LOMA LINDA UNIVERSITY
MEDICAL CENTER
Staff Development

Faculty Development Showcase Week:
Using the Item Analysis to Enhance Test Construction

Course Description:
This is a brief audiovisual presentation designed to introduce faculty to ways to enhance and improve multiple choice test construction using statistical data from the item analysis.The multiple choice written examination is a significant formative and summative evaluation tool for many health professions. Great valid examinations are difficulty to construct. Using statistical data from an Item Analysis can provide direction to both novice and experienced faculty members as well as strengthen the reliability and validity of these examinations.

Course Objectives:
The participant will:
1. State how to use an item analysis summary to determine the reliability and validity of the examination.
2. Develop and modify multiple choice examination questions
3. Define whether multiple choice questions are strong indicators of student knowledge retention.

Using Item Analysis to Enhance Test Construction

NS, RN
ssor

Item Analysis

bility Coefficient

e extent that the test is likely
istent results over time

s will have greater
wer items

subject

Reliability Coefficient

der-Richardson-20 (KR-20) is a
of the internal reliability
nation

tests that have
ers

m 0 to

KR-20 Determined By

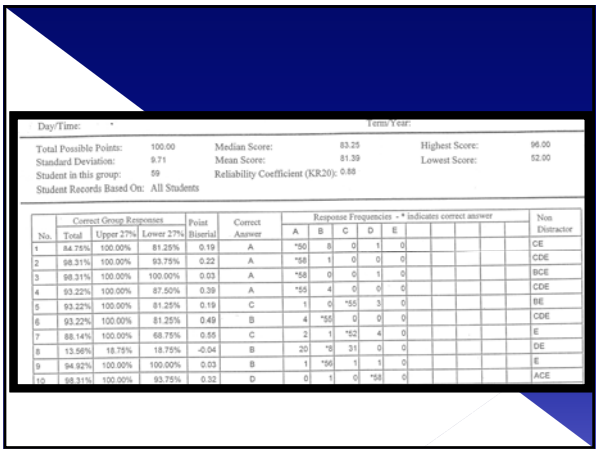
- l number of test items
- ion of the responses
- e correct
- sponses to

KR-20

- eliability
- uestions unrelated to each
- cause is that the
- ber of

KR-20

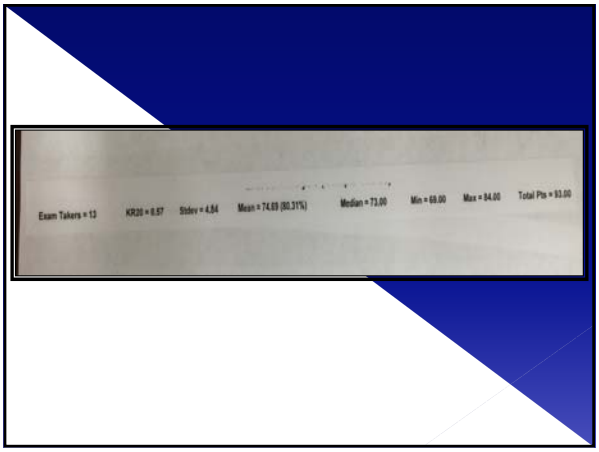
- ility coefficient of greater than
- good for a nursing
- be improved by
- of items
- e items to
- nts



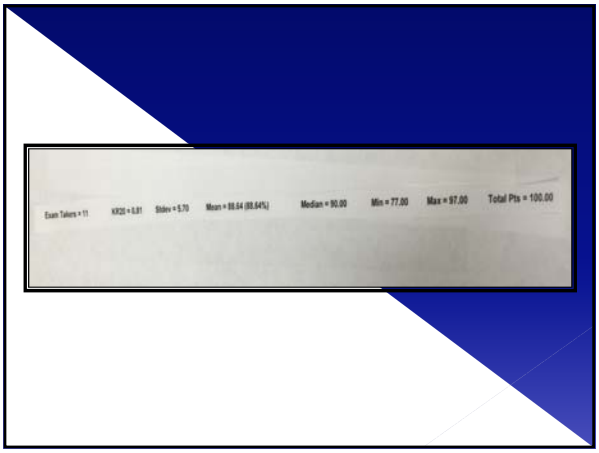
Day/Time: * Term/Year:

Total Possible Points:	100.00	Median Score:	83.25	Highest Score:	96.00
Standard Deviation:	9.71	Mean Score:	81.39	Lowest Score:	52.00
Student in this group:	59	Reliability Coefficient (KR20):	0.88		
Student Records Based On:	All Students				

