



Outcome Data: How is the Catalina Foothills School District succeeding as a 21st century learning environment?

DESIGNING LOCAL PERFORMANCE TASKS FOR DEEP LEARNING

The discussion surrounding the measurement of 21st century skills is all the rage, but conversations abound on how to operationalize the concept in schools. At Catalina Foothills, we've established that these skills matter, are relevant to the educational process, and will be measured as part of a broadened definition of accountability.

Faced with the challenge of trying to foresee students' future needs and how these new demands fit within the learning process, it became clear that we had to accomplish some critical tasks and processes:

- Identify a concrete list of **21st century skills** and define them,
- Establish a **curriculum design process** to better understand the relationship between the skills and academic outcomes,
- Create **rubrics** for each skill to define a continuum of performance,
- Develop a **unit design model** to integrate the skills and teach them,
- Provide **professional learning opportunities** and resources to build capacity and support improvement and implementation of the skills, and
- Create an assessment system to measure them.

Since 2010, professional staff has been involved in the design of high quality curriculum-embedded performance tasks. Their use is critical to measuring the district's academic standards and 21st century skills in cross-disciplinary contexts. The tasks require students to analyze and solve problems, make claims and support them with source-based evidence, and to reason effectively. These important college- and career-readiness skills are generally not addressed by the multiple-choice tests that have dominated the testing landscape for the past two decades. And, if we insist on judging classroom-based and/or district developed performance tasks/assessments with the "validity" and "reliability" criteria traditionally used by statisticians and psychometricians, it will be more difficult to move much beyond factual and procedural recall to achieve the kinds of higher-level student work that we need more of. We believe our teachers are thoughtful assessors of these tasks. Their involvement in developing, scoring, and analyzing the results of them is a powerful form of professional learning and provides a source of outcome data on our students' development and transfer of these skills.

SCALING UP AUTHENTIC PERFORMANCE TASKS

In order to scale-up the use of performance tasks, the district created an assessment team of teacher leaders to learn about the design and use of curriculum-embedded performance tasks. We decided to use the College and Work Readiness Assessment (CWRA) performance task as a model. Marc Chun, a leading authority on performance-based assessment, trained the team. This team provides

professional learning on the development of performance tasks using specific design parameters for use at the elementary, middle, and high school levels. New tools and templates aligned to CFSD language and practice were created, and classroom teachers participate in processes to construct and field-test classroom-based and common performance assessments. They learn how to design robust tasks and scoring rubrics, and then participate in a common scoring process to select student work that will anchor each of the performance levels. The goal has been to design a system of recurring tasks that provide students with ongoing opportunities to apply and transfer higher-order learning.

The table below displays the performance task/assessment design considerations currently being used by CFSD professional staff to create performance tasks.

Performance Task Design Considerations
• Identify an issue or problem
• Create a scenario or context that places the student in an authentic/real-world situation
• Give the student a specific role
• Include a decision to be made or a problem to solve
• Require understanding and transfer of learning
• Identify an authentic product that someone assuming the role would produce (the student will produce this product)
• Include appropriate stakes to add urgency
• Include appropriate opposition to make it compelling
• Create a task library (set of documents) that includes a range of information sources (students will use these documents to respond to the problem or issue)
• Require students to engage in critical thinking, problem solving, and communication (or similar 21 st century skills) in order to arrive at a solution or make a decision
• Create transparent evaluation criteria and rubrics that are made clear to the students so they can evaluate their work and receive feedback on strengths and weaknesses

CWRA+ AS A DRIVER FOR PROGRESS

Locally designed performance tasks provide us with outcome data on skills such as critical thinking, problem solving, and communication within the context of a specific content area. The ultimate test of how much students are learning is their ability to transfer that knowledge to a new context. With that in mind, we decided to search for an external measure of 21st century skills to further evaluate the district’s work on building these skills. We quickly found that standardized commercial assessments that assess 21st century skills were few, but learned about an innovative and highly touted performance task, the College and Work Readiness Assessment (CWRA). After an extensive review and discussions with staff at the Council for Aid to Education (CAE), where the CWRA is created, we selected it as a longitudinal measure. The CWRA, now called the CWRA+ provides the district with a value-added approach to the measurement of students’ higher-order thinking and written communication skills regardless of their academic focus.

The CWRA+ contains two major components: a performance task (PT) and a series of selected response questions (SRQ). The performance task challenges students to make decisions and synthesize information from a variety of sources to compose a response to a complex problem. Students must comprehend, evaluate, and integrate a variety of these sources through a digital interface, and prepare a detailed written response to an intended audience.

The second part of the assessment includes 25 selected response questions that measure scientific and quantitative reasoning, critical reading and evaluation, and critiquing an argument. Like the performance task, these questions are document-based and require students to draw information from provided materials.

Since the 2011-2012 school year, we have been assessing all of our freshmen and seniors and analyzing their performance. In Spring 2016, we began testing juniors instead of seniors. We determined that it would be more useful to individual students and the school if the scores were received while the students were still in high school. Students may choose to share their results with potential colleges as further evidence of the skills they have acquired during high school.

