



RULES FOR TEST ITEMS CONSTRUCTION

Professor Dr. Muhammad Saeed

Chairman, Department of Educational Research and Evaluation

Institute of Education and Research

and

Director, Quality Enhancement Cell

UNIVERSITY OF THE PUNJAB

LAHORE

AUGUST 2020

PREFACE

Assessment and evaluation is one of the key components of the educational process. Without effective assessment, classroom teaching cannot be made effective. That is why it is mostly said that ‘teaching is guided by assessment’. A number of the assessment techniques are used by the teachers in the classroom context such as oral questioning, quiz, observations, assignments and presentations, performance-based assessment on hands-on-activities etc. But tests is a more popular form of measuring students’ academic achievement at all levels of educational tier i.e. primary/elementary, secondary, and higher education.

Since these guidelines are basically written for the teachers at university level, therefore the focus will be on the types of test items which are more prevalent in the universities and other higher education institutions (HEIs). At University of the Punjab, under semester system, teachers develop tests at the mid of the semester and at its end, which are usually called as mid-term (35%) and final test (40%). Technically, the term mid-semester is more appropriate than mid-term, as internationally a term has short duration (usually 10-12 weeks) and a semester has 16-18 weeks. Anyhow, 75% of the assessment is usually based on tests, however, at M.Phil/MS and Ph.D levels teachers sometime use alternative means of assessment such as Term Paper and Project depending upon the nature of the discipline and course, but still the importance of tests is no more less. At Ph.D level, there is mandatory comprehensive examination, for which teachers usually develop essay type questions.

Keeping in view the above scenario, as Chairman, Department of Educational Research and Evaluation, and Director, Quality Enhancement Cell, thought to share my experience on Assessment and Testing and to draft this document on rules/principles for test items construction. It contains rules pertaining to multiple choice questions, restricted response items (usually called as short answer questions in our context), and essay type questions. Some examples in the form of questions are also given pertaining to the application of the rules in each type of test items category. These types of test items are more prevalent in the University of Punjab.

It would not be appropriate to acknowledge the visionary leadership of our Vice Chancellor, Professor Dr. Niaz Ahmad Akhter, who has high concern on quality of teaching and assessment. This was also a motive to draft these guidelines. I hope this document will be helpful for my colleagues while developing tests under semester system and annual system and will also provide academic support to teachers engaged in developing question papers in different disciplines for the Examination Department of the University. The postgraduate students interested to develop their skills in test items development may also get benefit of this work.

August 2020

Professor Dr. Muhammad Saeed
Chairman, Department of Educational Research & Evaluation
Director, Quality Enhancement Cell, University of the Punjab.

Table of Contents

Contents	Page No.
Introduction	4
Rules/Principles for MCQs Development	4
Rules/Principles for Short Answer Questions	8
Rules/Principles for Essay Type Question	9
References	11

INTRODUCTION

Constructing tests of good quality is a highly technical academic task. This needs the test developer for having subject command and understanding of basic rules/principles of test items construction i.e. psychometric traits.

In this document, I have shared the important test items construction rules or principles focusing MCQs, restricted response/short answer question and extended response/essay type questions which a test developer in any discipline may consider at least to ensure largely the validity and to some extent reliability of the test. A few advantages and limitations are also briefly mentioned in the opening paragraphs of these three types of test items.

The best way to ensure the content validity of the test is to draw table of specification prior to constructing the test as this will guide the test developer to include the desired contents (vertically, top to bottom), weightage of different cognitive abilities and the categories of questions (MCQs/restricted response or short answer and extended response or essay type) along horizontal side. The readers may consult the recommended books for further details on rules of test construction, table of specification and allied concepts (see References).

RULES FOR CONSTRUCTING MCQs

Multiple choice questions (MCQs) are quite popular while examining students at various educational levels i.e. primary, secondary, and higher. Although time consuming and difficult to construct but have advantages that these can cover the entire syllabi and easy to score/mark. These are considered most suitable form of questions when we intend to assess students' knowledge and comprehension, however application and analysis abilities can also be assessed, especially in the disciplines of sciences.

A multiple choice question (MCQ) has two main components: a) stem/ statement, and b) options/choices/alternatives. At school level, usually four alternatives/options/responses are developed while at higher education level, usually five alternatives are preferred. Amongst the 4 or 5 alternatives, one is the correct/best answer, while others are called 'distractors'.

The key principles for developing MCQs are discussed under four headings: a) Direction/instructions for the attempting MCQs; b) Writing stem/statements; c) Writing alternatives/responses; and d) Some examples of good MCQs reflecting application of the rules.

