth rubric for schools making or not making AGP represent an attempt to create tension in the system where it is needed to push performance, while still striking a fair balance related to normative growth. A key design feature for a quality accountability system is that it promote the desired educational outcomes, as opposed to simply reinforcing the results that the system achieved. The fact that a year's growth in a year's time (50th percentile growth) is simply not enough for students needing to catch up is a critical consideration to reflect in the Performance Frameworks. This mentality stems from the clear policy aims of state law, particularly SB 08-212 and SB 09-163.

Schools of any size only need to have MGPs of 55 in order to merit a "Meets Expectations" on the framework if they are not making their AGP. Schools already making AGP under the current Performance Frameworks get a bit of a break (i.e., MGPs of only 45) because on average their students are already catching up and keeping up; thus the urgency is not quite as great as in schools with a majority of students not on track. But it needs be made clear that schools with very low MGPs will get the lowest number of points for academic growth, and schools with high MGPs will get the highest number of points, regardless of their AGPs.

Regardless of their students' risk factors, districts are tasked with getting *their* students up to Colorado's proficiency standards, not the students 100 miles down the road or those from the higher-income university town in the next district over. Higher-risk groups of students will tend to have higher AGPs – such is

the reality of the situation, because these students have more ground to make up, and consequently need higher growth. The state accountability system does not require that districts or schools meet adequate growth among their students - this would be a difficult target in many cases because of the extent to which many of their students have fallen behind. AGPs describe whether students on average are on track or falling behind – they are not targets, and they are not used as such on the Performance Frameworks.

The argument that Adequate Growth penalizes large schools and districts is based the idea that the deck is stacked against them in two ways: that they are unlikely to be able to demonstrate median growth that differs much from 50, and that adequate growth calculations exacerbate the problem by punishing schools for their demographics (such as when they serve large populations of students with risk factors). As we have shown in this analysis, the first claim is erroneous, and the second is inaccurate in that it attributes great influence on Performance Framework ratings to AGPs, because the empirical evidence for this is nonexistent when the state's schools as a whole are considered.

Claim: The accuracy of Adequate Growth Percentiles (AGPs) has not been investigated and published, so there is no evidence that they tell us anything we don't already know about students and the schools they attend.

In this case, the position paper's claim is true – until now, CDE has not published any of its evidence around the utility of using Adequate Growth. CDE made AGPs a part of the two Academic Growth indicators on the Performance Frameworks; they are a first-pass filter for understanding if the observed growth was good enough. **Is there evidence that these AGPs reveal anything useful beyond what the MGPs and status measures already show?** We regret the delay, but would like to take this opportunity to produce some compelling evidence that adequate growth is worth calculating and including in our accountability system.

CDE has calculated AGPs for districts, schools and disaggregated subgroups back to the 2003-2004 school year, and in so doing has learned a great deal by looking at the data retrospectively. (These data are all currently available on SchoolVIEW and in particular, in the SchoolVIEW Data Lab.) The analyses performed indicated that AGPs are indeed telling a useful story.

AGPs for schools, districts and disaggregated subgroups are calculated by using the Colorado Growth Model to output the minimum level of growth that would bring each student up to or maintain proficiency in three years or by tenth grade. (To understand the details of the calculations, see the video tutorial here http://www.cde.state.co.us/media/training/SPF_Online_Tutorial/player.html). Every student is assigned an AGP – his or her "catch up" or "keep up" number. Schools that have students with large "catch up" numbers will have high median AGPs, and vice versa. Such is the reality of the situation that we face in our state.

A simple test of the validity of AGPs is to determine whether calculating them offers any advantage over not doing so. CDE conducted this test using historical data for each content area between 2007 and 2010. Using 2007 as the base year, only students in grades where they had four complete years ahead of them in which to catch up or keep up were included. The simple model predicted that students already scoring at the proficient level in a given content area in 2007 would continue to do so through 2010, and those scoring below proficient would not attain proficiency within the timeframe. CDE then checked those predictions against what actually happened with those students to get a sense of the accuracy of the base rate prediction—the percentage of the predicted outcomes that actually came true several years later. The AGP-based prediction, on the other hand, uses the statistical power of the Colorado Growth Model to look at score history and growth for each student in order to estimate whether or not a student is on track to catch up (starting out below proficient) or keep up (staying proficient). The AGP-based predictions are also compared against actual data (what really happened to those students) to arrive at a percentage of correct predictions. A summary of results is included in Table 1 below[1].

Table 1. Validating the predictions of catching up and keeping up using historical data – adequate growth versus prior achievement level

		Percentage of correct predictions (using only prior proficiency level)	Percentage of correct predictions (using AGPs)	Improvement in percentage of correct predictions
Math	Below proficient	77.7	88.6	10.9
	Proficient	58.2	75.5	17.2
Reading	Below proficient	55.8	76.2	20.5
	Proficient	78.1	82.6	4.4
Writing	Below proficient	56.4	78.8	22.4
	Proficient	68.7	78.7	9.9

[1] The complete dataset and output from this analysis are available upon request.

