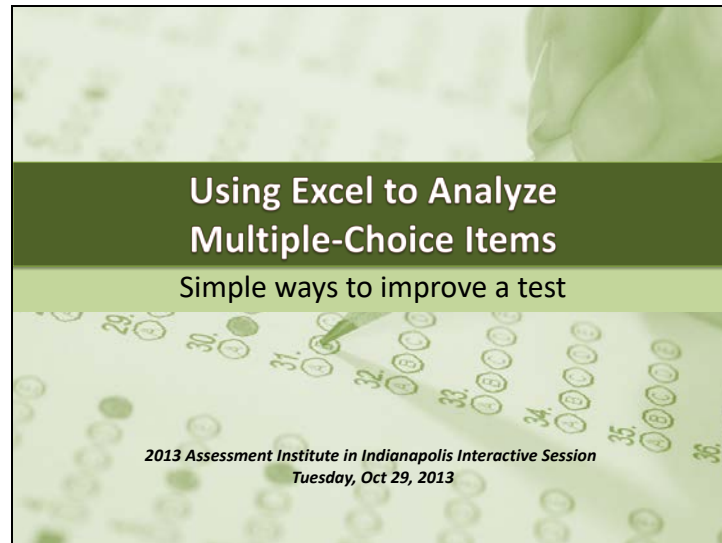
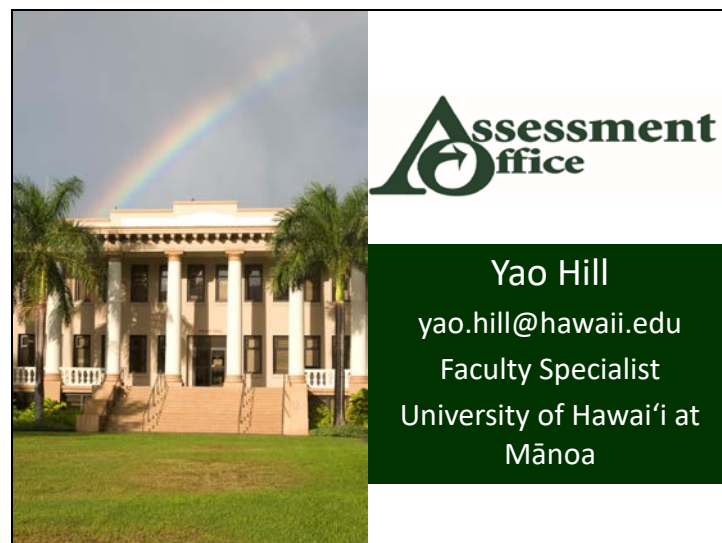


Slide 1




Hill, Y. Z. (2013, October). *Using Excel to analyze multiple-choice items: Simple ways to improve a test.* Interactive session presented at the IUPUI Assessment Institute Annual Conference, Indianapolis, IN.

Slide 2




Slide 3



### Intended outcomes

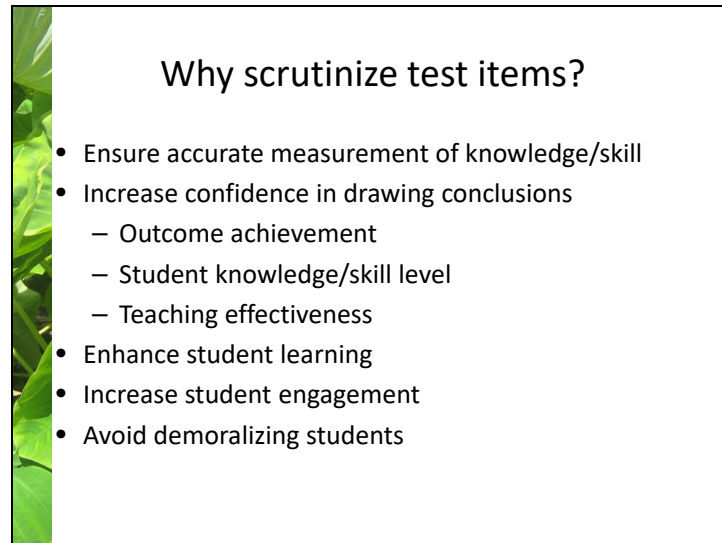
1. Articulate the purpose and need
2. Define the statistics:
  1. Item facility
  2. Difference index
  3. B-index
  4. Distractor efficiency index
3. Apply item analysis in Excel
4. Identify the proper index to use.
5. Interpret results