The CWRA+ measures skills that we know our students will need for postsecondary success and lifelong learning. It functions as a cross-curricular assessment to inform instructional practices, including the effects of changes to curriculum and assessment design. Our annual use of the CWRA+ makes clear that higher-order thinking and learning is valued in CFSD because that is what the assessment requires. Currently, CFSD is the only public school district in Arizona that uses the CWRA+ as a tool for assessing students' critical thinking and written communication skills, and measuring growth on these skills to determine student preparedness for college and the workplace.

FRESHMAN AND SENIOR/JUNIOR PERFORMANCE ON THE CWRA+

Assessing freshmen informs the district about students' specific strengths and weaknesses after their K-8 experience, making formative use of the results. The senior (now junior) results provide the district with information on college readiness (national comparison of seniors (now juniors) in CFSD to college freshmen), national comparison of our high school students with students at participating CWRA+ high schools, and internal growth at the high school – actual freshmen to senior (now junior) gains by cohort.

Although the district has been using the CWRA/CWRA+ data since 2011, the outcome data highlighted below displays the CWRA+ results for freshmen and seniors/juniors for the past three years. The 2013-2014 school year was the year the CAE expanded the scope of the assessment and renamed it the CWRA+. Though our freshmen do well, compared to other freshmen that take the test nationally, our seniors' average scores have been far above the average of all other CWRA/CWRA+ seniors taking the test.

Table 1 and Table 2 display the distribution of CFSD freshman and senior/junior performance in the subscore categories of Analysis & Problem Solving, Writing Effectiveness, and Writing Mechanics on the Performance Task (PT) component of the CWRA+. Also displayed are the overall PT mean scores for three consecutive school years beginning with school year 2013-2014. The subscores range from 1 (low) to 6 (high). The rubric continuum is best described as improvement over a Grades 9-16 spectrum, so the goal is for high schools to bring their students a certain distance (peaking at the 3s and 4s) with colleges doing the rest. The results show that CFSD freshmen at Catalina Foothills High School are performing, on average, above freshmen at all participating CWRA+ high schools. CFSD senior cohorts, on average, improved or maintained performance in the subscore categories on the performance task component of the CWRA+.

Table 1. Comparison of CWRA+ Performance Task Subscores and Mean Scores: CFSD Freshmen and All CWRA+ Schools

Summary of Subscores and Mean Scores for CFSD Freshmen with All Freshmen At CWRA+ Schools								
Analysis & Problem Solving			Writing Effectiveness		Writing Mechanics		Overall PT Mean Score	
Year	CFSD	All Schools	CFSD	All Schools	CFSD	All Schools	CFSD	All Schools
2013-2014	3.4	3.1	3.4	3.2	3.7	3.4	1017	976
2014-2015	3.5	3.1	3.5	3.2	3.7	3.4	1032	976
2015-2016	3.3	3.1	3.3	3.2	3.7	3.4	1015	976

Source: CWRA+ Institutional Reports for Catalina Foothills School District, Fall 2013 – Spring 2016

Table 2. Comparison of CWRA+ Performance Task Subscores and Mean Scores: CFSD Seniors and All CWRA+ Schools

Summary of Subscores and Mean Scores for CFSD Seniors/Juniors with All Seniors At CWRA+ Schools								
Analysis & Problem Solving			Writing Effectiveness		Writing Mechanics		Overall PT Mean Score	
Year	CFSD	All Schools	CFSD	All Schools	CFSD	All Schools	CFSD	All Schools
2013-2014	3.7	3.5	3.9	3.4	4.0	3.9	1126	1058
2014-2015	3.9	3.5	3.9	3.5	4.0	3.8	1127	1058
2015-2016*	4.0	3.5	4.2	3.5	4.2	3.8	1148	1058

*2016 scores reflect first year test administration for CFSD juniors.

Source: CWRA+ Institutional Reports for Catalina Foothills School District, Fall 2013 – Spring 2016

Table 3 and Table 4 display the freshman and senior/junior results for the selected response questions (SRQ) subscore categories and the overall SRQ mean score. The data in Table 3 show that freshmen results, overall, are improving in each subscore category. The overall SRQ mean score increased 26 points over three school years. Scientific & Quantitative Reasoning increased 20 points. Critical Reading & Evaluation increased 16 points. Critique an Argument increased by 5 points with a slight dip in the 2014-2015 school year. Senior scores in each of the SRQ categories increased 12-17 points and the overall SRQ mean score increased by 35 points between 2013 and 2015. The data also show that CFSD seniors are outperforming, on average, all CWRA+ schools that test seniors by 65 points.

Table 3. Comparison of CWRA+ Selected Response Questions: Mean Subscores and Overall Mean Scores for Freshmen

Selected Response Questions: Mean Freshmen Subscores and Overall SRQ Score								
Scientific & Quantitative Reasoning			Critical Reading & Evaluation		Critique an Argument		Overall SRQ Mean Score	
Year	CFSD	All Schools	CFSD	All Schools	CFSD	All Schools	CFSD	All Schools
2013-2014	468	475	474	479	475	476	973	985
2014-2015	481	475	484	479	468	476	992	985
2015-2016	488	475	490	479	480	476	1011	985

Note: The selected-response subscores are reported on a scale ranging approximately from 200 to 800.

