

Name: John Doe
 Project 2 – Analysis of a Test
 EDF 3430 Critical Assignment

Subject Matter Area: American History—The Cold War
 Grade: 11th

Distribution of Scores

Student Scores	
100	
100	
100	
90	
90	
80	
80	
80	
70	
70	
60	
60	
60	
60	
60	
50	
50	
40	
30	
30	
68	MEAN
60	mode
65	median
21.9089023	Standard deviation

Frequency Distribution

Score	Frequency
100	3
90	2
80	3
70	2
60	5
50	2
40	1
30	2

Item Distribution of Student Responses

	Item Number	1	2	3	4	5	6	7	8	9	10
U P P E R G R O U P	Student 1	A	C	B	D	D	T	T	F	T	F
	Student 2	A	C	B	D	D	T	T	F	T	F
	Student 3	A	C	B	D	D	T	T	F	T	F
	Student 4	A	C	B	A	D	T	T	F	T	F
	Student 5	A	D	B	D	D	T	T	F	T	F
	Student 6	A	C	B	D	D	T	F	F	T	T
	Student 7	A	C	B	A	D	T	T	F	T	T
	Student 8	A	C	A	D	D	F	T	F	T	F
	Student 9	A	D	B	D	D	T	T	F	F	T
	Student 10	A	D	A	A	D	T	T	F	T	F
	Key	A	C	B	D	D	T	T	F	T	F
	Number correct	10	7	8	7	10	9	9	10	9	7
	Number incorrect	0	3	2	3	0	1	1	0	1	3
L O W E R G R O U P	Student 11	A	C	C	A	D	F	T	F	T	T
	Student 12	A	C	C	A	D	F	T	F	T	T
	Student 13	B	C	A	D	D	F	T	F	F	F
	Student 14	A	C	B	D	D	F	F	F	F	T
	Student 15	A	C	B	C	C	T	T	T	F	F
	Student 16	A	C	D	D	B	F	T	T	F	F
	Student 17	C	B	A	D	D	T	T	F	F	T
	Student 18	A	B	B	C	D	F	T	T	F	T
	Student 19	A	D	C	C	D	F	F	T	F	F
	Student 20	A	C	B	A	B	F	F	T	F	T
	Key	A	C	B	D	D	T	T	F	T	F
	Number correct	8	7	4	5	7	2	7	5	2	4
	Number incorrect	2	3	6	5	3	8	3	5	8	6

Distribution of Student Responses & Item Difficulty

Item #	A (T)	B (F)	C	D	Item Difficulty	
1	18*	1	1	0	18/20=0.9	Easy
2	0	2	14*	4	14/20=0.7	Optimal
3	4	12*	3	1	12/20=0.6	Optimal
4	6	0	3	11*	11/20=0.55	Difficult
5	0	2	1	17*	17/20=0.85	Easy
6	11*	9	0	0	11/20=0.55	Difficult
7	16*	4	0	0	16/20=0.8	Optimal
8	5	15*	0	0	15/20=0.75	Optimal
9	11*	9	0	0	11/20=0.55	Difficult
10	9	11*	0	0	11/20=0.55	Difficult

Item Discrimination

Item 1

	A*	B	C	D	D Index	Item evaluation
Top Group	10	0	0	0	$10/10 - 8/10 = 0.2$	Fair item.
Bottom Group	8	1	1	0		

Item 2

	A	B	C*	D	D Index	Item evaluation
Top Group	0	0	7	3	$7/10 - 7/10 = 0$	No discriminating item. Poor item.
Bottom Group	0	2	7	1		

Item 3

	A	B*	C	D	D Index	Item evaluation
Top Group	2	8	0	0	$8/10 - 4/10 = 0.4$	Very good item.
Bottom Group	2	4	3	1		

Item 4

	A	B	C	D*	D Index	Item evaluation
Top Group	3	0	0	7	$7/10 - 4/10 = 0.3$	Good item.
Bottom Group	3	0	3	4		

Item 5

	A	B	C	D*	D Index	Item evaluation
Top Group	0	0	0	10	$10/10 - 7/10 = 0.3$	Good item.
Bottom Group	0	2	1	7		

Item 6

	T*	F	D Index	Item evaluation
Top Group	9	1	$9/10 - 2/10 = 0.7$	Very good item.
Bottom Group	2	8		

Item 7

	T*	F	D Index	Item evaluation
Top Group	9	1	$9/10 - 7/10 = 0.2$	Fair item.
Bottom Group	7	3		

Item 8

	T	F*	D Index	Item evaluation
Top Group	0	10	$10/10 - 5/10 = 0.5$	Very good item.
Bottom Group	5	5		

