



CONSTRUCTION AND DEVELOPMENT OF AN ACHIEVEMENT TEST; THE ACHIEVEMENT TEST FOR PROSPECTIVE TEACHERS OF EDUCATIONAL GUIDANCE AND COUNSELLING (EGCAT)

Fiza Fatima
ffiza8743@gmail.com

MPhil Scholar, Department of Education, University of Gujrat,
Pakistan.

Dr. Saira
drsaira.ijaz@uog.edu.pk

Lecturer, Department of Education, University of Gujrat, Pakistan.

Abstract

An achievement test is a measure of student's learning progress and performance. This study was aimed at developing and standardizing an Achievement Test, namely achievement test in Educational Guidance and Counselling (EGCAT), to measure prospective teachers' educational guidance and counselling concepts. The test was based on five main areas: basic knowledge, application of knowledge, communication skills, ethical practices and self-improvement techniques. The test was a mixed-type (MCQs and Descriptive). The first draft was read by the subject specialists for clarity and relevance. The test was then pilot tested and statistically analyzed with aspiring teachers from various schools after expert evaluation. Items were subjected to item analysis procedures with respect to difficulty value and discrimination power to improve the test items by deleting, modifying or replacing as appropriate. Face and content validity of the instrument were established by 10 teacher educators. Split-half reliability was calculated by Spearman–Brown Prophecy Formula to determine reliability. Final test reliability coefficient was 0.89 which is high reliability. A representative sample of 60 prospective teachers were used for the collection of quantitative data and norms were set by using z-scores and percentile ranks. The aim of this paper is to report the full development and calibration of the Achievement Test of Educational Guidance and Counselling (ATEC).

Keywords: *Achievement Test, Educational Guidance and Counselling, Prospective Teachers, Standardization.*

Corresponding Author: Fiza Fatima (MPhil Scholar, Department of Education, University of Gujrat (UOG), Pakistan.)

Email: ffiza8743@gmail.com

1. Introduction

In Educational and Psychological Research, the assessment of academic achievement is an important aspect that allows the measurement of knowledge, skills and competencies acquired by the learner in a given subject area. Students' mastery and educational progress are typically measured with achievement tests. The tests allow teachers and scientists to assess the outcomes of learning and take decisions about education.

Achievement tests are standardized tests designed to measure the knowledge and skills that learners have gained in a specific subject area. Often these tests are applied for assessing comprehension in disciplines like mathematics, science, language and social studies in schools and universities. Achievement is a result of instruction and experience that demonstrates a person's level of proficiency and mastery. Thus, a measuring instrument must be reliable, valid and objective.

According to educational psychologists, achievement tests are instruments which measure what has been achieved at the end of the course of instruction. These tests are used to determine students' present performance levels, and can assist with placement, promotion and academic evaluation. They also offer teachers useful feedback on the effectiveness of the teaching methods used.

In the present study, the achievement test was constructed, and standardized in the field of Educational Guidance and Counselling. This was to assess the prospective teachers' knowledge and understanding of the concepts of guidance and counselling in general.

2. Literature Review

Achievement testing is one of the key pillars of educational assessment, which involves evaluating how much students know, can do and have been able to learn as a result of a course of study. Achievement tests are especially significant in teacher preparation programs, as they are not just used to assess student knowledge but also to assess teacher readiness for the task of teaching. But in the sphere of Educational Guidance and Counselling (EGC), achievement testing comes with greater meaning as it evaluates both theoretical knowledge and applied skills in counselling for would-be teachers. To develop an Educational Guidance and Counselling Achievement Test (EGCAT) system, it is important to have a solid base of literature, both in educational measurement and test construction procedures, as well as in validity and reliability, and teacher competency assessment.

Achievement tests typically refer to standard tests which are administered to assess a person's achievement in a particular area of learning. The tests are frequently administered in schools, colleges, and in professional training to assess the effectiveness of teaching and learning goals. Measurement experts say the purpose of achievement tests

is to measure “what has been learned” not aptitude or natural ability. In learning environments, they can be used as diagnostic tools to identify learning gaps, to assess teaching success, and to provide evidence of achievement levels. In teacher preparation programs, achievement tests are used as a means to assure prospective teachers have the necessary content and teaching skill knowledge.