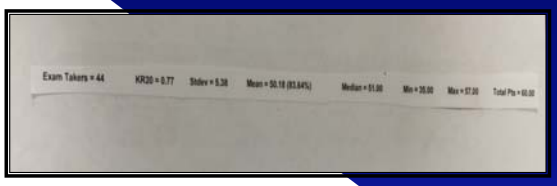
No.	Correct Group Responses			Point Biserial	Correct Answer	Response Frequencies - * indicates correct answer									Non Distractor
	Total	Upper 27%	Lower 27%			A	B	C	D	E					
1	84.75%	100.00%	81.25%	0.19	A	*50	8	0	1	0					CE
2	98.31%	100.00%	93.75%	0.22	A	*58	1	0	0	0					CDE
3	98.31%	100.00%	100.00%	0.03	A	*58	0	0	1	0					BCE
4	93.22%	100.00%	87.50%	0.39	A	*55	4	0	0	0					CDE
5	93.22%	100.00%	81.25%	0.19	C	1	0	*55	3	0					BE
6	93.22%	100.00%	81.25%	0.49	B	4	*55	0	0	0					CDE
7	88.14%	100.00%	68.75%	0.55	C	2	1	*52	4	0					E
8	13.56%	18.75%	18.75%	-0.04	B	20	*8	31	0	0					DE
9	94.92%	100.00%	100.00%	0.03	B	1	*56	1	1	0					E
10	98.31%	100.00%	93.75%	0.32	D	0	1	0	*58	0					ACE



Exam Takers = 13 KR20 = 0.57 Stdev = 4.84 Mean = 74.69 (80.31%) Median = 73.00 Min = 69.00 Max = 84.00 Total Pts = 93.00



Exam Takers = 11 KR20 = 0.81 Stdev = 5.70 Mean = 89.64 (89.64%) Median = 90.00 Min = 77.00 Max = 97.00 Total Pts = 100.00



Item Analysis

k the effectiveness of test items

e exam and sorts the results

mber of students
7% on this
is exam
se

Item Analysis

well-written item

tion of students in the
d have selected the

s in the
d each

Item Difficulty Level: Definition

tage of students who answered the item correctly

High (Difficult)	Medium (Moderate)	Low (Easy)
<= 30%	> 30% AND < 80%	>=80%

Item Difficulty Level: Examples

f students who answered each item = 50

rrect	% Correct	Difficulty Level
	30	High
		Medium
		Medium
		ow

Item Difficulty Level: Discussion

- est question that no one o easy?
- ion on which no one difficult?
- "too easy" out?

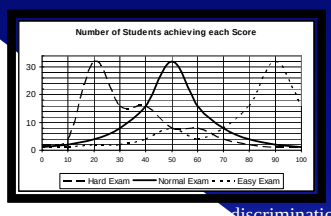
What is Item Discrimination?

rally, students who did well on
should select the correct
given item on the exam.
n Index
item between
ts who did
who

What is Item Discrimination?

Discrimination Index is
decimal format and
en -1 and 1
n Index is
serial

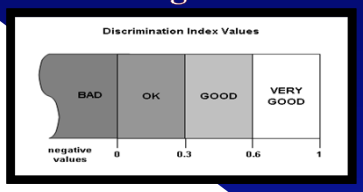
What is a "good" value?



discrimination of

ment

What is a “good” value?