Source: CWRA+ Institutional Reports for Catalina Foothills School District, Fall 2013 – Spring 2016

Table 4. Comparison of CWRA+ Selected Response Questions: Mean Subscores and Overall Mean Scores for Seniors/Juniors

Selected Response Questions: Mean Senior/Junior Subscores and Overall SRQ Score								
Scientific & Quantitative Reasoning			Critical Reading & Evaluation		Critique an Argument		Overall SRQ Mean Score	
Year	CFSD	All Schools	CFSD	All Schools	CFSD	All Schools	CFSD	All Schools
2013-2014	519	507	523	508	516	508	1088	1058
2014-2015	533	507	540	508	528	508	1123	1058
2015-2016*	524	507	538	508	531	508	1113	1058

*2016 scores reflect first year test administration for CFSD juniors.

Note: The selected-response subscores are reported on a scale ranging approximately from 200 to 800.

Source: *CWRA+ Institutional Reports for Catalina Foothills School District, Fall 2013 – Spring 2016*

Mastery Levels are criterion-referenced indicators of performance on the CWRA+. They provide distinctions in student performance relative to students’ proficiency in critical thinking and written communication. The Mastery Levels are determined by students’ Total CWRA+ scores. As a district, they are determined by each class level’s mean Total CWRA+ score. Each mastery level – Below Basic, Basic, Proficient, Accomplished, and Advanced – corresponds to specific evidence of critical thinking and written communication skills.

Table 5 and Table 6 display the Total CWRA+ mean scores and Mastery Levels for freshmen and seniors/juniors. Both freshmen and seniors scored within the proficient mean mastery level. Fifty-six to sixty-one percent of freshmen scored at the proficient, accomplished, and advanced mastery levels. The Total CWRA+ mean score for freshmen increased by 18 points over a three-year period.

The seniors, on average, scored within the proficient mean mastery level with 82 percent of seniors scoring at proficient and advanced mastery levels compared to 57 percent of freshmen. The Total CWRA+ mean score for seniors increased by 18 points over a two-year period of time.

Table 5. Comparison of Total CWRA+ Mean Score and Mean Mastery Levels of Freshmen

Summary of Total CWRA+ Mean Scores and Mastery Levels: Freshmen							
Year	Total CWRA+ Mean Score	Mean Mastery Level	Percent Below Basic	Percent Basic	Percent Proficient	Percent Accomplished	Percent Advanced
2013-2014	995	Proficient	2	42	55	N/A	1
2014-2015	1012	Proficient	1	38	58	N/A	2
2015-2016	1013	Proficient	2	37	31	28	2

Source: *CWRA+ Institutional Reports for Catalina Foothills School District, Fall 2013 – Spring 2016*

Table 6. Comparison of Total CWRA+ Mean Score and Mean Mastery Levels of Seniors/Juniors

Summary of Total CWRA+ Mean Scores and Mastery Levels: Seniors/Juniors							
Year	Total CWRA+ Mean Score	Mean Mastery Level	Percent Below Basic	Percent Basic	Percent Proficient	Percent Accomplished	Percent Advanced
2013-2014	1107	Proficient	1	20	72	N/A	7
2014-2015	1125	Proficient	1	14	75	N/A	9
2015-2016*	1131	Accomplished	0	16	21	51	11

*2016 scores reflect first year test administration for CFSD juniors.

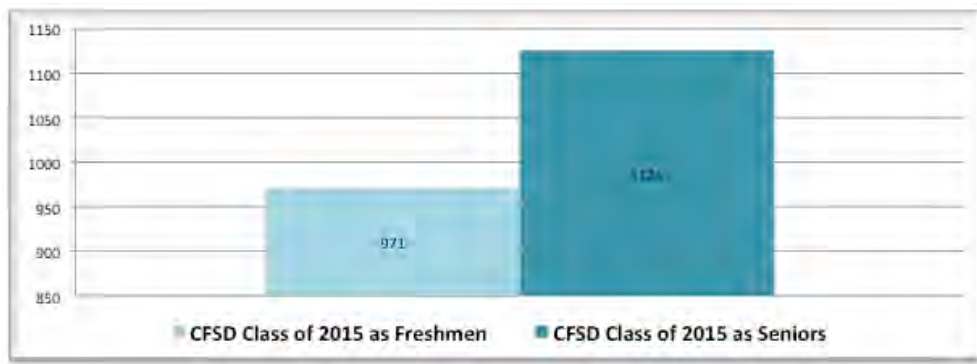
Source: *CWRA+ Institutional Reports for Catalina Foothills School District, Fall 2013 – Spring 2016*

Noteworthy are the first year scores for juniors also displayed in Tables 5 and 6. One can surmise that focused efforts on 21st century skill building and the use of curriculum-embedded performance tasks at the elementary, middle, and high school levels are paying off. Eighty-three percent (83%) of CFSD juniors scored at the proficient, accomplished, and advanced mastery levels. The 2015-2016 junior cohort is the same group shown as the 2013-2014 freshman cohort in Tables 5 and 6. The Total CWRA+ mean score increased 136 points between freshman and junior year and only 16 percent of students scored at the Below Basic and Basic mastery levels at junior year compared to 44 percent during freshman year.

Further comparison of the results of the 2015-2016 junior cohort and their results as freshmen during the 2013-2014 school year reveals more gains! The overall mean SRQ score displayed in Table 3 for freshmen increased 140 points by junior year. Mean freshman results for the SRQ subscore categories also increased. Scientific & Quantitative Reasoning increased 56 points. Critical Reading & Evaluation increased 64 points, and Critique and Argument increased 56 points. The overall performance task mean score increased by 131 points!