Achievement testing has undergone a tremendous evolution over the years. In a different era of schooling, assessment was mostly reliant on traditional-formal essay tests focusing on recalling and memorizing facts. But, contemporary assessment emphasizes the measurement of higher order thinking skills like analysis and application, evaluation and synthesis. The change is due to the fact that learning is a multidimensional process, not just a retention process. Modern achievement tests are thus developed in a standardized fashion ensuring fairness, validity, and reliability. The development has also been shaped by the field advances of psychometrics and educational psychology that place a greater focus on the scientific measurement of learning outcomes.

One of the important topics mentioned in the literature is the systematic construction of tests. The process of developing an achievement test consists of a number of steps that can be clearly specified and which guarantee the validity and reliability of the achievement test. Instructional goal identification is the first step. These objectives are the ones the test seeks to measure and offer a clear structure for writing test items. For the EGCAT Course, counselling theories, psychological principles, ethical practices in counselling and skills in student guidance and career counselling are some of the instructional objectives that may be covered.

Once objectives have been identified content specification is undertaken to be sure that the test adequately addresses all aspects of the subject domain. This step is similar to that of content validity, which involves the degree to which test items reflect the curriculum or content area. Test content is typically reviewed by experts in the field to guarantee content is thorough and well-balanced. This is important in teacher education since gaps and inaccuracies in the content area might result in misjudgments of competencies.

An additional essential part of test building is the making of a Table of Specifications (TOS). The TOS provides a blueprint that connects learning targets to test items, and guarantees a balanced representation of content areas and cognitive levels. It allows for balancing between topics and makes sure that both lower-order and higher-order thinking skills are measured. Literature in education has pointed out that a well-constructed TOS is important because it contributes to the validity and reliability of achievement tests through systematic item distribution.

Another important step in the test construction process is called item writing. The items must be clear, succinct, and in tune with the learning targets. Items that are not well

put together can result in confusion, misinterpretation and measurement errors. Achievement test questions are often in multiple choice, short answer or essay format. There are pros and cons to each format. Multiple-choice questions are good for covering a broad spectrum of content knowledge efficiently; essay questions better for assessing critical thinking and expressive skills. For EGCAT, a mix of item types could be suitable for a theoretical knowledge test and an applied counselling skills test.

In literature, pilot testing is strongly recommended as an indispensable development of a test. A pilot test is a test administered to a small subset of the target population that is used to find out if there are any problems with the items that are too difficult, too clear, or too long. Pilot testing provides feedback to enhance the quality of the items and to ensure that the final test will be effective. It also serves as initial information for item analysis.

The statistical technique that is used to assess the quality of test items is called item analysis. It contains difficulty index of the items, discrimination index etc., effectiveness of distractors etc. The difficulty index provides a measure of the difficulty of an item for the respondents, and the discrimination index provides a measure of the separation between high scorers and low scorers. Distractors in multiple choice questions help to ensure that only knowledgeable students choose the correct answer. Based on literature, it is found that item analysis is important in the psychometric of achievement tests so that it can improve the psychometric achievement test.

Validity is one of the most crucial in test construction. It is a measure of the degree of correspondence between a test and the content of a test. According to modern validity theory, validity is considered as a single concept in which there are three types of validity such as content validity, construct validity, and criterion related validity. Content validity is the degree to which test items cover the content. Construct validity determines whether the test actually measures the theoretical construct that it is intended to measure. The criterion-related validity of a test is its relation to outside criteria of performance. Validity is an important aspect in achievement tests, since invalid tests result in false conclusions regarding learner's competence.

Reliability is also an important criterion for achievement testing. It means that the scores within a test are similar in one test session to those in another. A good test will give consistent results when repeated under the same conditions. Test-Retest reliability, Internal consistency (Cronbach's alpha) and Split-half reliability are common methods of estimating reliability.