Slide 4



### Workshop agenda

1. Purpose
2. Introducing item facility, difference index, & B-index
3. Excel demo 1
4. Introducing distractor analysis
5. Excel demo 2
6. Introducing item analysis software
7. Questions & Evaluation



### Why scrutinize test items?

- Ensure accurate measurement of knowledge/skill
- Increase confidence in drawing conclusions
  - Outcome achievement
  - Student knowledge/skill level
  - Teaching effectiveness
- Enhance student learning
- Increase student engagement
- Avoid demoralizing students

When we develop a multiple-choice test, we include many items on the test. For each item that we include on the test, we have an idea what it should measure. Conducting item analysis give us clues on whether the item measures what we want it to measure. If we are happy with how item functions, then we have higher confidence in making inferences based on the results. We will be more confident to say whether students achieved a certain outcome because they all scored high on, say, item 1, 2, 3 that are related to that outcome. We will be more confident to judge student knowledge and skill levels. And when we use the test before the class and after the class, if we see a lot of improvement on the post-test, it provides strong evidence for teaching effectiveness.

A test in and of itself is a great learning tool. Through the thought-process examining each answer choice and deciding whether each one is correct or not, students will need to recall and apply knowledge and skills. A good multiple-choice will intellectually engage students, for them to identify what they know, what they don't know, and what they still feel not so sure about. On the other hand, a bad test can be very demoralizing. Our office actually received a phone call from a student complaining about the multiple-choice test given by an instructor. She said: half of the questions on the test are not related to what the teacher taught in the class. 2/3 of the students got D or F on the test. She felt the teacher was very unfair, that the students were mistreated, and she was extremely frustrated. This is case, the test severely demoralized students, which is quite detrimental to their learning experience.

Slide 6

Anatomy of a multiple choice item

• Which city is the U.S. capital? **Stem**

(A) Seattle  
(B) New York  
(C) Los Angeles  
(D) Washington DC → **KEY**

**distractors**

**Options/choices**



$v \rightarrow 1; x \rightarrow 0$

Testing and item analysis is an academic field of its own. So there are names for each part of the item.

Slide 7


Conduct Content & Format Review  
Before Statistical Analysis:

- Unintentional clues avoided?
- Distractors plausible?
- Redundancy avoided in the options?
- Ordering of the options carefully considered?
- Correct answers randomly assigned?
- Eliminated “none of the above” and “a. and b. only” options?



There are many components of item analysis. The first step is content and format analysis of an item.

## Slide 8



**How Can a Multiple-Choice Item Go Wrong?**

1. Not assessing target outcomes/knowledge/skills.
  - Background knowledge, intuition, guessing
2. Assessing unintended knowledge/skills.
3. Test-taking strategies alone can answer an item correctly
4. Confusing stem or options.
5. Multiple correct answers.
6. A distractor is too close to being correct.

- The item in the earlier slide cannot be used to assess the geographic location of Washington D.C.. If you are teaching a geography class and use that item to assess students, you may find that student scores on the pre-test and post test are not different from each other. You may also find that students with high level of achievement scored the same with students with low level of achievement. Item analysis can help us to identify problematic problems and the statistics can help us investigate the reasons behind a problematic item.
- Many of the word problems to assess math skills confound math skills and English ability for English language learners. Item analysis can show that highly achieving students scored equal or lower than lower achieving students.

### Item Facility

- Definition: % of students who answered the item correctly.

$$IF = \frac{\text{Number Correct}}{\text{Total Number}}$$

- Range: 0 – 1
  - 0: no one answered correctly. (Pre-test)
  - 1: everyone answered correctly. (Post-test)
  - 0.50: half answered correctly.

An ideal item for assessing achievement is one that has an IF of .00 at the beginning of instruction and an IF of 1.00 at the end of instruction. Such pretest and posttest IFs indicate that everyone missed the item at the beginning of instruction (that is, they needed to study the content or skill embodied in the item) and everyone answered it correctly at the end of instruction (that is, they had completely acquired whatever was being taught). Of course, this example is an ideal item, in an ideal world, with ideal students, and an infallible teacher.