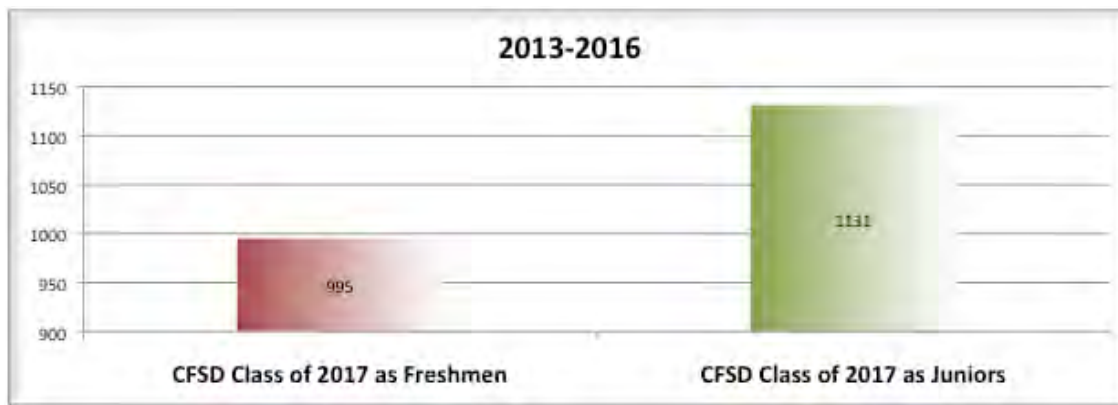
The following graph compares the total CWRA/CWRA+ mean score of the first freshman cohort to their scores as seniors. The 2011-2012 freshman class was the first cohort to take the CWRA. As seniors during the 2014-2015 school year, their performance increased by 155 points.

Growth of CFSD Students from Grade 9 to Grade 12: Total CWRA/CWRA+ Mean Scores



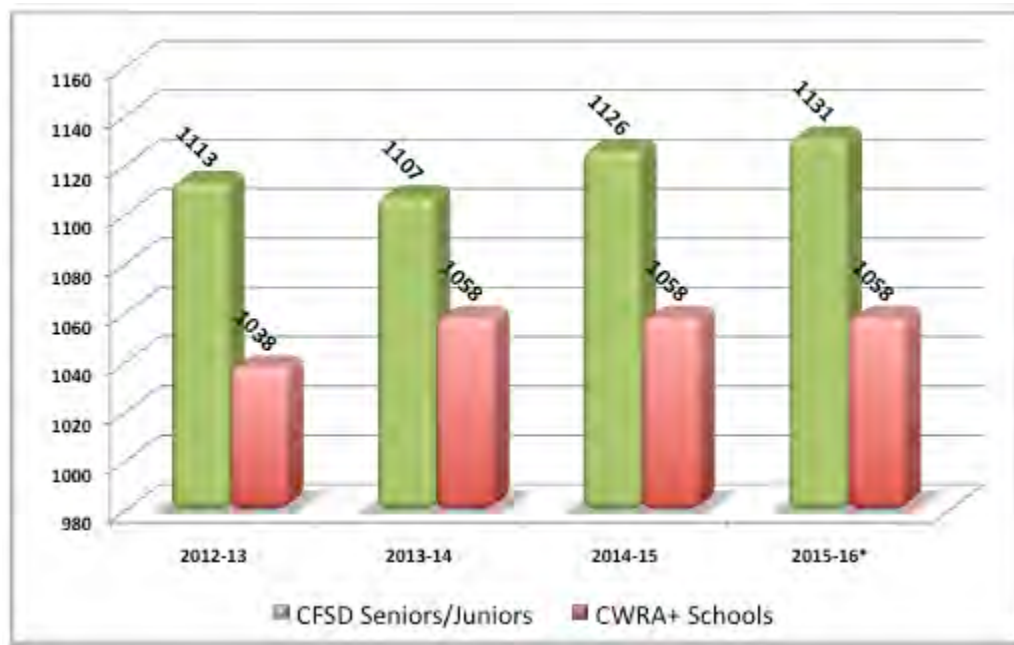
Note: CWRA Scores Converted to CWRA+ Scores for 2012-2013

The graph below compares the total CWRA+ mean score of the class of 2017 as freshmen and juniors from 2013-2016. Spring 2016 was the first test administration of the CWRA+ to juniors. As juniors during the 2015-2016 school year, the total mean CWRA+ mean score increased by 136 points since those same students were freshmen.



An analysis of the total CWRA/CWRA+ scores from school year 2012-2013 to 2015-2016 shows that CFSD seniors and juniors at Catalina Foothills High School are outperforming seniors at participating CWRA/CWRA+ schools. The table below displays the scores for CFSD and all CWRA+ schools for this four-year period.