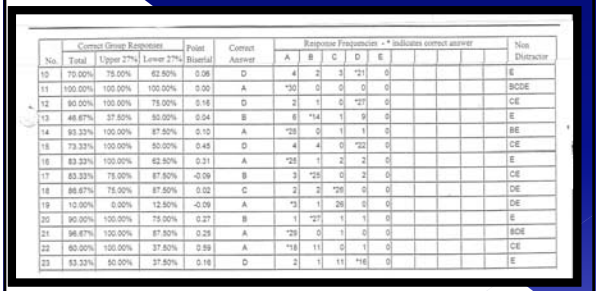


d and rewritten for

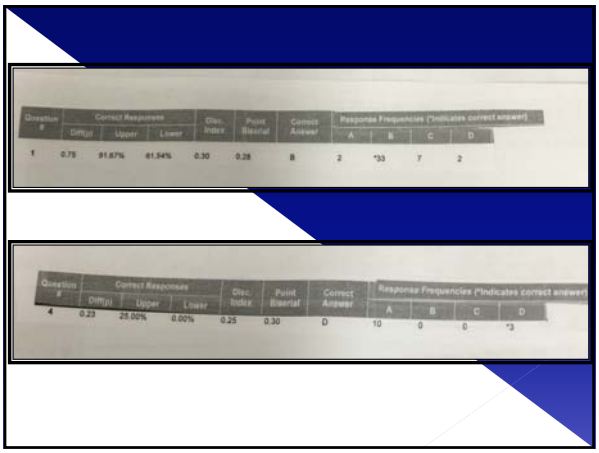
d the same on

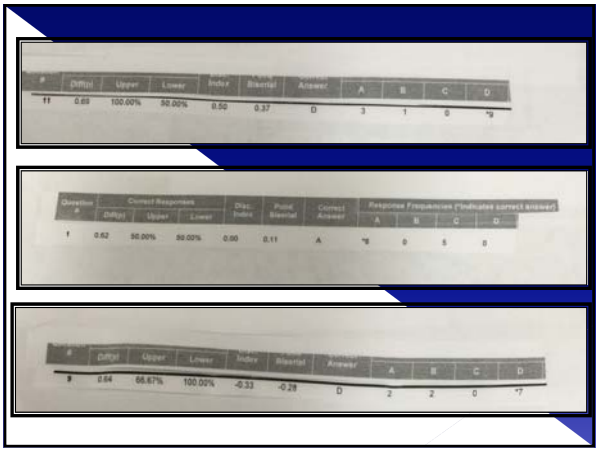
ing

No.	Correct Group Responses			Point Biserial	Correct Answer	Response Frequencies - * indicat					
	Total	Upper 27%	Lower 27%			A	B	C	D	E	
1	84.75%	100.00%	81.25%	0.19	A	*50	8	0	1	0	
2	98.31%	100.00%	93.75%	0.22	A	*58	1	0	0	0	
3	98.31%	100.00%	100.00%	0.03	A	*58	0	0	1	0	
4	93.22%	100.00%	87.50%	0.39	A	*55	4	0	0	0	
5	93.22%	100.00%	81.25%	0.19	C	1	0	*55	3	0	
6	93.22%	100.00%	81.25%	0.49	B	4	*55	0	0	0	
7	88.14%	100.00%	68.75%	0.55	C	2	1	*52	4	0	
8	13.56%	18.75%	18.75%	-0.04	B	20	*8	31	0	0	
9	94.92%	100.00%	100.00%	0.03	B	1	*56	1	1	0	
10	98.31%	100.00%	93.75%	0.32	D	0	1	0	*58	0	
11	59.32%	87.50%	50.00%	0.18	D	2	17	5	*35	0	



No.	Correct Group Responses			Point Biserial	Correct Answer	Response Frequencies - * indicates correct answer								Non Distractor
	Total	Upper 27%	Lower 27%			A	B	C	D	E				
10	70.00%	75.00%	62.50%	0.06	D	4	2	3	*21	0				E
11	100.00%	100.00%	100.00%	0.00	A	*30	0	0	0	0				BCDE
12	90.00%	100.00%	75.00%	0.16	D	2	1	0	*27	0				CE
13	46.67%	37.50%	50.00%	0.04	B	6	*14	1	9	0				E
14	93.33%	100.00%	87.50%	0.10	A	*28	0	1	1	0				BE
15	73.33%	100.00%	50.00%	0.45	D	4	4	0	*22	0				CE
16	83.33%	100.00%	62.50%	0.31	A	*25	1	2	2	0				E
17	83.33%	75.00%	87.50%	-0.09	B	3	*25	0	2	0				CE
18	86.67%	75.00%	87.50%	0.02	C	2	2	*26	0	0				DE
19	10.00%	0.00%	12.50%	-0.09	A	*3	1	26	0	0				DE
20	90.00%	100.00%	75.00%	0.27	B	1	*27	1	1	0				E
21	96.67%	100.00%	87.50%	0.25	A	*29	0	1	0	0				BDE
22	60.00%	100.00%	37.50%	0.59	A	*18	11	0	1	0				CE
23	53.33%	50.00%	37.50%	0.16	D	2	1	11	*16	0				E





Item Discrimination: Examples

Number of Correct Answers in Group		Item Discrimination Index
%	Lower 27%	
	20	0.7
	70	0.1
		1
		0
		0

Item Discrimination: Discussion

ors could contribute to low
nation between the two
s?
for a negative

Distracter Analysis: Definition

he performance of the highest-
ring 27% of the students on
ns (i.e. the incorrect
he exam)
should choose
nswer

Distracter Analysis: Examples

m 1	A*	B	C	D	E
ner	20	5	0	0	0
		10	10	10	5
			5	10	0

Distracter Analysis: Discussion

he purpose of a good

should you
t?