Slide 10

item facility = No. Correct / Total

Students	Item1	Item2	Item3	Item4	Item5
01_Robert	1	1	1	1	1
02_Millie	1	0	1	1	1
03_Dean	1	0	0	1	1
04_Shenan	1	1	0	1	1
05_Cuny	1	1	1	1	1
06_Corky	1	0	1	1	1
07_Randy	1	1	0	1	1
08_Jeanne	1	1	0	0	1
09_Iliana	1	1	1	0	1
10_Lindsey	0	0	0	0	1
Item Facility =	0.9	0.6	0.5	0.7	1.0

9/10

Which one is the easiest?

Which one is the most difficult?

Does this more like pre-test result or post test result?

Which item would be worrisome if this is for the post test?

Slide 11

### Comparing performance on items

- Pre- and post-test comparison – Difference index
- Contrasting group comparison – *B*-index

Slide 12

### Difference Index (DI)

- Definition: The difference in item facility between the pre- and post-tests.

$$DI = IF_{post} - IF_{pre}$$

- Possible range: - 1 to 1
  - 1: Post – all x; Pre – all v ☹
  - 1: Post – all v; Pre – all x ☺
- Acceptable value: higher than 0

Slide 13

### Difference Index = $IF_{post} - IF_{pre}$

Items	1		2		3		4		5	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Monica	0	1	1	1	0	0	1	0	0	1
Yao	0	1	0	1	0	0	1	0	1	1
Jenna	0	1	1	1	0	0	1	0	0	1
Katie	0	1	0	1	0	0	1	0	0	1
<i>IF</i>	0.0	1.0	0.5	1.0	0.0	0.0	1.0	0.0	0.25	1.0
<i>DI</i>	1.0		0.5		0.0		-1.0		0.75	



Slide 14

### B-index

- Definition: The difference of item facility between the those who succeeded (masters) and those who failed the test (non-masters).
 

$$B\text{-index} = IF_{\text{master}} - IF_{\text{non-master}}$$
- Possible range: - 1 to 1
  - 1: Masters – all x; Non-masters – all v ☹️
  - 1 : Masters – all v; Non-masters – all x ☺️
- Acceptable value: higher than 0

Slide 15

$B\text{-Index} = IF_{\text{master}} - IF_{\text{non-master}}$

Students	Item1	Item2	Item3	Item4	Total
01_Robert	1	0	1	1	90
02_Millie	1	0	1	1	85
03_Dean	1	0	1	1	80
04_Shenan	1	0	0	1	80
05_Cuny	1	0	1	1	70
					70% cut-point
06_Jeanne	0	1	0	1	60
07_Iliana	0	1	1	1	40
08_Lindsey	0	1	0	1	20
IF <sub>master</sub>	1.0	0.0	0.8	1.0	
IF <sub>nonmaster</sub>	0.0	1.0	0.3	1.0	
B-index	1.0	-1.0	0.5	0.0	

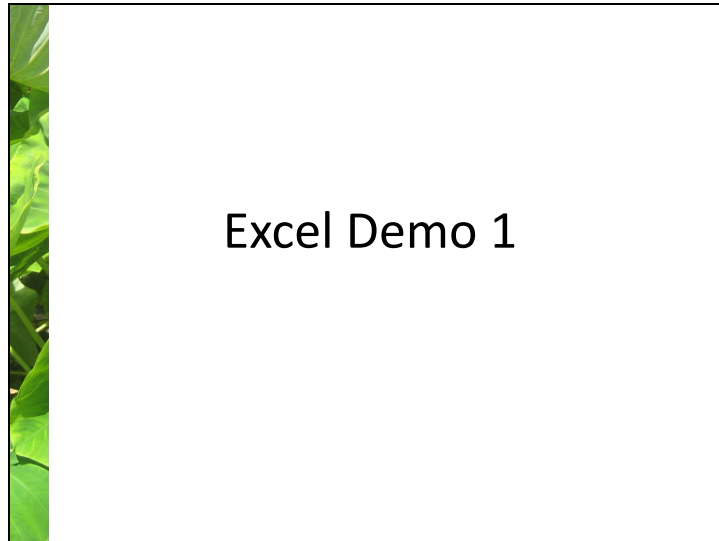
Which one is the easiest?

Which one is the most difficult?