**Comparison of CFSD Seniors/Juniors to Seniors at CWRA/CWRA+ High Schools:
Total CWRA+ Mean Scores**

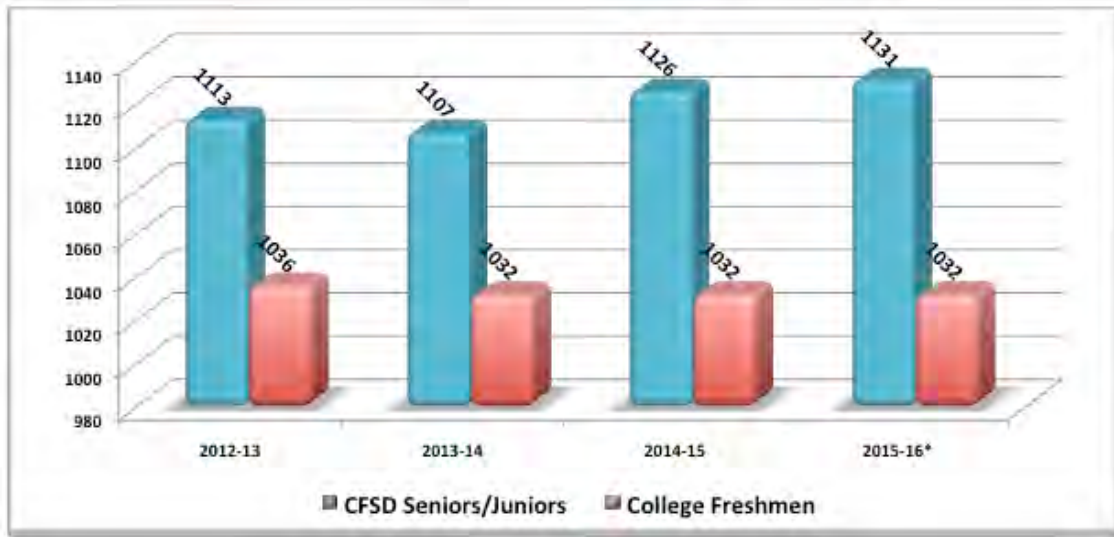


* CFSD Juniors Tested

Note: CWRA Scores Converted to CWRA+ Scores for 2012-2013

An analysis of the Total CWRA and CWRA+ mean scores from 2013-2016, as displayed below, shows that CFSD seniors/juniors overall, are college ready! That is, they are outperforming college freshmen at participating colleges/universities who are administering the Collegiate Learning Assessment (CLA and CLA+), which parallels the design of the CWRA and CWRA+. Below are the mean scores for CFSD seniors/juniors and college/university freshmen that are taking the CLA and CLA+ at participating schools.

College Readiness: CWRA/CWRA+ Score Comparison of CFSD Seniors/Juniors to Freshmen at CLA/CLA+ Colleges/Universities



* 2015-2016 Juniors Tested

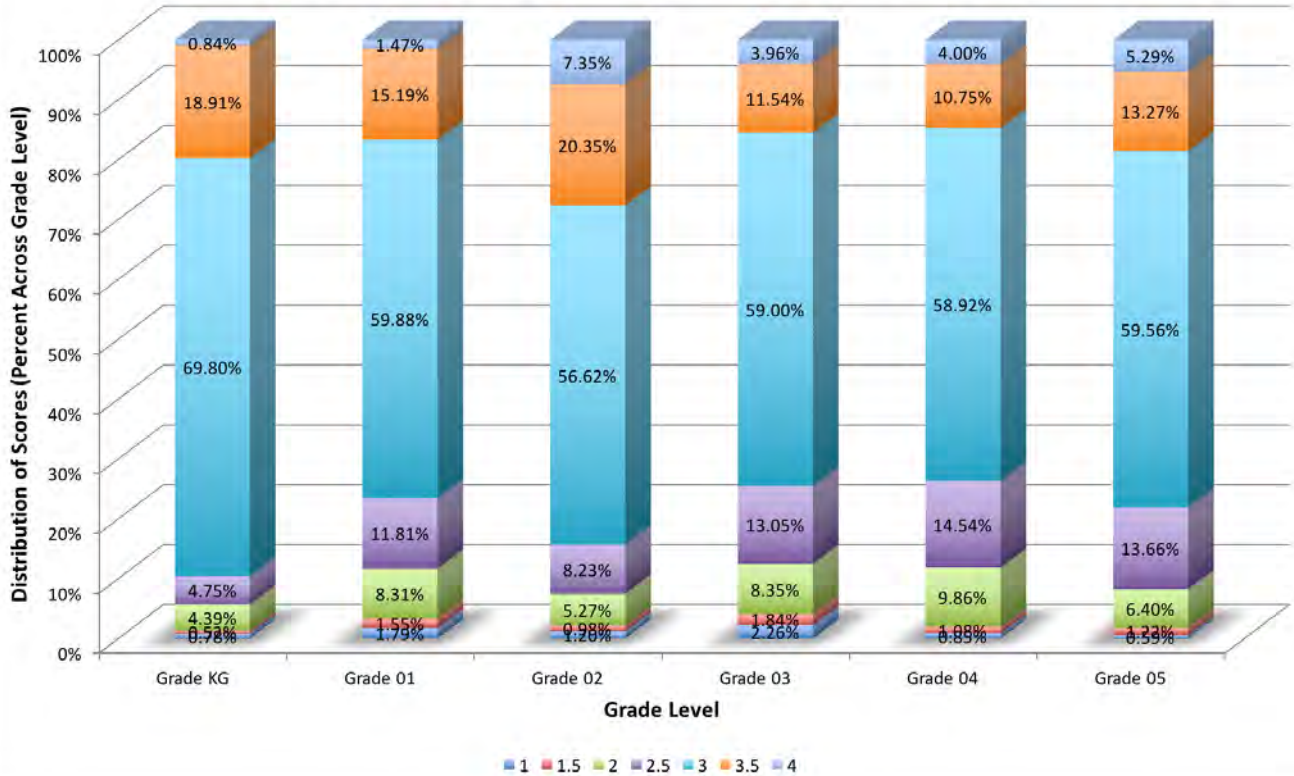
Note: CWRA and CLA Scores Converted to CWRA+ Scores for 2012-2013