Item 9

	T*	F	D Index	Item evaluation
Top Group	9	1	$9/10 - 2/10 = 0.7$	Very good item.
Bottom Group	2	8		

Item 10

	T	F*	D Index	Item evaluation
Top Group	3	7	$7/10 - 4/10 = 0.3$	Good item.
Bottom Group	6	4		

This test covered the Cold War unit. The unit included the entire Cold War from start to finish. The test assesses on the Bay of Pigs Invasion, the Cuban Missile Crisis, and also the Korean and Vietnam Wars. The students of this test are all in eleventh grade. The students are a diverse group. The class has twelve females compared to ten males, making the class a female majority. Of those students, majority are white. Six of them are Hispanic, two of them are African-American, and one is Asian. Every student is proficient in English. Scores for this test were very spread out, ranging from scores of 30-100. Looking at the standard deviation of test scores, the test does not seem to be very reliable. Since the standard deviation is higher, it means that the SEM would be higher as well. Based on that indication, a higher SEM would mean that the reliability is lower. However, I do believe there is a strong case for validity for this test. The test measures the students' sole understanding of Cold War topics. The test is not culturally biased as well. Perhaps the vocabulary may have been advanced for the class. I do not see that as an issue though.

Overall as a class, the scores of the students did not meet expectations. Throughout the unit, students were extremely engaged and interested in the topics. Through group activities, each student showed a level of interest and participated in the activities. The mean of the scores was 68, so just two points away from a C average in the class. The expectations of the class mean was to be at about the 75 mark. As mentioned before, the standard deviation of scores was high too, which can be explained by the wide range of scores in the classroom. The boys performed better than the girls did on the test. The boys scored a mean of 64, whereas the girls scored a mean of 60. Based on the mean score of the classroom, I will have to further investigate the test questions.

Although the mean score was below a C average, I still believe the test was effective. I believe that it challenged the class. There were no trick questions on the assessment. The questions were straightforward content-related questions that were made to prove an understanding. It was either you knew the answer or you did not. Obviously, there were students who seemed to not understand the unit as a whole. The test had its share of very good items. Only one item was poor. Item two was the poor item. Both the bottom and top group found success with seven answering correctly in each group. Items six, eight, and nine were all very good items. These items particularly saw the top group flourish while the bottom group struggled. Those items had the largest separation between the top and bottom group. The best way to close such separation is to alter future lesson plans to make them more engaging and fun. I firmly do not believe that the diversity in the classroom played a role in the mean of the scores. The questions were not culturally biased. A good assessment strategy for the next test will offer a type of review. I think a "Jeopardy" style assessment review would have each student engaged with the test material. The implications from this assessment is that students need to know the material. Students need to take their homework seriously and parents need to ensure that their children are absorbing the information. As a teacher, the implications I have taken are to get the class more engaged in the material.

Rubric

Criteria	Exceeds Expectations 10-8 Points	Meets Expectations 7-5 Points	Does Not Meet Expectations 4-0 Points	Not Attempted	Your score
Statistics	No errors in statistical analysis	Fewer than 5 errors in statistical analysis.	No frequency distribution for total scores of individual items. No mean, median, mode or standard deviation	Failed to submit or submitted late.	
Item difficulty & Item Discrimination	Both indexes were computed with no errors and interpretations were provided	Both indexes were computed with some errors.	Only one or no indexes were computed.	Failed to submit or submitted late.	
Paragraph: Written Description Assessment: FL-FAU-FAEP 1.2.	A thorough, accurate and precise description of the assessment process was provided including types of students that might have taken (i.e. grade, gender, SES, ESOL or ESE.	An accurate description which requires more clarification was provided including types of students that might have taken (i.e. grade, gender, SES, ESOL or ESE.	No conclusion or written description of the purpose for the assessment provided.	Failed to submit or submitted late.	
Paragraph: Overall Description	An accurate description and conclusion provided that reflects the entire assessment process and what strategies were used to ensure validity and reliability	The description provided covers the assessment process and statistical analysis.	No comment on the assessment process or overall statistical analysis is provided.	Failed to submit or submitted late.	
Paragraph: Discussion of the Assessment Process Continuous improvement:	Discussion includes (1) weaknesses and strengths of the assessment process (2) item analysis (item difficulty and item discrimination	A general conclusion for the assessment process using at least three items is	No discussion of the implications or how to improve the quality of the assessment process is	Failed to submit or submitted late.	

FL-FAU- FEAP.3.1	for at least 4 items) and if the variance and outliers found may be indicative of having a diverse population in the classroom. (3) implications for students, parents (4) strategies to be used to meet all students needs.	provided together with general implications and reflection	provided.		
-----------------------------	--	--	-----------	--	--