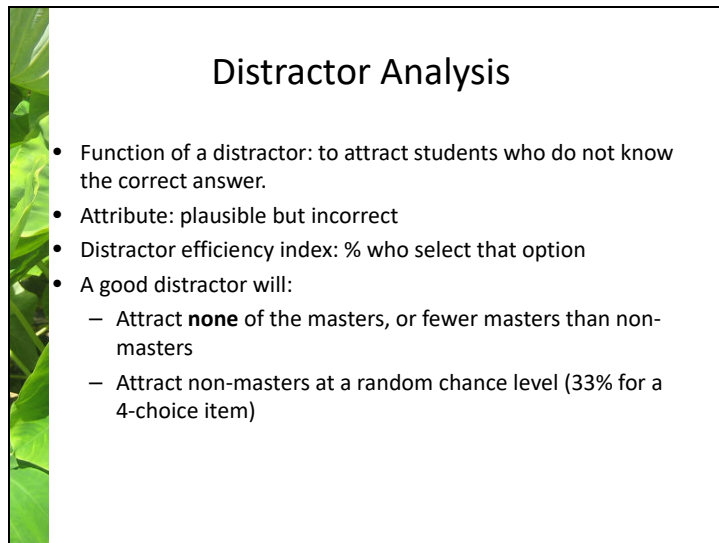
Does this more like pre-test result or post test result?

Which item would be worrisome if this is for the post test?

Slide 16



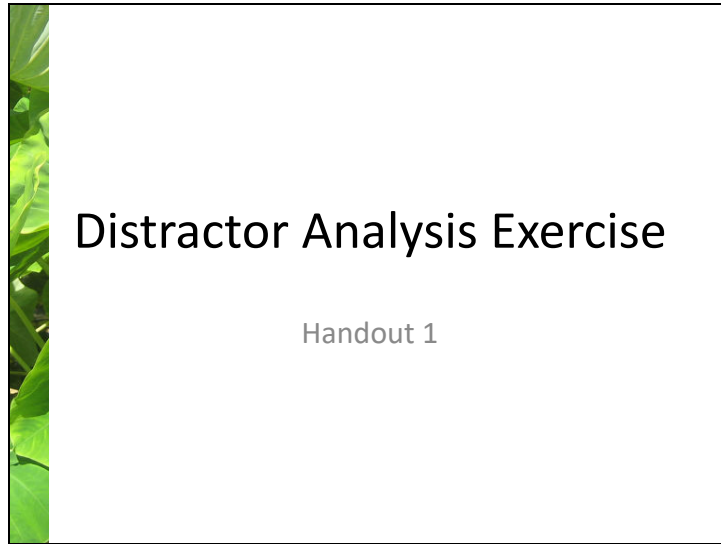
Slide 17

A rectangular slide with a black border. The left side features a vertical strip of green leaves. The main area is white with the title "Distractor Analysis" centered at the top. Below the title is a bulleted list of points.

### Distractor Analysis

- Function of a distractor: to attract students who do not know the correct answer.
- Attribute: plausible but incorrect
- Distractor efficiency index: % who select that option
- A good distractor will:
  - Attract **none** of the masters, or fewer masters than non-masters
  - Attract non-masters at a random chance level (33% for a 4-choice item)

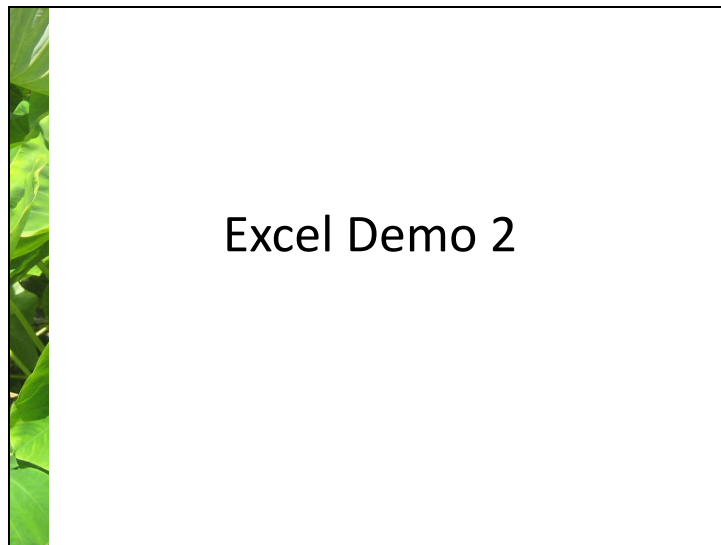
Slide 18

A rectangular slide with a black border. On the left side, there is a vertical strip of green leaves. The main area is white and contains the text "Distractor Analysis Exercise" in a large, bold, black font, and "Handout 1" in a smaller, grey font below it.

