

**MQM Comprehensive Examination
Measurement Minor**

Day 1

Directions: On this exam you must answer four questions. The first two (in Part I) are required, and you must choose two from among the remaining four questions in Part II. Begin each response on a new page, and clearly number the item to which you are responding.

Part I. Answer both items.

1. Define *consequential aspect of validity* as used by Messick (1989, 1995). Identify two different consequential concerns that might arise when a statewide high school graduation examination is implemented and results from that examination are used to determine whether high school seniors will be awarded a high school diploma. For each concern:
 - A. Describe why that concern represents a threat to the consequential aspect of validity of the graduation test.
 - B. Explain empirical and/or rational bases for evaluating the degree to which the consequential aspect of validity is supported for the graduation test.
2. The concept of information in item response theory (IRT) is closely related to the standard error of measurement in classical test theory.
 - A. Define information from an IRT perspective. Give a formula for information and define all of the terms used in the formula.
 - B. Define standard error of measurement from a classical test theory perspective. Give a formula for standard error of measurement and define all of the terms used in the formula.
 - C. Explain the connection between standard error of measurement and information. Provide at least one mathematical expression that shows the connection.

Part II. Answer two of the following four items.

3. Define the term *reliability* in language suitable for a lay audience. Also, define the term in technical language, providing an equation representing the reliability index as it is formulated in true score test theory, generalizability theory, and item response theory (be sure to identify all terms included in each equation).

4. Provide guidelines for interpreting an items' p-value, D index, and r_{biserial} as they are used in evaluate test items. The table presents item analysis index values for a single four-option multiple-choice achievement test item. Note that option (A) is the keyed (correct) option. State whether you believe it would be appropriate to include this test item on a norm-referenced achievement test, citing specific evidence from the table to support your decision.

Index	Option			
	A	B	C	D
p-value	.15	.35	.25	.25
D	.03	-.05	-.05	-.30
R_{biserial}	.45	-.30	-.25	-.40

5. An educational policy maker wants to set a performance standard for the state 8th grade mathematics exam. This examination consists of multiple-choice questions.
- Make a recommendation for the standard setting method that you think will work well in this situation and give the reasons for your recommendation.
 - Describe in detail the process for using this standard setting procedure beginning with the statement of policy and ending with the approved numerical score on the test.
6. Large scale testing programs regularly run Differential Item Functioning (DIF) analyses as part of their quality control procedures. Respond to the following questions from the context of such testing programs.
- What DIF analysis procedure would you recommend for a large scale testing program? Give the reasons for your recommendation.
 - What does it mean if an item is determined to have significant DIF? Give an answer that would help a policy maker understand what DIF is all about.
 - When would you delete an item that has a large DIF statistic from a test? Give the rationale for your answer.