REPORTING OUT ON 21ST CENTURY SKILLS

The district creates a variety of reports to show cumulative progress in 21st century skill building and to communicate student performance on common performance tasks. The scores from teachers’ electronic grade books are archived each grading period. Common assessment scores are entered using specific naming conventions and can be extracted in a variety of ways to analyze progress across programs of study. For example, we can create reports that show aggregate scores of common assessment tasks or for a specific skill such as critical thinking and problem solving across multiple grade levels and/or subject areas.

The graph below, *Score Distribution for DLP: Critical Thinking & Problem Solving*, displays the percentage of rubric scores earned by elementary students, grades kindergarten through fifth grade, on summative assessments for critical thinking and problem solving during the 2015-2016 school year. Critical thinking and problem solving includes performance indicators such analysis and interpretation, reasoning, and problem solving/solution finding. A score of “3.0” (blue) represents proficient performance in the context of a particular assignment or assessment. At third grade, 59% of earned scores are at “3.0” – Proficient, 11.54% are at “3.5” and 3.96% are at “4.0” – Advanced, as reported by teachers. Overall, 74.5% of third grade students are scoring proficient and above in tasks that require students to demonstrate critical thinking and problem solving. 25.5% of students are not yet proficient. Since students take multiple classes, it can be assumed that some students received a summative score for critical thinking in more than one class.

Score Distribution for DLP: Critical Thinking & Problem Solving

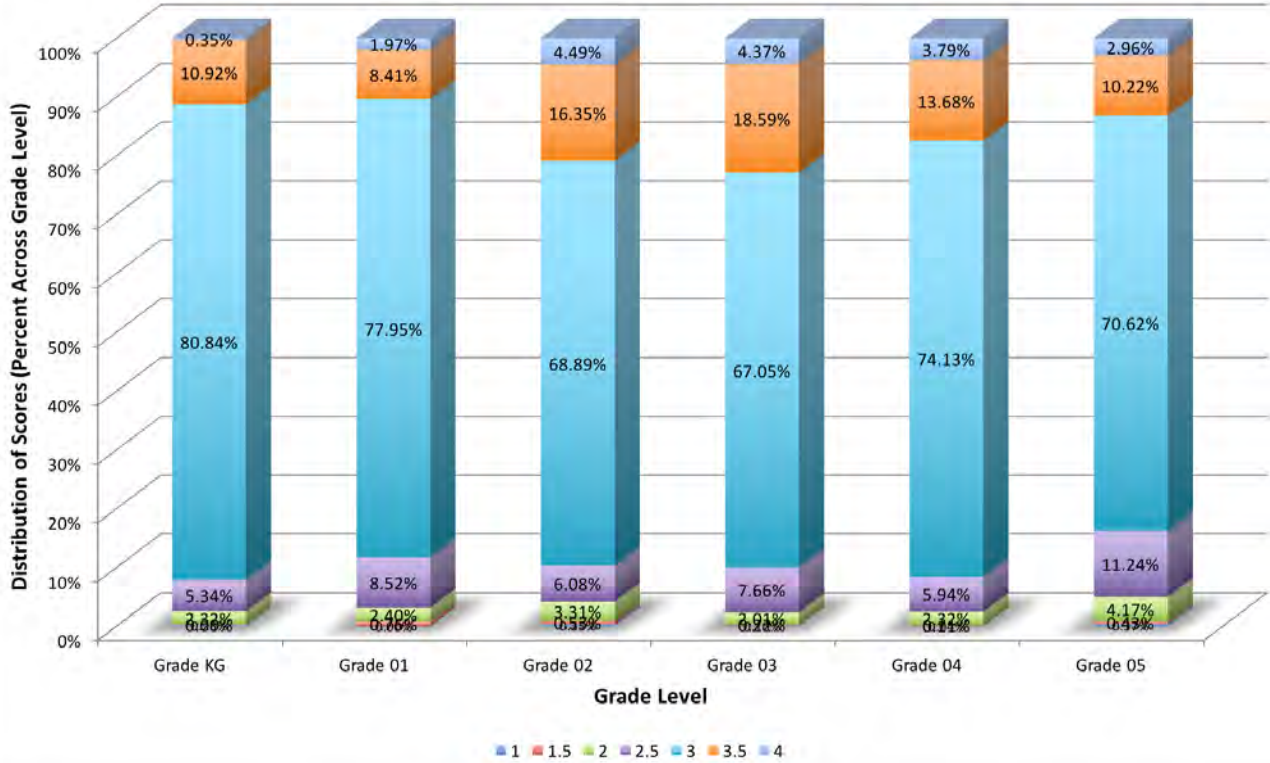


The performance areas and indicators for Critical Thinking and Problem Solving are defined in the district’s rubric for Critical Thinking and Problem Solving. Rubrics that show a continuum of performance have been created for grades K-2, 3-5, 6-8, and 9-12. Rubrics are available for each of CFSD’s 21st century skill (now titled deep learning proficiencies or DLPs).