As depicted above, merely using the simple prior proficiency model gives pretty good predictions in several cases – predicting that a below-proficient student will remain below proficient in math is accurate 77.7 percent of the time. However, AGP-based predictions incorporating the most recently observed level of growth for each student are better in all cases. The improvement in the percentage of correct predictions (the final column in Table 1) is quite impressive, providing evidence of the validity and usefulness of the AGPs. Most importantly, the results suggest that the AGPs are most useful at discerning which students are beating the odds and catching up, because the improvements in correct predictions are highest for the Below Proficient rows; this is directly attributable to the power of the Colorado Growth Model and its extension to AGPs.

The percentages of correct predictions are unlikely to approach 100 even under the best of circumstances, because of the great number of situations that can affect what happens in a student's life and schooling in the years subsequent to the growth calculation made by the state. Indeed, these levels of prediction are quite remarkable by themselves, showing how useful the Colorado Growth Model can be.

Claim: Large districts and schools do not have equal access to the Exceeds Expectations points for growth and growth gaps on the Performance Frameworks.

A central claim made in Poudre's position paper is that large, diverse districts have much more difficulty in receiving the "Accredited with distinction" label than others. According to their argument, the growth and growth gaps indicators are the drivers of this supposed bias. The claim is that MGPs and AGPs inadvertently provide a steeper hill to climb for large districts than for smaller ones, and consequently it is nearly impossible for them to get a high enough point tally on these two indicators to get into the Distinction category. The evidence provided in favor of this claim is the state's recent performance on the new District Performance Frameworks. As we have demonstrated in other sections of this response to the position paper:

- There is no evidence, statistical or otherwise, that large districts or schools are penalized by the use of the MGP metric. Smaller districts and schools may have more volatility in their MGPs, but there is no straightforward approach to a "correction." Sampling error cannot exist where there is no sampling.
- The calculation of the adequacy of student growth at the district and school level is not only required by the Education Accountability Act of 2009, but it has been shown in our analysis to be a valuable metric at the individual student level.
- There is solid evidence that it is the MGPs that overwhelmingly drive points on the growth indicator of the Performance Frameworks, rather than the AGPs. There is therefore no validity to Poudre School District's claim that "the growth metric, as implemented in the Performance Frameworks, suffers from the same inequity as the status measures it was meant to supersede" (page 14).

The fact that the state's largest districts did not earn as many points on the growth indicators and consequently in total framework points for the first year of implementation the new accountability system does not constitute a proof that they are unable to do so. We would like to take issue with the implied claim that the only districts that can be called "large" are the state's absolutely largest ones. In statistics, fairly smooth distributions tend to form with 200-300 observations. When examining datasets larger than this, changes in distribution shape diminish dramatically. For all practical purposes, districts with n-counts greater than 1000 are of a size that makes their datasets able to "cancel out" the effects of the unmeasured variables mentioned earlier in this paper. There is no need to appeal to very, very large numbers to see the effect of this smoothing; the differences between an n-count of 1,000 and of 10,000 are minimal compared with those between n-counts of 100 and 1,000.

The position paper implies that, in order for a district to be accredited with distinction, the results of its school levels on the Performance Frameworks must approach the maximum number of points available in the growth and growth gaps indicators. The evidence is abundant that this is not the case. In 2010, of the 14 districts accredited with distinction, only three had an Exceeds rating on growth, and not one earned Exceeds on their growth gaps indicator. In fact, for the 512 school levels with public data across all districts in 2010, only in 15 instances did a school level (E, M or H) on the district Performance Framework earn Exceeds in the growth gaps indicator. So it should not be assumed that exceptional performance on growth or growth gaps is a prerequisite for accreditation with distinction. Meeting expectations on growth is nearly always good enough, as long as high numbers of points are also scored on other framework indicators.

Large districts in Colorado tend to have diverse populations that are historically lower-performing. It is the reality that to serve all students, these systems will likely need to make greater improvements than systems with less diverse, higher-performing populations. In 2010 only seven districts total (3.8 percent) were accredited with turnaround plans, among them several large districts. Similarly, a small number of districts were accredited with distinction, the top 7.7 percent. Academy 20 School District was accredited on the basis of its three-year performance results; three years of data effectively triple its size as a district. The 31,274 student growth percentiles that went into its math median make it a large district by anyone's reckoning; it was accredited with distinction. We are confident that all fourteen of these districts amply demonstrated their

qualifications to be accredited with distinction on the basis of strong results among their students in academic performance and postsecondary and workforce readiness.

The Colorado Department of Education continues to believe that the Performance Frameworks currently in use form an equitable base that promotes the desirable conversations around performance management and improvement needed in an accountability system. The department also recognizes that there may be better approaches to this work, and openly welcomes constructive ideas for improving the current frameworks.

Colorado Department of Education
Unit of Research and Evaluation
201 E. Colfax Ave.
Room 502
Denver, CO 80203
303-866-6600
http://www.cde.state.co.us/research/home.htm