A. Direction/Instruction

Prior to writing the multiple choice questions, we need to give a clear direction/instruction. This may include as following;

1. Whether the correct/best alternative should be tick mark or encircled. An example may be given in this regard.
2. In case, two or more options/alternatives are encircled, no score will be awarded.
3. If there is negative scoring, then its weightage should be mentioned.
4. Any clue or over-writing/cutting etc. will lead to Zero or any other penalty may be mentioned as per policy of the institution.

B. Rules for Overall Item Writing and Statements/Stem

1. The stem should contain the specific problem/content area. The entire stem must always precede the alternatives/options.
2. All parts of the question should be on the same page.
3. The stem should be preferred in a simple sentence format rather interrogative form. But if interrogative form seems more understandable, that can be used in order to make the question clear for the examinee.
4. Avoid opinion-based items.
5. The stem should be meaningful, simple and clear. There should be no ambiguity. Keep vocabulary simple.
6. The stem should be as short and verbally uncompleted as possible. Give as much context as possible as is necessary to answer the question but do not include superfluous information.
7. Avoid omissions/blanks in the stem. If somewhere needed, these should appear towards the end of the stem.
8. Keep each item independent from other items. Do not give answer away to another item. If items require computation, avoid items that are dependent upon each other.
9. Do not give any clue to the correct answer in the statements.
10. Try to test/measure a different point in each question.
11. The stem should be free of grammatical errors.
12. Avoid extreme words in stem and alternatives such as 'never, 'not at all' 'always' etc.
13. Avoid copying the same statements from any reading material.
14. Avoid negative statements, unless the significant learning outcome is required. If to be used, gather such statements together and make these either capitalized/boldface/Italic. If written by hand then these may be underlined.

C. Rules for Writing Alternatives/Responses/Options

1. The options/alternatives should be plausible i.e. all options should be close to each others, but there should only one correct/best answer.
2. The stem should include as much of the item as possible. If, after writing the item, you notice that each of your alternatives begins with the same word or phrase, incorporate the word or phrase into the stem, deleting it from each of the alternatives. Sometime, use of article, especially 'a or an' leads to guessing, so the test developer needs to take care in terms of decreasing the chances of guessing.
3. Alternatives should be presented in logical order, where applicable (alphabetical, chronological, quantity/ numbers, etc.)
4. All alternatives should be homogenous in terms of content, form and grammatical structure.
5. Alternatives should not overlap in meaning or be synonymous with one another.
6. If one or more alternatives are partially correct, ask for the 'best' answer.

7. Use capital letters A, B, C,... for the responses/alternatives signs.
8. Randomly distribute the correct alternative/response in all MCQs.
9. All alternatives should be of approximately the same length and as brief as possible.
10. All alternatives should be consistent with the stem.
11. Avoid options such as “all of the above, none of the above, Both B and D, Both C and D” etc.
12. Avoid abbreviations, except otherwise internationally accepted by the readers.
13. Review items carefully after developing to keep these free of any conceptual, technical and/or grammatical errors.

D. Some Examples of Relatively Good MCQs

The examples are given below from a course at master/Hons. level, ‘Research Methodology’ as it is assumed that all teachers of higher education institutions possess desired competency in this course.

- **The key steps in scientific method are:**
 - A. Research problem, experiment, findings, recommendations
 - B. Problem, hypothesis, experiment, results, recommendations
 - C. Observation, hypothesis, experiment, conclusions
 - D. Observation, hypothesis, experiment, conclusions, law or theory
 - E. Hypothesis, experiment, results, recommendations

- **The types of sampling techniques associated with each other are:**
 - A. Simple random and quota
 - B. Simple random and stratified
 - C. Purposive and stratified
 - D. Systematic and convenient
 - E. Snow ball and stratified

- **Chi-square distribution is MOST appropriate when sample is:**
 - A. Small size and variables are two
 - B. Moderate and variables are two
 - C. Large and data are dichotomous
 - C. Small and data are nominal scale
 - D. Large and data are at ratio scale

- **For experimental studies, the BEST range of sample is:**
 - A. Less than 15
 - B. 15-25
 - C. 25-35
 - D. 35-45
 - E. 45-above

- **The qualitative data is gathered through:**
 - A. Interviews and structured observations
 - B. Focus group discussion and interviews
 - C. Content analysis and rating scale
 - D. Likert scale and focus group discussion
 - E. Likert scale and Guttman scale