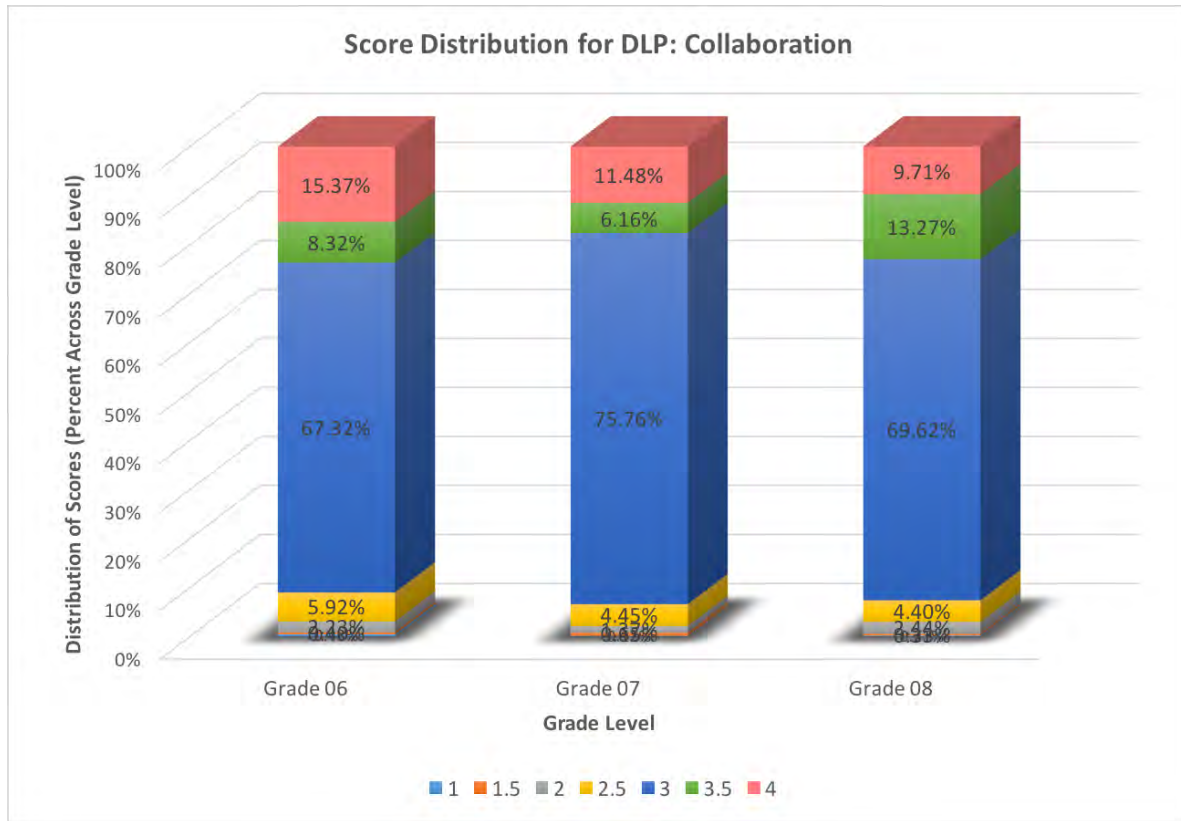
The graphs that follow show the distribution of rubric scores in percentages for Collaboration at the elementary and middle school levels. Collaboration refers to skills such as cooperating as a member of a highly successful group, working together with one or more people to achieve a goal (e.g., solve problems, create novel products, learn and master content) and revising work in response to feedback. The demands of 21st century learning and working increasingly call for cooperative and collaborative efforts.

On average, the data at the elementary level show that 84% to 92% of students are demonstrating skills of effective collaboration based on summative scores for the 2015-2016 school year.