**Distractor Analysis Exercise**


Handout 1

Slide 19

A rectangular slide with a black border. On the left side, there is a vertical strip of green leaves. The main area is white and contains the text "Excel Demo 2" in a large, bold, black font.

**Excel Demo 2**


Slide 20



### Item analysis software

- TAP (free):  
[http://www.ohio.edu/people/brooksg/tap\\_download.htm](http://www.ohio.edu/people/brooksg/tap_download.htm)
- CITAS (free):  
<http://www.assess.com/xcart/product.php?productid=407>
- Iteman 4 (demo version limited to 50 items and 50 examinees):  
<http://www.assess.com/xcart/product.php?productid=417&download=1&url=Iteman4212.zip>
- Web-based Attainment Calculator:  
<http://attainmentcalculator.fancyfoxpublications.net/Start.aspx>

Slide 21



### Alternative terms

- Item Facility → item difficulty,  $p$  (proportion) value
- Difference Index → Instruction sensitive item analysis (Crocker & Algina, 2008), intervention strategy (Brown, 1996)
- B-index → Differential group strategy (Brown, 1996)
- Distractor analysis → distractor efficiency analysis

## Resources

1. Brown, J.D. (1996). *Testing in language programs*. Upper Saddle River, NJ: Prentice Hall.
2. Crocker, L., & Algina, J. (2008). *Introduction to classical & modern test theory*. Mason, OH: Cengage Learning.
3. Elvin, C.(n.d.). *Test item analysis using Microsoft Excel spreadsheet program*. Retrieved from <http://www.eflclub.com/elvin/publications/2003/itemanalysis.html>
4. Fulcher, G. (n.d.). *Excel spreadsheets for classical test analysis*. Retrieved from <http://languagetesting.info/statistics/excel.html>
5. Matlock-Hetzel, S. (1997). *Basic concepts in item and test analysis*. Retrieved from <http://www.ericae.net/ft/tamu/Espy.htm>
6. Quirante, S. (n.d.). *Item & distracter analysis*[PowerPoint slides]. Retrieved from <http://www.slideshare.net/suequirante/item-and-distractor-analysis>

## Excel Demo 1

### Part 1: Calculating Item Facility (IF)

1. Open the sheet **IF\_Data** in your data file: **Item\_Analysis\_Practice1**.
2. Prepare your data. Student names or IDs should be in Column A. Each item occupies a column. An item is scored 1 if it is correctly answered and 0 for an incorrect answer. The last column shows the total score for each student.
3. Sort the data on the **Total** variable in the descending order (largest to smallest).
4. To calculate *IF* of Item 1 for the 16 test-takers, in Cell B18, enter the following formula: = **SUM(B2:B17)/16** and hit enter. An Excel formula starts with the equal sign (=). **SUM** is a function to add up all the item scores specified in the parenthesis. **B2:B17** specifies the data range starting from **B2** and ending at **B17**. The colon (:) translates as "to." The back slash (/) is a division symbol. An alternative formula to use in Cell B18 is =**AVERAGE(B2:B17)**, which provides an average for the values in the range between Cell B2 to B17. Copy Cell **B18** and paste this cell to **C18** to **K18**.
5. Examine the results and identify the best and worst items if these items were used on a post-test.

### Part 2: Calculating Difference Index

1. Open the sheet **ID\_Data** in your data file.
2. The *IFs* are already calculated for the pre- and post-tests. Calculate the difference by entering the following formula for Item 1 in Cell **F3**: =**B3-D3**. The formula says: deduct the value in **D3** from the value in **B3**.
3. Copy the formula in Cell **F3** to the rest of the items (ranging from **F4** to **F22**).
4. Examine the results and identify the best and worst items.

