**Item Development and Analysis Worksheet**

**Student Name:**  **Section:** PSYC421- 

**PART 1: Writing Multiple Choice Test Items (Cohen et al., 2013, pg. 252)**

Develop one multiple choice question that covers content from each of the four chapters listed below. When writing your sample questions, please keep in mind the specifications regarding item construction discussed in the textbook. Also, remember the importance of carefully crafted distractor options. Finally, please limit the number of response options to 4 (1 correct response and 3 distractors), and avoid the options of “all of the above,” none of the above,” or the like. Be sure to indicate which of the response options is the correct one.

**Chapter 3 Multiple Choice Question** (2.5 points)



**Chapter 4 Multiple Choice Question** (2.5 points)



**Chapter 5 Multiple Choice Question** (2.5 points)



**Chapter 6 Multiple Choice Question** (2.5 points)



**PART 2: Item Analysis: Item Difficulty Index** **(Cohen et al., 2013, pg. 263)**

A test is only as good as its questions! When researchers, test constructors, and educators create items for ability or achievement tests, we have a responsibility to evaluate the items and make sure that they are useful and high-quality. The process that we use to evaluate test items is known as Item Analysis. When bad items are identified and eliminated from a test, that increases the efficiency, reliability and validity of the entire test! One way that we can distinguish among good and bad items is with the **Item Difficulty Index**.

**Part 2A: Calculating Item Difficulty**

Using the data below, calculate the **Item Difficulty Index** for the first 6 items on *Quiz 1* from a recent section of PSYC101. For each item, “1” means the item was answered correctly and “0” means it was answered incorrectly. Type your answers in the spaces provided at the bottom of the table. (1 pt. each)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **PSYC101 Quiz 1 Item Distribution and Total Scores** | | | | | | | |
| **Examinee** | **Item 1** | **Item 2** | **Item 3** | **Item 4** | **Item 5** | **Item 6** | **Total Score** |
| Andre | 1 | 1 | 1 | 1 | 1 | 1 | 16 |
| Allison | 1 | 1 | 1 | 1 | 0 | 0 | 7 |
| Heather | 1 | 1 | 1 | 1 | 0 | 0 | 10 |
| Corey | 1 | 1 | 0 | 1 | 1 | 1 | 17 |
| Christina | 0 | 0 | 1 | 1 | 0 | 1 | 3 |
| Jeffrey | 0 | 1 | 1 | 1 | 0 | 0 | 11 |
| Shawn | 1 | 1 | 1 | 1 | 0 | 1 | 14 |
| Dana | 0 | 0 | 1 | 1 | 0 | 1 | 10 |
| Megan | 1 | 1 | 1 | 1 | 0 | 1 | 13 |
| David | 0 | 1 | 1 | 1 | 0 | 1 | 12 |
| Isabel | 0 | 0 | 0 | 1 | 0 | 0 | 4 |
| Lance | 1 | 1 | 1 | 1 | 0 | 0 | 9 |
| Aliyah | 1 | 1 | 1 | 1 | 0 | 1 | 15 |
| Blaire | 0 | 1 | 1 | 1 | 0 | 1 | 12 |
| Gabriel | 0 | 0 | 1 | 1 | 0 | 0 | 6 |
| **Item Difficulty** |  |  |  |  |  |  |  |

**Part 2B: Calculating Optimal Item Difficulty** (.5 pt. each)

1. For a test item with ***two response options*** (e.g., true/false), what is the probability of selecting the correct answer *by chance*?

 %

1. Calculate the **optimum** level of difficulty for a test questions with ***two response options***.



1. For a test item with ***three response options***, what is the probability of selecting the correct answer *by chance*?

 %

1. Calculate the **optimum** level of difficulty for a test questions with ***three response options***.



1. For a test item with ***four response options***, what is the probability of selecting the correct answer *by chance*?

 %

1. Calculate the **optimum** level of difficulty for a test questions with ***four response options***.



1. For a test item with ***five response options***, what is the probability of selecting the correct answer *by chance*?

 %

1. Calculate the **optimum** level of difficulty for a test questions with ***five response options***.



**PART 3: Item Analysis: Item Discrimination Index** **(Cohen et al., 2013, pg. 265–266)**

Another way that test creators can distinguish between good and bad items is with an analysis called the **Discrimination Index**. The discrimination index measures how well an individual test item distinguishes between high scorers and low scores on the test. An item is considered to be “good” if most of the high scorers get it right, and most of the low scorers get it wrong.

