

3710 Mid-Term, Fall 2009

Reflection Time

It is now time to determine if I am satisfied with how the test functioned in the context of my class goals.

My goal was to use the test *to distinguish between those who know the material best and those who do not know it as well*. The “material” was the knowledge, understandings and applications related to Chapters 1-8.

Thus, the test needs to be difficult (but not unreasonably difficult).

It also needs to be fair.

The steps I use for test reflection:

- First, I determine internal reliability.
- Second, I reflect on validity (Table of Specifications, face validity)
- Third, I perform item analysis.
- Fourth, I determine final test scores.
- Fifthly, I do a little research a curious question . . .

Item Analysis: Let's Begin -

15 people took the test;
Below are the number right by item

01. ___ 13 ___

03. ___ 12 ___

05. ___ 15 ___

07. ___ 9 ___

09. ___ 8 ___

11. ___ 14 ___

13. ___ 13 ___

15. ___ 8 ___

17. ___ 6 ___

19. ___ 7 ___

21. ___ 12 ___

23. ___ 11 ___

25. ___ 12 ___

27. ___ 7 ___

29. ___ 7 ___

31. ___ 12 ___

33. ___ 3 ___

35. ___ 14 ___

37. ___ 12 ___

39. ___ 10 ___

41. ___ 10 ___

43. ___ 3 ___

45. ___ 12 ___

47. ___ 10 ___

49. ___ 10 ___

02. ___ 13 ___

04. ___ 6 ___

06. ___ 12 ___

08. ___ 14 ___

10. ___ 12 ___

12. ___ 11 ___

14. ___ 4 ___

16. ___ 13 ___

18. ___ 9 ___

20. ___ 9 ___

22. ___ 10 ___

24. ___ 5 ___

26. ___ 11 ___

28. ___ 6 ___

30. ___ 12 ___

32. ___ 4 ___

34. ___ 9 ___

36. ___ 12 ___

38. ___ 3 ___

40. ___ 13 ___

42. ___ 13 ___

44. ___ 5 ___

46. ___ 13 ___

48. ___ 10 ___

50. ___ 7 ___

Item Difficulty Index: I will set acceptable item difficulty (aka “p-value”) at around 75%; that is, if 75 percent of the class gets the answer correct, then the item is not too difficult. This translates into 11 or more people getting the item right. Using this criterion, we are left with 26 items to analyze further.

07. ___ 9 ___

09. ___ 8 ___

15. ___ 8 ___

17. ___ 6 ___

19. ___ 7 ___

27. ___ 7 ___

29. ___ 7 ___

33. ___ 3 ___

39. ___ 10 ___

41. ___ 10 ___

43. ___ 3 ___

47. ___ 10 ___

49. ___ 10 ___

04. ___ 6 ___

14. ___ 4 ___

18. ___ 9 ___

20. ___ 9 ___

22. ___ 10 ___

24. ___ 5 ___

28. ___ 6 ___

32. ___ 4 ___

34. ___ 9 ___

38. ___ 3 ___

44. ___ 5 ___

48. ___ 10 ___

50. ___ 7 ___

NEXT - Item Discrimination Index: A difficult item is good . . . as long as it differentiates between those who know and those who do not. So, let’s examine the 26 items to see if they “discriminate” properly.

We do this by comparing the best tests to the worst tests. I decide to do this by comparing the top quarter to the bottom quarter. I will take the top 20% (3 tests), compare them to the bottom 20% (3 tests), and compute the discrimination index (aka D).

TABLE

Item Number	# Right in High Grp	# Right in Low Grp	% in High	% in Low	“D Value” High%-Low%
07.	3	0	1.00	.0	+1.00
09.	3	1	1.00	.33	+.67
15.	3	1	1.00	.33	+.67
17.	3	1	1.00	.33	+.67
19.	3	0	1.00	.0	+1.00
27.	2	1	.67	.33	+.34
29.	2	0	.67	.0	+.67
33.	0	1	.0	.33	-.33
39.	3	1	1.00	.33	+.67
41.	2	1	.67	.33	+.34
43.	0	0	.0	.0	0
47.	3	2	1.00	.67	+.33
49.	3	0	1.00	.0	+1.00
04.	2	1	.67	.33	+.34
14.	1	1	.33	.33	0
18.	2	1	.67	.33	+.34
20.	2	2	.67	.67	0
22.	2	1	.67	.33	+.34
24.	2	0	.67	.0	+.67
28.	2	1	.67	.33	+.34
32.	1	0	.33	.0	+.33
34.	2	2	.67	.67	0
38.	0	1	.0	.33	-.33
44.	1	2	.33	.67	-.34
48.	2	2	.67	.67	0
50.	1	1	.33	.33	0

My analysis shows that I have nine items that are not functioning as I intended. They are poor test items . . . they are unfair.

Item Number	# Right in High Grp	# Right in Low Grp	% in High	% in Low	“D Value” High%-Low%
33.	0	1	.0	.33	-.33
43.	0	0	.0	.0	0
14.	1	1	.33	.33	0
20.	2	2	.67	.67	0
34.	2	2	.67	.67	0
38.	0	1	.0	.33	-.33
44.	1	2	.33	.67	-.34
48.	2	2	.67	.67	0
50.	1	1	.33	.33	0

What should I do about these findings vis-à-vis the students who took this exam?

I decided to award each student 4.5 points (.5 point added back in for each poor item) to their otherwise final point total since I “own” this amount of overall “error” in students’ scores. This was done for all students, whether they missed these nine items or not.