

Section 3: Consistency

The district outlines the process and procedures used to ensure **consistency**. As a result of choosing a course-based/common assessment approach, the district must provide a process and evidence for **all three criteria**.

Overview of Consistency

Fremont County School District #1 ensures scoring consistency by adopting common assessments and common scoring protocol. This includes common rubrics, answer keys and exemplars. Teachers are trained in the use of the common assessment techniques on a yearly basis. This training is conducted at the common core department level by department heads, mentors, or instructional facilitators.

All faculty record the results of all assessments into the district wide student information system. Regular monitoring of each teacher's "grade book" will allow administration to keep track that all assessments are being given and scores recorded in a timely manner. If scores on assessments appear unreasonable or other irregularities are seen discussion with that teacher occurs. A copy of one teacher's "grade book" follows this narrative. This sample is a screen shot of Mr. Patton's grade book. As you can see, the assessments are divided into common and "other". The transparent nature of this system allows the district to monitor implementation and consistency.

Infinite Campus - Windows Internet Explorer

https://icampus.fre1.k12.wy.us/campus/main.xsl

File Edit View Favorites Tools Help

Google Search

msn Powered by Live Search Search

News Entertainment Video Sports

Infinite Campus Infinite Campus



Year 08-09 School Lander Valley High School Section 1) LA1041-110 Lang Arts 12 Sem 1

Index Search Help

Charles Patton

- Student Information
 - General
 - Counseling
 - Grad Planner
- Instruction
 - Attendance
 - Gradebook
 - Grading By Task
 - Grading By Student
 - Roster
 - Daily Planner
 - Lesson Planner**
- Admin
 - Reports
- Behavior
 - Behavior Referral
- Ad Hoc Reporting
 - Filter Designer
 - Report Designer
 - Report Builder

LA1041-110 Lang Arts 12 Sem 1
Teacher: Patton, Charles Edward

Lesson Planner

New Task Group New Assignment New Activity Lesson Plan Copier

Lesson Plan

- Term S1 Progress
- Term S1 Quarter
- Term S1 Semester
 - Common Assessments (70%)
 - Common Speech [100pts]
 - Common Paper [100pts]
 - Common Final [100pts]
 - Common Lit Analysis Paper [50p
 - Common Exam [100pts]
 - Other (30%)

Infinite Campus - Windows Internet Explorer

https://icampus.fre1.k12.wy.us/campus/main.xsl

File Edit View Favorites Tools Help

Google Search

msn Powered by Live Search Search

News Entertainment Video Sports

Infinite Campus Infinite Campus

Infinite Campus

Year **08-09** School **Lander Valley High School** Section **1) LA1041-110 Lang Arts 12 Sem 1**

LA1041-110 Lang Arts 12 Sem 1
Teacher: Patton, Charles Edward

Lesson Planner

New Task Group New Assignment New Activity Lesson Plan Copier

Lesson Plan

- Term S1 Semester
 - Common Assessments (70%)
 - Essay/Syllabus [20pts]
 - Unit 1.1 - Study Guide [20pts]
 - Unit #1.1 - Exam [100pts]
 - Journals [30pts]
 - Unit #1.2 - Study Guide [20pt]
 - Unit #1.2 - Exam [100pts]
 - Journals [30pts]
 - Reading [100pts]
 - Unit #2.1 - Study Guide [20pt]
 - Unit #2.1 - Exam [100pts]
 - Unit #2.2 Study Guide [20pts]
 - Unit #2.2 Exam [100pts]
 - Word Web/Outline [25pts]
 - Rough Draft - BOE [25pts]
 - Other (30%)

Index Search Help

- Charles Patton
 - Student Information
 - General
 - Counseling
 - Grad Planner
 - Instruction
 - Attendance
 - Gradebook
 - Grading By Task
 - Grading By Student
 - Roster
 - Daily Planner
 - Lesson Planner**
 - Admin
 - Reports
 - Behavior
 - Behavior Referral
 - Ad Hoc Reporting
 - Filter Designer
 - Report Designer
 - Report Builder

3a Open-Ended Assessments

- Procedures to be used to ensure inter-rater reliability
- Defines a desired acceptable rate
- Present data from implementation of stated procedures

On-going processes and procedures to ensure consistency on **open-ended assessments** are clearly described.

When entire assessments or items are open-ended in design, we address inter-rater reliability through well established methods. These include training teachers in the use of the rubrics and exemplars and assigning teachers new to the district a mentor to assist with initial scoring responsibilities. We identify 3 of the nine content areas to be reviewed each school year. This approach guarantees that all areas are reviewed in a 3 year period. We randomly select 1 indicator course from each of the 3 identified areas. Then we randomly select 10% of student work from the open-ended common assessments. We double score each of those assessments for a check on reliability.

We've considered the issue of acceptable rates for inter-rater reliability. After deliberation among the BOE Leadership Team we came to the decision that our desired rate for inter-rater reliability is 80% having an exact agreement and/or 95% having an exact and adjacent agreement. We define exact match as when two scorers independently score a piece of student work and arrive at the same exact score. Adjacent agreement is when the two independent scorers are within 1 point of each other. We feel that 80% exact or 95% adjacent is a challenging, yet obtainable goal.

Desired, **acceptable rate** is defined

Data relating to inter-rater reliability has been collected and analyzed. Results of the inter-rater reliability work for Language Arts and Health conducted during the 2008-2009 school year are displayed in the tables below.