### Part 3: Calculating *B*-Index

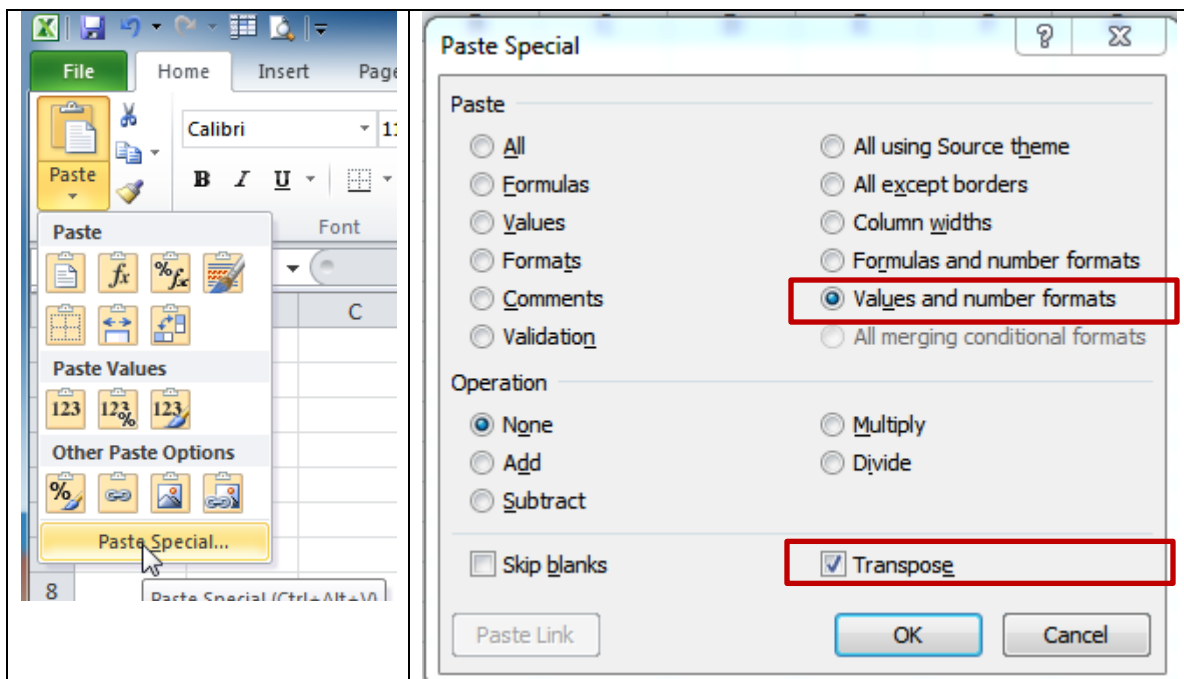
1. Open the sheet **B-Index\_Data** in your data file.
2. Notice that the students' scores are already sorted in the descending order. Those who passed the course and those who failed have been identified and separated into two groups: masters and non-masters.
3. Calculate *IF*<sub>master</sub> for Item 1: In Cell **B25**, type the formula: =**AVERAGE(B4:B17)**.
4. Calculate *IF*<sub>non-master</sub> for Item 1: In Cell **B26**, type the formula: =**AVERAGE(B19:B24)**.
5. Calculate *B*-index for Item 1: In Cell **B27**, type the formula: =**B25-B26**.
6. Select cells **B25**, **B26**, and **B27**, copy, and paste to the cells for the rest of the items.
7. Examine the results and identify the best and worst items.

### Excel Demo 2: Distractor Analysis

1. Open the sheet **distractor** in the file **distractor analysis template**. This sheet has the raw data.
2. Copy the data including ids and responses for Item 1 to Item 10. Be sure to include the first row with column headers. Paste the data in the top portion of the second sheet **calculation\_template**. You can see the results automatically updated at the bottom portion.
3. The sheet **calculation\_template** was split into two windows. The top window from Row 1 to Row 1001 is the section that you can paste your own data. The template allows for up to 1000 examinees and up to 300 items. The distractor efficiency indices are calculated automatically in row 1008 to Row 111.
4. Copy the distractor efficiency indices (Row 1008 to Row 1011)

1007	Count D	7	5	2	3	7	4	8	5	8
1008	% A	50.0%	45.0%	0.0%	40.0%	30.0%	5.0%	10.0%	25.0%	5.0%
1009	% B	10.0%	0.0%	85.0%	30.0%	30.0%	10.0%	25.0%	10.0%	20.0%
1010	% C	5.0%	30.0%	5.0%	15.0%	5.0%	65.0%	25.0%	40.0%	35.0%
1011	% D	35.0%	25.0%	10.0%	15.0%	35.0%	20.0%	40.0%	25.0%	40.0%
1012										

5. **Paste special as values and number format + Transpose** in the sheet **Report**.