Interpreting the Discrimination Index (d)

* The discrimination index can range from -1.0 to 1.0.
* The closer d is to 1.0, the better the item discriminates between high and low scorers
* The closer d is to 0, the more poorly the item discriminates between high and low scorers.
* An item with a negative discrimination index is considered a “negative discriminator” because more low scorers get the item correct than high scorers.
* A discrimination index of 1.0 means all the high scorers got the item correct and all of the low scorers got it incorrect.
* A discrimination index of -1.0 means all of the low scorers got the item correct and all of the high scorers got it incorrect.
* Items with d’s close to 0 or with negative d’s ought to be eliminated from the test!

**Calculating the Item Discrimination Index (d)**

Calculate the item discrimination index (d) for the 7 hypothetical test items presented below. Type your answers in the spaces provided at the right of the table (1 pt. each).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Item #** | **U** | **L** | **n** | **d** |
| Item 1 | 21 | 17 | 25 |  |
| Item 2 | 23 | 7 | 25 |  |
| Item 3 | 25 | 0 | 25 |  |
| Item 4 | 3 | 24 | 25 |  |
| Item 5 | 22 | 3 | 25 |  |
| Item 6 | 0 | 25 | 25 |  |
| Item 7 | 19 | 6 | 25 |  |

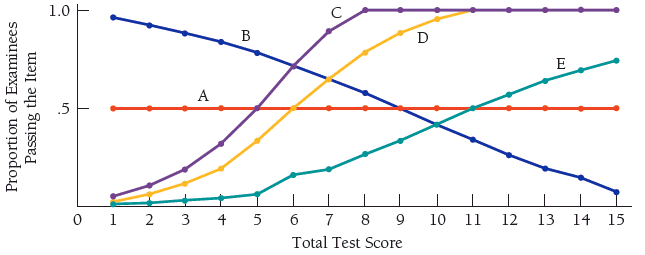
Based on your calculations above, answer the following questions (1 pt. each).

1. Which item discriminates the best? 
2. Which item discriminates most poorly? 
3. Based on your analysis, identify which two items would you choose to eliminate from this test and explain why you would eliminate each.



**Part 4: Item Characteristic Curves** **(Cohen et al., pg. 268–270)**

Another method that test creators can use to assess the usefulness of test items is with **Item Characteristic Curves**. Item characteristic curves provide a graphical depiction of examinees’ performance on individual test items. As indicated in the figure below, Total Test Score is plotted on the x-axis of the graph, while proportion of examinees who got the item correct is plotted on the y-axis.



Using the figure above, provide a written description of how test items A–D discriminate among examinees at various levels of performance. In your responses, discuss why each item would be considered a “good” or a “bad” item. EXAMPLE: “This item discriminates well among high scores, but doesn’t discriminate well among low scorers. So this item would be considered a good item because it discriminates at the highest levels of performance.” (2 pt. each)

|  |  |
| --- | --- |
| Item A: |  |
| Item B: |  |
| Item C: |  |
| Item D: |  |
| Item E: |  |

**Part 5: Qualitative Item Analysis** **(Cohen et al., pg. 272–274)**

Qualitative item analysis refers to a set of non-statistical procedures used to gather information about the usefulness of test items. These analyses typically involve interviews, panel discussions, questionnaires and other forms of verbal exchange with test-takers to explore how individual test items work.

As an online student, you have a very different test-taking experience than residential students. Based on your readings from Chapter 8, identify 4 topics related to online test taking, and create 4 qualitative questions that you could ask online test-takers to gain an understanding of their experiences with test-taking. Also, as students at a Christian institution of higher education, course assignments/assessments are supposed to give students an opportunity to integrate course content with their Christian worldview. Given the topic of *faith and learning*, create one qualitative question that you could ask test-takers.

|  |  |
| --- | --- |
| **Topic** (1 pt. each) | **Sample Question for Test-Takers** (1 pt. each) |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

Assignment Scoring (Instructor Use Only)

Part 1 Subtotal:      /10

Part 2 Subtotal:      /10

Part 3 Subtotal:      /10

Part 4 Subtotal:      /10

Part 5 Subtotal:      /10

**TOTAL SCORE:** **/50**