No.	Correct Group Responses			Point Biserial	Correct Answer	Response Frequencies - * Indicates correct answer										Non Distractor
	Total	Upper 27%	Lower 27%			A	B	C	D	E						
10	70.00%	75.00%	62.50%	0.06	D	4	2	3	*21	0						E
11	100.00%	100.00%	100.00%	0.00	A	*30	0	0	0	0						BCDE
12	90.00%	100.00%	75.00%	0.16	D	2	1	0	*27	0						CE
13	46.67%	37.50%	50.00%	0.04	B	6	*14	1	9	0						E
14	93.33%	100.00%	87.50%	0.13	A	*28	0	1	1	0						BE
15	73.33%	100.00%	50.00%	0.45	D	4	4	0	*22	0						CE
16	83.33%	100.00%	62.50%	0.31	A	*25	1	2	2	0						E
17	83.33%	75.00%	87.50%	-0.09	B	3	*25	0	2	0						CE
18	86.67%	75.00%	87.50%	0.02	C	2	2	*26	0	0						DE
19	10.00%	0.00%	12.50%	-0.09	A	*3	1	26	0	0						DE
20	90.00%	100.00%	75.00%	0.27	B	1	*27	1	1	0						E
21	96.67%	100.00%	87.50%	0.25	A	*29	0	1	0	0						BDE
22	60.00%	100.00%	37.50%	0.59	A	*18	11	0	1	0						CE
23	53.33%	50.00%	37.50%	0.16	D	2	1	11	*16	0						E

No.	Correct Group Responses			Point Biserial	Correct Answer	Response Frequencies - * Indicates correct answer										Non Distractor
	Total	Upper 27%	Lower 27%			A	B	C	D	E						
1	84.75%	100.00%	81.25%	0.19	A	*50	0	0	1	0						CE
2	98.31%	100.00%	93.75%	0.22	A	*58	1	0	0	0						CDE
3	98.31%	100.00%	100.00%	0.03	A	*58	0	0	1	0						BCE
4	93.22%	100.00%	87.50%	0.38	A	*55	4	0	0	0						CDE
5	93.22%	100.00%	81.25%	0.19	C	1	0	*55	3	0						BE
6	93.22%	100.00%	81.25%	0.49	B	4	*55	0	0	0						CDE
7	88.14%	100.00%	68.75%	0.55	C	2	1	*52	4	0						E
8	13.56%	18.75%	18.75%	-0.04	B	20	*8	31	0	0						DE
9	94.92%	100.00%	100.00%	0.03	B	1	*56	1	1	0						E
10	98.31%	100.00%	93.75%	0.32	D	0	1	0	*58	0						ACE
11	59.32%	87.50%	50.00%	0.18	D	3	17	5	*35	0						E
12	100.00%	100.00%	100.00%	0.00	C	0	0	*59	0	0						ABDE

What to do with test questions

Upper 27%	Lower 27%	Biserial	Nondistractors Present	Action
Correct	Incorrect	+	-	No action
Correct	Incorrect	+	+	Revise nondistractors
Incorrect	Correct	-	-	Revise test item
Incorrect	Correct	-	+	Revise test item & nondistractors
Incorrect	Incorrect	+ or -	+ or -	Reteach Content

Importance to Education

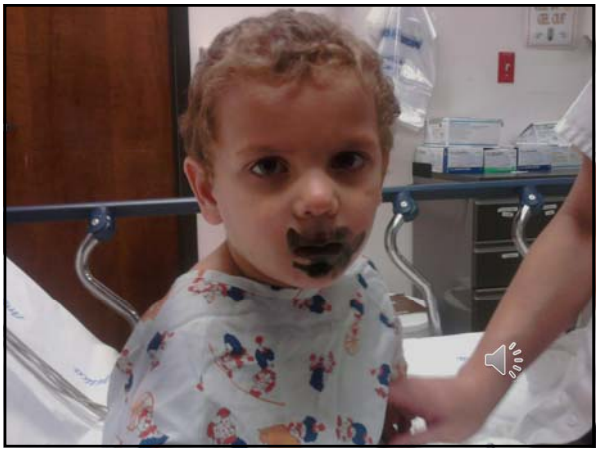
Importance to Education

s the quality of examinations by
he quality of the items
minations have reliability and

to review exams and
t emotion
cussion that will

at

Importance to Nursing Education



Packers

References

- Clifton, S. L. & Schriner, C.L. (2010). Assessing the quality of multiple-choice test items. *Nurse Educator*, *35* (1), 12-16. doi: 10.1097/NNE.0b013e3181c41fa3
- Malau-Aduli, B.S. & Zimitat, C. (2012). Peer review improves the quality of MCQ examinations, *Assessment & Evaluation in Higher Education, 37 (8), 919-931.* doi.org/10.1080/02602938.2011.586991
- McGahee, T. W & Ball, J. (2009). How to read and really use an item analysis. *Nurse Educator*, *34* (4), 166-171. doi: 10.1097/NNE.0b013e3181aaba94