

UNIVERSITI SAINS MALAYSIA

Peperiksaan Semester Pertama
Sidang Akademik 2003/2004

September/Oktober 2003

HXE 309 – Testing and Evaluation

Masa : 3 jam

THIS EXAMINATION PAPER COMPRISES **TWO** PAGES.

Answer any FOUR questions.

1. Study the Item Statistics table below and then answer the questions that follow:

	Item Number									
Item Statistic	1	2	3	4	5	6	7	8	9	10
IF (total)	.94	.58	.44	.50	1.00	.63	.32	.61	.37	.00
IF(upper)	1.00	.67	.60	.100	.80	.75	.19	.46	.20	.00
IF(lower)	.85	.67	.18	.00	.25	.30	.30	.40	.30	.00
ID	.15	.00	.42	1.00	.55	.45	-.11	.06	-.10	.00

- [a] Comment on the **IF(total)** values for all the ten items.
- [b] Comment on the Item Discrimination (ID) values for the following items on the test: **Item 1, Item 2, Item 4, Item 7 and Item 9.**
- [c] If your task was to choose the best four (4) items on the above test, which would you choose? Why?

- [d] Which items should be omitted from the next test? Justify your answer.
- [e] What are the advantages of displaying test statistics in the form of a table?
[100 marks]
2. Most language testing professionals consider the mean and standard deviation as two important statistical devices. Discuss the use of these two devices in describing test results.
[100 marks]
3. Write short notes on any **four (4)** of the following in relation to language testing:
- [a] Productive response items
- [b] The B-index
- [c] Skewed Distributions
- [d] Achievement tests
- [e] T scores
- [f] Distractor Efficiency Analysis
[100 marks]
4. "In recent years, the distinction between norm-referenced and criterion-referenced tests has gradually entered the language testing literature" (Brown, 1996). Discuss how this knowledge can help language teachers design and match tests to decision-making purposes.
[100 marks]
5. In order to investigate content validity, testers must decide whether the test is a representative sample of the content of whatever the test was designed to measure. Discuss, with examples.
[100 marks]
6. "Measurement scales are also used in psychological and educational measurement" (Henning, 1987). Discuss, with examples, how the four major types of psychological measurement scales are used in language testing.
[100 marks]