- **When we have to find the difference in the mean scores of two variables of the same group at different intervals, the suitable inferential statistics is:**
 - A. Paired samples t-test
 - B. Independent samples t-test
 - C. Chi-square distribution
 - D. One way analysis of variance
 - E. Two way analysis of variance

- **The BEST method for measuring reliability of a questionnaire when items are at Likert type scale is:**
 - A. Test retest
 - B. Parallel forms
 - C. KR-20
 - D. Cronbach Alpha
 - E. Inter-rater

- **At the time of proposal development, the researcher may narrow scope of research keeping in view resources and time factor. This in the context of research, is called:**
 - A. Assumption
 - B. Hypothesis
 - C. Limitation
 - D. Limit of the study
 - E. Delimitation

- **Which edition of the American Psychological Association style is currently prevalent in the world?**
 - A. 9th
 - B. 8th
 - C. 7th
 - D. 6th
 - E. 5th

- **The book ‘Educational Research: Competencies and Analysis’ is authored by:**
 - A. Bailey
 - B. Cohen
 - C. Creswell
 - D. Gay
 - E. Fraenkal

RULES FOR CONSTRUCTING SHORT ANSWER QUESTIONS

Short answer questions are quite popular in the sense that many concepts can be covered and scoring is not as difficult as in essay/long answer questions. If designed skillfully, assess students' knowledge, understanding/comprehension, and problem solving abilities. These are also called as restricted response items/questions.

1. The marks of short answer questions shall be determined on the basis of length and difficulty level of each question. If each question carries different marks, then these must be indicated at the end of each question.
2. Where needed, the limit of words or lines may be indicated to guide the respondent to be brief and concise. But this must match with the marks allocated to such type of question/s.
3. Questions of different cognitive abilities, especially problem solving and analytical reasoning should be focused. If a teacher is preparing a question paper and intends to add MCQs as well, then he/she should prefer that short answer questions should be at least above Knowledge level which only assesses memory and recall. This is because it is somehow difficult to develop such MCQs that can assess students' higher cognitive abilities.
4. The questions should not of the similar nature i.e. use what, why or how in different questions. Use different verbs/words like differentiate, compare, describe, explain, identify etc., to assess different cognitive abilities of the students.
5. The questions should be asked from the entire curriculum or syllabus of the prescribed course.
6. There should be no ambiguity in the statements of the questions. These should be free of all grammatical errors. Use simple vocabulary, except particular/technical terms which are used in a certain discipline
7. Questions should not be quoted exactly from any reading material. These should be rephrased in your own words properly.
8. There should be some space of subjectivity in the responses. It means the respondent may have the choice to explain the concept in his/her own way, rather a fixed answer.
9. A Marking Key/Scoring rubric should be developed for each question at the time of constructing test, as afterwards some points the test developer may forget.
10. After developing all questions should be revised carefully to ensure: a) Are different cognitive levels addressed? b) .Do entire course outline covered? , c) Do the questions grammatically and conceptually error-free, and d) Do desired direction/instruction clear at the beginning?

Sampled short answer questions in the course of Research Methodology

Direction/Instruction: Each question carries THREE marks

1. Distinguish between quantitative and qualitative research by giving an example of each.
2. Draw a flow chart showing steps for data analysis of semi-structured interviews on assessing teachers' classroom teaching.
3. State two significant problems of action research nature in the context of course/subject of your interest.
4. Why literature review is considered important in a research work?
5. Define co-relational research. What kind/s of statistics is generally applied to test a hypothesis of co-relational research?

RULES FOR CONSTRUCTING ESSAY TYPE QUESTIONS

Essay type questions are also called as extended response or subjective type or long answer questions. They are constructed to assess the students' higher cognitive abilities such as analysis, synthesis and evaluation. Essay type questions can also be developed that assess lower cognitive abilities such as knowledge, comprehension and application, but literature does not recommend it because these lower level abilities can be assessed in effective way through objective type (especially MCQs) and restricted response/short answer questions. Especially, at higher education level, including essay type questions (more or less) in the test is recommended.

Essay type items are relatively easy to construct by the teachers in comparison to MCQs but are difficult to score as the examinees are also required to give their detailed point of view. As by constructing only essay type questions, the entire course curriculum is difficult to cover, so in such cases, the test developer should be more vigilant.