3. Research Methodology

The study aimed at designing and standardizing an achievement test for prospective teachers in Educational Guidance and Counselling. The purpose of the test

was to assess achievement on five broad areas: Basic Knowledge, Application of Knowledge, Communication Skills, Ethical Behaviour and Self-Improvement.

A quantitative research approach was adopted for the study. A sample of 60 prospective teachers of Education, studying in the University of Gujrat, was taken. Purposive sampling technique was used to select the participants.

There were 40 multiple-choice questions and one descriptive question in the preliminary paper of the test. The quality of the test items was evaluated by means of the statistical procedures. The revised and refined version of the achievement test was then given to the selected participants. The collected data was then analyzed for the achievement of norms and reliability and validity of the instrument.

4. Results and Discussion

Step I: Planning the Achievement Test:

The preparation of the test was the first step in test construction, and planning was required carefully in terms of what the test would measure, what the test would assess, and the target population. In this step the following aspects were taken into consideration:

1. Setting up the purpose and scope of the test
2. Defining Content and Learning Outcomes
3. Create Table of Specification (TOS)

1. Setting up the purpose and scope of the test

The intended users of this study were B.Ed. prospective teachers. The aim of the test was to measure their achievement in Educational Guidance and Counselling.

2. Defining Content and Learning Outcomes

The researchers concentrated on cognitive skills (understanding, application and analysis) when developing test items. Content areas were chosen based on the curriculum.

3. Create Table of Specification (TOS)

A Table of Specification was developed to ensure an even balance of content areas and cognitive levels. The TOS provided the distribution of questions by domains of knowledge, understanding and application.

Table no 1: Table of specification (TOS) of the Achievement Test (EGCAT)

Content Area	Weightage %	Knowledge (K)	Understanding (U)	Application (A)	Total Questions	Marks
Unit 1: Introduction to Guidance & Counselling	20%	4	4	2	10 MCQ	10
Unit 2: Guidance &	40%	4	4	2+1	10 MCQ + 20 1 Long	

Counselling in Pakistan						
Unit 3: Stakeholders (Parents, Community)	20%	3	5	2	10 MCQ	10
Unit 4: Planning & Action Plan	20%	3	4	3	10 MCQ	10
TOTAL	100%	14	17	10	41 Items	50

Step II: Construction of Test Items:

In the second phase, test items were developed in the form of multiple-choice questions. 40 MCQs were prepared initially. The items have been developed to test various cognitive functions and to provide a broad spectrum of test material.

Multiple-choice questions were chosen because they are efficient, objective and convenient to use for mass testing. The items were constructed following conventional guidelines for educational measurement.

4.1. Evaluation of Test Items

Qualitative Evaluation

The test items were examined for clarity of language, sentence structure, relevance of content and overall quality by ten teacher educators from Universities and Colleges. Some changes and enhancements were made based on their ideas.

A quantitative evaluation was carried out, and a pilot test was conducted.

60 prospective teachers from University of Gujrat Department of Education were given the test during its pilot. Each question had four options of which one was correct. One mark for every correct answer was given and no marks were given for incorrect or omitted answers. The results of the pilot study were statistically analysed by the item analysis procedures.

4.2. Item Analysis

1. Difficulty Value

The difficulty level is the percentage of students who answered an item correctly. The value ranges from 0.00 to 1.00. Those items that were between 0.20 and 0.80 were classified as acceptable and were used in the final test.

It was used the following formula:

$$P = \frac{R}{N}$$

Where:

P = Difficulty value

P = Number of questions answered correctly

N = Total number of respondents

Those items that were too easy or too difficult were eliminated from the test.

2. Discrimination Power

Discrimination power is the degree to which an item discriminates between good students and poor students. It was calculated using the following formula:

$$DP = \frac{RU - RL}{NU}$$

Where:

RU = Correct answers from the top group

RL = Correct answers from the bottom group

NU = Number of participants in the upper group (A3)

Those with a discrimination power less than 0.20 were rejected or modified.