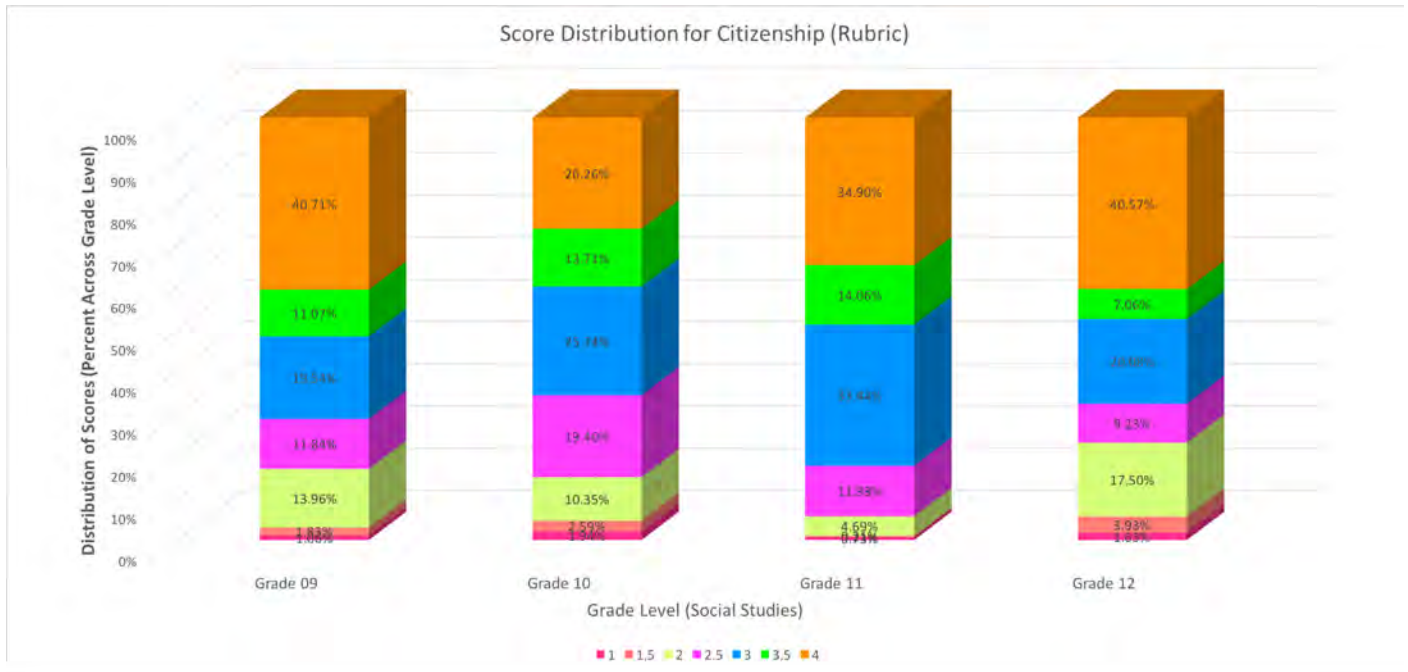
Score Distribution for DLP: Collaboration



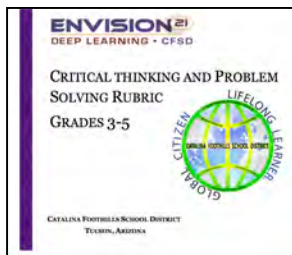
At the middle level, 91%-93% of students are demonstrating effective collaboration skills, as reported by teachers in their electronic grade books.



The graph below shows the percentage of cumulative scores on summative assessments at each rubric level for Citizenship across all social studies courses at the high school. The data show that students had higher scores at the 4.0 – Advanced level than at the 3.0 Proficient level. For example, 40.7% of students enrolled in Global Issues at grade 9 scored 4.0 in measures of Citizenship compared to 11% at 3.0. Skills associated with this proficiency include cultural literacy and civic literacy and engagement.



CFSD’s focus on teaching and measuring 21st century skills will help our students navigate their environment, work with others, perform well, and achieve their goals.



Learn more about the performance areas and indicators for Critical Thinking and Problem Solving and all of the CFSD rubrics on CFSD’s website.

WHAT’S IN A GRADE?

Since the 2008-2009 school year, CFSD has been using a standards-referenced report card at the elementary and middle school levels. In our standards-referenced grading and reporting system, grades/scores reflect what students know and can do relative to the Catalina Foothills School District standards for learning. This means that students and parents receive explicit feedback about their progress on content area knowledge *and* 21st century skills (DLPs). At the high school, students receive grades in the targeted 21st century skills that have been designated as the focus for each subject area/department.

The report card lists the essential categories or measurement topics for each subject area. Students receive indicators of progress based on a four-point rubric to show how well they are performing in those areas. A score of 3.0 means that a student is consistently demonstrating proficiency in the learning expectations associated with a specific category. For example, students will receive a score for the category of Ratio and Proportions in mathematics. At the middle level, students also receive an overall score for each subject area. Standards-referenced grade/scores help teachers plan their instruction so they can challenge and support all students. They help parents know the academic

areas in which a student meets or exceeds the standard, needs challenge, or needs support. Below is a section of the report card that shows the essential categories of learning for middle school mathematics. Each subject area is similarly represented on the report card. A student may earn an overall score of 3.5 for mathematics, but scores in the individual categories will show which categories are strengths and which may be in need of improvement.

	TERM	Q1	Q2	Q3	Q4
MATHEMATICS	OVERALL GRADE				
Ratio and Proportional Relationships					
The Number System					
Expressions and Equations					
Functions					
Geometry					
Statistics and Probability					
Mathematical Practices (Problem Solving, Reasoning, etc.)					
DLP / TECHNOLOGY / PERSONAL RESPONSIBILITY					
DLP: Critical Thinking and Problem Solving					
PR: Class Participation					
PR: Self-regulation					

Students also receive separate grades/scores on their development and growth in Personal Responsibility (e.g., self-regulation, class participation, work completion/effort), and the Educational Technology standards. The goal is to identify, for every student, the next appropriate “stretch” to move students toward higher levels of learning.

At Catalina Foothills, the measurement of student success is driven by the outcomes we value and the type of evidence that best demonstrates success. We are using locally-designed performance tasks, an externally validated performance assessment (CWRA+), and cumulative classroom grades/scores on summative tasks/assessments to gauge progress and report out on 21st century skills. We won’t know if we are succeeding or need to improve in our efforts unless we can document and analyze the work in a variety of contexts.

The advice of a broad range of business and professionally engaged citizens helped us agree on the knowledge and skills that matter in our 21st century world. We are committed to fully engaging our students in their learning and evaluating program results to ensure our students move on from their PreK-12 education confident that they are well prepared to prosper in post-secondary education, on the job, and in life.