As you will see in those tables some of our inter-rater percentages are below our targets. The process we use to address those low reliability scores is to repeat the double score process later in the school year. If those teachers continue to score the same papers very differently they will be asked to participate in additional training that focuses on performance standards and examination of common student work. That training is conducted by our instructional facilitator and/or department chair.

Data are presented that support implementation of the stated procedures. Also, the district has looked at the results and has found low reliability scores and outlines a plan to address this issue.

Content Area	Indicator Class	Common Assessment	Number Papers Double Scored	#/% Exact Match	#/% Exact and Adjacent Match
Language Arts	LA 11	Literary Analysis	8	7 88%	8 100%
		Consortium	9	8 89%	9 100%
		Culminating Writing Project	11	5 46%	11 100%
	LA 12	Consortium	8	3 38%	8 100%
		Final Essay	8	5 63%	8 100%

Content Area	Indicator Class	Common Assessment	Number Papers Double Scored	#/% Exact Match	#/% Exact and Adjacent Match
Health	Health	49	10	8 80%	10 100%
		50	10	1 10%	9 90%
		51	10	5 50%	8 80%
		52	10	8 80%	10 100%
		53	10	3 30%	9 90%
		54	10	3 30%	10 100%
		20	11	4 36%	11 100%
		21	11	7 64%	11 100%
		22	11	2 18%	11 100%
		23	11	3 27%	11 100%
		24	11	2 18%	11 100%
		25	11	6 55%	11 100%

3b. Closed-Ended Assessments

A clear procedure to ensure reliability and **desired acceptable rates** is presented.

- Clear procedures to ensure reliability
- Defines a desired acceptable rate
- Present data from implementation of stated procedures

Closed ended assessments are reviewed by the district curriculum office using Chronbach’s Alpha procedures. We attended the BOE training sponsored by the Wyoming Assessment Activities Consortium and were provided a Reliability Calculator tool developed by Jason Nicholas, Ph.D. We have used that tool to conduct reliability studies on several Language Arts and Health assessments. The results of those studies are in the table below. Multiple choice items or assessments that are determined to be less reliable than we desire will be replaced.

FCSD1 chooses .7 reliability coefficient as the minimum acceptable rate of reliability. We made this decision after researching the options, reading professional publications, consulting with recognized experts in the field and considering the impact of any one assessment in our BOE system. In a publication from the National Council on Measurement in Education David Frisbie reported that teacher made tests have reliability coefficients around .5 while commercially prepared ones are around .9. Using the .7 is a reasonable level between teacher made and commercially prepared. Conversations with Jason Nicholas, Ph.D. confirm that .7 is a reasonable level given the nature of our BOE system (many short assessments over multiple years) and the length of most of our closed-ended assessments.

The approach, **acceptable rates** and rationale are described.

The data in the table below shows that our assessments in Language Arts are not at the .7 level of reliability. We are currently investigating each of those assessments and will replace the items and/or entire assessment prior to next administration. The two lowest assessments, those administered to our honors and AP classes, may not be as unreliable as the Cronbach’s Alpha score indicates. Those classes contain students that are very homogenous in ability with many high scoring students. Conversations with Dr. Nicholas indicate that those scores may be lower due to statistical effects.

Data is provided to support implementation

Content Area	Indicator Class	Common Assessment	Cronbach's Alpha Score
Language Arts	AP English 12	First Semester Final	.3322
	LA 11	First Semester Final	.6390
	Honors English 11	First Semester Final	.1842
	LA 12	First Semester Final	.6424

Content Area	Indicator Class	Common Assessment	Cronbach's Alpha Score
Health	Health	Assessment # 9	.7129
		Final Assessment	.8726

3c. Teacher Judgment

- Procedures to ensure reliability of judgment
- Clear documentation that judgment is anchored to performance standards
- Present data from implementation of stated procedures

A clear procedure to ensure reliability of **teacher judgment** across assessments within a course and across teachers is outlined.

The 30% assessments are allowed to vary from teacher to teacher, year to year. These assessments may be seen as “teacher judgment”. They are designed to allow teachers to modify assessment practices in response to student needs and circumstances. The 30% assessments are required to be aligned with state and district standards and benchmarks within the content area. The building principals monitor those assessments for consistency by having teachers maintain documentation that addresses consistency issues. When the assessments are open-ended a scoring rubric is required.

Each semester, to confirm consistency between the 30% teacher specific assessments and the 70% common assessments, the district will select 25% of the teachers of indicator classes and calculate a correlation coefficient between the overall percentage and the percentage determined from the 30% teacher specific assessments. The district will randomly select an indicator class(es) from those taught by that teacher. This will allow us to determine the impact of the 30% in relation to the student’s overall grade. Our goal is to see a high correlation (.80) between the 30% assessments and the overall percentage. If we find a low correlation, the building principals will work with the teacher to determine the cause and develop a plan to close the gap. The Language Arts and Health data from the fall semester of 2008 is displayed in the chart that follows. As the data shows, the reliability of teacher judgment is very high and we see solid evidence that allowing teacher’s to have 30% at their own discretion is not changing the outcome for students.

Content Area	Indicator Class	Teacher	Correlation Coefficient
Language Arts	LA 11	Hall	.833
	Honors English 11	Primrose	.875
	LA 12	Primrose	.719
		Patton	.855
		Patton	.909
Health	Health	Comings	.859
		Brasel	.789

Procedures to ensure reliability of **teacher judgment** across assessments and across teachers are described. Data to show implementation is included.