### Distractor Analysis Exercise Handout

Distractor Efficiency			Options			
Item Number	IF	Group	<i>a.</i>	<i>b.</i>	<i>c.</i>	<i>d.</i>
1	0.94	High	1.00*	0.00	0.00	0.00
		Low	0.80*	0.20	0.00	0.00
2	0.56	High	0.40	0.00	0.60*	0.00
		Low	0.13	0.07	0.60*	0.20
3	0.44	High	0.12	0.60*	0.13	0.15
		Low	0.21	0.20*	0.27	0.32
4	0.50	High	1.00*	0.00	0.00	0.00
		Low	0.00*	0.34	0.32	0.34
5	1.00	High	0.00	0.00	0.00	1.00*
		Low	0.00	0.00	0.00	1.00*
6	0.44	High	0.06	0.00	0.80*	0.11
		Low	0.49	0.00	0.20*	0.31
7	0.50	High	0.00*	0.80	0.08	0.12
		Low	1.00*	0.00	0.00	0.00
8	0.63	High	0.08	0.12	0.80*	0.00
		Low	0.20	0.19	0.40*	0.21
9	0.38	High	0.72	0.08	0.00	0.20*
		Low	0.13	0.13	0.14	0.60*
10	0.00	High	0.84	0.00*	0.13	0.03
		Low	0.17	0.00*	0.37	0.46

\*Correct option.

Adapted from Brown (1996, p. 72).



### List of Item Analysis Indices Covered in the Workshop

Item Analysis Index	Definition in Words	Calculation Formula
<b>Item Facility (IF)</b>	% of students who answered the item correctly	$= \frac{\text{Number Correct}}{\text{Total Number}}$
<b>Difference Index (DI)</b>	Difference in IFs between the pre and post-tests	$= IF_{post} - IF_{pre}$
<b>B-Index</b>	Difference in IFs between the masters and non-masters	$= IF_{master} - IF_{non-master}$
<b>Distractor Efficiency Index</b>	% of examinees who chose that option	$= \frac{\text{Number who chose}}{\text{Total Number}}$

### List of Excel Formulas Covered in the workshop

Formula	Example	Explanation
<b>=SUM(DATA RANGE)</b>	=SUM(B2:B17)	Add up the values in the data range from B2 to B17
<b>=AVERAGE(DATA RANGE)</b>	=AVERAGE(B2:B17)	Average the values in the data range from B2 to B17
<b>=COUNTA(DATA RANGE)</b>	=COUNTA(B2:B17)	Count the number of text values in the data range from B2 to B17.
<b>=COUNTIF(DATA RANGE,CRITERION)</b>	=COUNTIF(B2:B17,"A")	Count all the occurrences of text A in the data range from B2 to B17. If the criterion is a number, don't use the quotation mark around it.
<b>=CELL A/CELL B</b>	=B2/B17	The value in B2 divided by the value in B17

#### Common rules of Excel formula:

1. Always start with the equal sign (=).
2. Specify data range in the parenthesis.

#### Item Analysis Software:

- TAP (free): [http://www.ohio.edu/people/brooksg/tap\\_download.htm](http://www.ohio.edu/people/brooksg/tap_download.htm)
- CITAS (free): <http://www.assess.com/xcart/product.php?productid=407>
- Iteman 4 (demo version limited to 50 items and 50 examinees):  
<http://www.assess.com/xcart/product.php?productid=417&download=1&url=Iteman4212.zip>
- Web-based Attainment Calculator:  
<http://attainmentcalculator.fancyfoxpublications.net/Start.aspx>

#### Resources

- Brown, J.D. (1996). *Testing in language programs*. Upper Saddle River, NJ: Prentice Hall.
- Crocker, L., & Algina, J. (2008). *Introduction to classical & modern test theory*. Mason, OH: Cengage Learning.
- Elvin, C.(n.d.). *Test item analysis using Microsoft Excel spreadsheet program*. Retrieved from <http://www.eflclub.com/elvin/publications/2003/itemanalysis.html>
- Fulcher, G. (n.d.). *Excel spreadsheets for classical test analysis*. Retrieved from <http://languagetesting.info/statistics/excel.html>
- Matlock-Hetzel, S. (1997). *Basic concepts in item and test analysis*. Retrieved from <http://www.ericae.net/ft/tamu/Espy.htm>
- Quirante, S. (n.d.). *Item & distractor analysis*[PowerPoint slides]. Retrieved from <http://www.slideshare.net/suequirante/item-and-distractor-analysis>