**EduMP** – 707 Test Theories and Design

This very specialised course is to make research students review test theories (Classical, Generalizability and Item Response) and the history of test development in international perspective. It explains how Item Banking is made possible with the help of item generation using table of specification; and thereafter, the process of standardisation of test items particularly multiple choice questions (MCQs). Use of software is an important characteristic of the course.

**Objectives**

After studying the course, the students will be able to:

1. Review the history of test development in international perspectives
2. Critically review the test theories (Classical, Generalizability and Item Response)
3. Understand the concepts of NRT and CRT
4. Understand the concept of standardized testing
5. Critically review the models of assembling test items in regard to IRT
6. Explain the concepts of item generation and item banking using IRT
7. Critically review the models of assembling multiple tests
8. Understand basics of computer-based test design models (CAT, MST)

**Contents**

|  |  |  |
| --- | --- | --- |
| **1** | **Test Theories in Historical Perspectives** | |
|  | 1.1 | Classical test theory |
|  | 1.2 | Theory of generalizability |
|  | 1.3 | Item response theory |
|  | 1.4 | Comparative view of test theories |
|  | 1.5 | IRT and new notion of test standardization |
|  | 1.6 | Item banking using IRT |

|  |  |  |
| --- | --- | --- |
| **2** | **Test Development Process (both NRT and CRT Perspectives)** | |
|  | 2.1 | Understanding of norm standard / NRT and criterion standard / CRT |
|  | 2.2 | Reviewing course curricula and reading materials |
|  | 2.3 | Writing objectives and competencies (Activity) |
|  | 2.4 | Designing specification grid/test specification (Activity) |

|  |  |  |
| --- | --- | --- |
| **3** | **Test Equating and DIF** | |
|  | 3.1 | Test score equating using IRT |
|  | 3.2 | Differential item functioning (DIF) |

|  |  |  |
| --- | --- | --- |
| **4** | **Models for Assembling Test Items** | |
|  | 4.1 | IRT-based test assembly (absolute and relative targets, cut-off scores) |
|  | 4.2 | CRT-based test assembly (maximizing test validity and reliability) |
|  | 4.3 | Matching observed score distributions |

|  |  |  |
| --- | --- | --- |
| **5** | **Models of Assembling Multiple Tests** | |
|  | 5.1 | Sequential and simultaneous assembly |
|  | 5.2 | Big-shadow test method |
|  | 5.3 | Optimizing Bib design |

|  |  |  |
| --- | --- | --- |
| **6** | **Models of Assembling Tests with Item Sets** | |
|  | 6.1 | Power-set method |
|  | 6.2 | Edited-set method |
|  | 6.3 | Pivot-item method |
|  | 6.4 | Two-stage method |

|  |  |  |
| --- | --- | --- |
| **7** | **Item Generation and computer-based test designs** | |
|  | 7.1 | Foundation of item generation of mass testing |
|  | 7.2 | Item generation models for lower and higher order cognitive abilities |
|  | 7.3 | Computer-based test designs (CAT, MST) |
|  | 7.4 | Foundation of item generation of mass testing |

|  |  |
| --- | --- |
| **8** | **Seminars on Contemporary Relevant Theme** |

**Suggested Readings**

Cizek, G. J. and Sternberg, R. J. (2001) *Setting performance standards: Concepts, methods, and perspectives.* Mahwah, New Jersey: Lawrence Erlbaum Associates, Publishers.

Van der Linden, W. J. (2003) *Linear models for optimal test design: Statistics for social science and behavioural sciences*.