Table no: 2 Discrimination power

Index of discrimination	Item evaluation
0.40 and up	Very good items
0.30-0.39	Reasonably good but possibly subject to improve
0.20-0.29	Marginal items, usually needing and being subject to improvement
Below 0.19	Poor items, to be rejected or improved by revision

Table no: 3 Item Analysis: Item Difficulty Value and Item Discrimination Power

Item No.	Score	Difficulty Value (P)	RU	RL	Difference (RU-RL)	Discrimination Power (DP)	Item Decision
1	17	0.53	16	1	15	0.94	Retain
2	16	0.5	12	4	8	0.5	Retain
3	22	0.69	14	8	6	0.38	Revise
4	20	0.62	11	9	2	0.12	Reject
5	19	0.59	14	5	9	0.56	Retain
6	17	0.53	11	6	5	0.31	Revise
7	21	0.66	14	7	7	0.44	Retain
8	20	0.62	13	7	6	0.38	Revise
9	20	0.62	13	7	6	0.38	Revise
10	18	0.56	10	8	2	0.12	Reject
11	20	0.62	14	6	8	0.5	Retain
12	20	0.62	14	6	8	0.5	Retain
13	18	0.56	12	6	6	0.38	Revise
14	18	0.56	10	8	2	0.12	Reject
15	24	0.75	15	9	6	0.38	Revise
16	23	0.72	15	8	7	0.44	Retain
17	21	0.66	13	8	5	0.31	Revise
18	24	0.75	16	8	8	0.5	Retain
19	20	0.62	15	5	10	0.62	Retain

20	16	0.5	11	5	6	0.38	Revise
21	18	0.56	13	5	8	0.5	Retain
22	20	0.62	15	5	10	0.62	Retain
23	18	0.56	11	7	4	0.25	Revise
24	20	0.62	15	5	10	0.62	Retain
25	21	0.66	15	6	9	0.56	Retain
26	17	0.53	10	7	3	0.19	Reject
27	18	0.56	12	6	6	0.38	Revise
28	24	0.75	16	8	8	0.5	Retain
29	24	0.75	16	8	8	0.5	Retain
30	19	0.59	13	6	7	0.44	Retain
31	22	0.69	16	6	10	0.62	Retain
32	17	0.53	11	6	5	0.31	Revise
33	22	0.69	16	6	10	0.62	Retain
34	22	0.69	16	6	10	0.62	Retain
35	17	0.53	11	6	5	0.31	Revise
36	24	0.75	16	8	8	0.5	Retain
37	18	0.56	12	6	6	0.38	Revise
38	24	0.75	16	8	8	0.5	Retain
39	20	0.62	16	4	12	0.75	Retain
40	20	0.62	14	6	8	0.5	Retain

Step IV: Reliability and Validity

1. Reliability

Reliability is the consistency and correctness of test scores. The split-half reliability method was used to determine the internal consistency of the test. Reliability coefficient was calculated using the Spearman–Brown Prophecy Formula.

The final reliability coefficient was 0.89 which showed a high level of reliability of the test.

2. Validity

Ten subject experts have been used to determine the face and content validity. They checked to see if the test was adequate in content areas and learning objectives of Educational Guidance and Counselling.

The experts verified the test items were successful in sampling the desired concepts and cognitive domains.

3. The final draft of the Achievement Test

Following the item analysis, items were selected that met the criterion saturation level (difficulty value) and discrimination power, with 41 items being incorporated into the final test.

Marking of answers: One mark was awarded for each correct answer, zero for incorrect and no attempt for un attempted answers. There was no negative marking.

Table no 4: Final Draft of Achievement Test

Units	Weightage	Knowledge (K)	Understanding (U)	Application (A)	Question	Marks
Unit 1	20%	4	4	2	10	10
Unit 2	40%	4	4	2+1	11	20
Unit 3	20%	3	5	2	10	10
Unit 4	20%	3	4	3	10	10
Total	100%	14	17	10	41	50

Step V: Establishment of Norms:

Students' scores were interpreted effectively with the development of norms. For this, a representative sample of the B.Ed. students was taken.

Three kinds of norms were set up:

Z-score norms

Percentile rank norms

Stanine scores

These norms were used to categorize students' performances and to provide comparisons between students' achievement and group achievement.