The key rules/principles of constructing essay type questions are as following;

1. The direction / instruction should be clear. For example, if all essay type questions are of the same marks, then it should be indicated at the beginning. If each essay type question carries different marks, then these must be indicated at the end of each question. The allocation of marks should be in accordance with the length of answer desired and difficulty level in terms of cognitive ability to be assessed.
2. The time duration should also be mentioned keeping in view the number of questions and their difficulty level and description.
3. Where applicable, the direction may also be given to focus the particular dimension so that the discussion may be relevant.
4. Restrict the use of essay questions to those learning outcomes that cannot be measured by objective type items. These should focus higher cognitive abilities where true/false, completion items, MCQs and short answer type questions do not serve the purpose.

5. Write the questions clearly enough for the students to understand the task they are required to carry out. These should be free of grammatical errors.
6. Start essay questions with such words or phrase as compare, contrast, give reason, predict what would happen if, analyze, synthesize, appraise etc. Sometimes you may write a statement and ask the respondent to argue, discuss, give your arguments, comment, analyze or prove logically, etc.
7. Avoid giving choice in essay questions. Students should be asked to attempt all questions in the test. Otherwise it decreases the validity and the basis for comparison among students.
8. Avoid assessing those concepts which have already included in the short answer questions (provided if such types of questions are included in the test).
9. Make sure questions are sharply focused on a single key concept/issue. Do not give examinee too much freedom in determining what the answer should be? But they should be allowed to search for ideas and concepts that are not restricted to a predetermined set of responses.
10. A Marking Key/Scoring rubric should be developed for each question at the time of constructing items, as afterwards some points the test developer may forget.
11. After developing all questions should be revised carefully to ensure: a) Are different cognitive levels addressed? b) .Do difficult areas of the course outline covered? , c) Do the questions grammatically and conceptually error-free, and d) Do desired direction/instruction clearly given at the beginning?

Some Examples of Essay type Questions in the Course of Research Methodology

1. “Research is creation of new knowledge”. Give your logical point of view on this version. (5)
2. Describe the characteristics of a good literature review? (5 marks)
3. What ethical considerations a researcher should keep in mind at the stage of data collection, analysis and reporting results of a study? (5)
4. Distinguish between single and multiple case study. Identify a research problem for each type of these types of case study and write two or three relevant objectives for each. (7 marks)
5. Discuss in detail the procedure of developing a questionnaire to measure students’ satisfaction on learning environment in the university. (8 marks)

References

- Athanasou, J. & Lamprianou, I. (2002). *A teacher's guide to assessment*. Thomson-Social Science Press.
- Clay, E. & Root, E. (2001). *Is this a trick question: A short guide to writing effective questions*. Kansas Curriculum Center, Kansas State Department of Education.
- Cohen, A. & Wallack, J.A. (n.d.). *Handbook on test development: helpful tips for creating reliable and valid classroom tests*. Testing and Evaluation Services, University of Wisconsin-Medison. .
- Downing, S. M. & Haladyna, T. M. (2006). *Handbook of test development*. Lawrence Erlbaum Associates, Publishers.
- Fulcher, G. & Davidson, F. (2007). *Language testing and assessment – An advanced resource book*. Routledge.
- Gronlund, N. E. (2006). *Assessment of student achievement*. New York: Pearson Education.
- Gronlund, N.E. & Linn, R.L. (2010). *Measurement and evaluation in teaching*. Maxwell Macmillan Publishers.
- Hussain, H., Bias, B., Hussain, A., & Samad, S.A. (2011). How to construct open ended questions. *Procedia – Social and Behavioural Sciences*, 60, 456-462. doi: 10.1016/j.sbspro.2012.09.406
- Kar, S.S., Lakshminarayanan, S. & Mahalakshmy, T. (2015). Basic principles of constructing multiple choice questions. *Indian Journal of Community and Family Medicine*, 1(2), 65-69. <http://www.researchgate.net/publication/304349654>.
- Kubiszyn, T. & Borich, G. (2004). *Educational testing and measurement: Classroom application and practice*. John Wiley & Sons.
- Nasir, M. & Saleem, T. (2003). *Test development: Item writing, analysis and banking*. Unpublished document for 8-week training workshop for the faculty of Virtual University of Pakistan. (Technical review team was Prof. Dr. Rizwan Akram Rana & Dr. Muhammad Saeed)
- Saeed, M. (2002). *Assessment and examinations: A manual for teachers and teacher educators*. Lahore: Directorate of Staff Development (now Quaid-i-Azam Academy for Educational Development).

Note: For any suggestions/comments or help you may contact at saeed.ier@pu.edu.pk / director.qec@pu.edu.pk

This document is free for professional use but not allowed for commercial purpose.

Professor Dr. Muhammad Saeed

Chairman, Department of Educational Research and Evaluation, IER /
Director, Quality Enhancement Cell,
University of the Punjab, Lahore.