5. Conclusion

The present study described the whole process of development and standardization of an Achievement Test for Educational Guidance and Counselling. The test has been designed to evaluate the potential teacher's mastery of important concepts in guidance and counselling.

The instrument was built following the standard procedures in instrument development such as planning, item writing, pilot test, item analysis, reliability test, validity test and norm development. The test showed a high reliability coefficient of 0.89, meaning that there was great consistency in the test.

This final achievement test can be used by teacher educators, researchers and educational institutions in order to assess the understanding of Educational Guidance and Counselling of prospective teachers. Moreover, the research offers a helpful model for the creation of standardized achievement assessments for other content areas and disciplines in education.

References

- Anastasi, A. & Urbina, S. (2024). Psychological Testing (Eleventh Impression). Person India Education Services Private Limited.
- Arslan, K., & Ari, A. G. (2024). Development of Ecology Achievement Test for Secondary School Students. Shanlax International Journal of Education, 12(1), 166–179.

- Bala, S. M., Bunkure, I. Y. & Zayum, S. D. (2024). Development and Validation of Electricity and Magnetism Achievement Test (EMAT). *Vunoklang Multidisciplinary*
- Bennett, R. E., Rock, D. A. & Wang, M. (1991). Equivalence of free-response and
- Best, J. W., Kahn, J. V. & Jha, A. K. (2016). *Research in Education* (Tenth Edition).
- Brown, F. G. (1983). *Principles of Educational and Psychological Testing* (Third Edition), Holt, Rinehart & Winston.
- Chaudhary, S. & Tyagi, S. K. (2017). Construction and Standardization of Achievement
- Deupa, M.S. & Deupa, J. (2023). Construction of attitude scale: Attitude of school teachers towards teaching profession. *Sudurpaschim Spectrum*, 1(1), 1-21.
DOI: <https://doi.org/10.3126/sudurpaschim.v1i1.63384>
- Ebel, R. L. & Frisbie, D. A. (1991). *Essentials of Educational Measurement* (Fifth Edition). Prentice-Hall of India Private Limited.
- Freeman, F. A. (2018). *Theory and Practice of Psychological Testing* (Third Edition). Oxford and IBH Publishers.
- Gangwar, S. & Singh, S. P. (2019). Vigyaan Upalabdhi Pareekshan: Nirmaan Aur Maanakeekaran. *Bhaarateey Aadhunik Shiksha*, (39)4, 46-55.
- Gupta, S. P. & Gupta, A. (2018). *Statistical Methods in Behavioural Sciences*. Sharda Pustak Bhavan.
- Journal of Science and Technology Education*, 12(2), 270–283.
- Lalhruaitluangi & Kaur, N. (2024). Construction of an Achievement Test for Students of Collegiate Level. *Mizoram Educational Journal*, X(3), 54–62.
- Lyman, H. B. (1998). *Test scores and what they mean* (6th Ed.). Allyn & Bacon.
- Miller, L. A., Lovler, R. L. & McIntire, S. A. (2015). *Psychological Testing: A Practical Approach* (Fourth Edition). SAGE Publications India Private Limited.
- Multiple choice items. *Journal of Educational Measurement*, 28(1), 77-92.
- Patel, R. K. & Singh, S. P. (2018). Sanskrit Uplabdhi Parikshan ka Nirman tatha Mankikaran. *Pariprekshay*, 25(2), 89-110.
Person India Education Services Private Limited.
- San Martín-Rodríguez, L., Beaulieu, M. D., D'Amour, D. & Ferrada-Videla, M. (2005).
- Singh, A. K. (2019). *Tests, Measurements and Research Methods in Behavioural Sciences*. Bharati Bhawan Publishers & Distributors.
- Singh, G. & Sharma, A. (2024). Construction and Standardization of achievement Test in Science. *TIJER- International Research Journal*,
- Test in Educational Psychology. Educational Quest: An Int. J. of Education and Applied Social Science*, 8(3), 817-823.
- The determinants of successful collaboration: a review of theoretical and empirical studies. *Journal of Interprofessional Care*, 19(sup1), 132-147.