

Proposing A Protocol Schemata for Developing Web-Based Question Item Bank for Summative Assessment

Dr. Ghulam Dastgeer

Principal,
Federal Directorate of Education
Islamabad

gdastgeer@hotmail.com

Dr. Zarina Akhtar

Assistant Professor,
Department of Teacher Education,
International Islamic University,
Islamabad

zarina.akhtar@iiui.edu.pk

Received: 31-May-2024

Accepted: 10-June-2024

Abstract

Introduction: This study aims to propose a protocol schemata for developing a web-based question item bank (QIB). It can be developed by educational experts in collaboration with assessment experts. These protocol schemata propose all the security and safety features required for developing QIB.

Methodology: The prevailing procedures for paper setting at educational boards in Pakistan and their needs to establish QIB were studied. Available web-based QIBs and their features were also studied. The knowledge gained was utilized for research and development for proposing some protocol schemata for developing a secure and comprehensive web-based QIB.

Results/Findings: The studies of prevailing procedures for paper setting and features of available QIB software led the researchers to propose a protocol schemata essential for establishing QIB for examining bodies. They

need to ensure security measures and propriety rights to save their QIBs from hacking.

Future Direction: *The proposed schemata can be helpful in establishing a web-based QIB software having multiple types of question items developed as per Bloom's Taxonomy cognitive and difficulty levels. It may be multilingual, and mathematical equations and scientific notation supported. It can ensure security from hacking when kept offline after development.*

Keywords: *Assessment; Bloom's taxonomy; Protocol; Question item bank (QIB); Schemata; Standardized testing; Web-based.*

Introduction

Examining bodies across the globe and especially in Pakistan face difficulties in assessing learners while devising question papers or tests consisting of varied items. These assessments should present a true picture of learners' and candidates' basic knowledge as well as higher-order thinking skills for promoting them to the next higher grades. The question items should be selected from a wide range of content and objectives to have some balanced assessment tools. Many of these examining bodies, such as education boards, have realized the need to establish question item banks (QIBs) for devising credible tests (IBCC, 2020-21), so that some reliable and valid tests could be generated for the assessment of candidates. Recently, the Federal Board of Intermediate & Secondary Education Islamabad, Agha Khan Board Karachi, and some other boards in Sindh and KPK, and the Punjab Examination Commission (PEC) have adopted to use some software such as REDMarker (2019) for the development of a question item bank and generation of question papers. Yet there is a lot needed in the field for having a complete solution to have some software with comprehensive interface construct options, applications, and security features.

Background and Context of the Study

Poor quality of education in Pakistan is always questioned by researchers, educationists, recruitment agencies, and the public. The factors might be many, such as lesser resource allocation; Pakistan spends 1.7 percent of its Gross Domestic Product (GDP) on education (Economic Adviser's Wing, 2023), access to education, lack of facilities, less motivated teachers and learners, uncongenial learning environment, bad instructional techniques and strategies, poor supervision and monitoring, and faulty assessment and evaluation mechanism. IBCC (2020-21) in its meeting pointed out that question papers are not balanced and result in inflated marks. Portions of understanding and application as per Blooms Taxonomy are often missing. A large number of

candidates, even after qualifying these exams with good grades, lack required levels of competency. This shows drawbacks and deficiencies in the prevailing examination system.

The study was delimited to analyze the prevailing assessment systems at the education boards in Pakistan. The researchers studied and reviewed relevant literature, and conventional as well as web-based systems to analyze the features and pros and cons of these systems. One conventional system prevailing at BISE Rawalpindi and one QIB system at FBISE Islamabad were studied to see the process of setting question papers for summative assessment. The relevant literature review gave the researchers an insight into their features and functioning and deficiencies. They have proposed a protocol schemata for developing web-based QIB.

Objective of the Study

The objective of the study was to propose a protocol schemata for the development of software for web-based development of question item bank that could auto-generate question papers for summative assessment.

Significance of the Study

This study has proposed a protocol schemata for developing web-based QIB software having comprehensive interface and construct options. The proposed software can enable a large number of question item developers, item reviewers, and item selectors to work simultaneously. They can work for developing and storing a large number of question items in various subjects, thus enabling the assessors to generate valid and reliable tests enhancing the quality of summative assessment. This would lead to better quality of education in the future.

Literature Review

Assessment is a systematic approach for knowing learners' achievements and for decisions for their promotion. Erwin (1994) defines assessment as a process of defining, selecting, designing, collecting, analyzing, interpreting, and using information to increase students' learning and development. Assessment occurs whenever a person performs some interaction with the other for getting and interpreting his competencies (Rowntree, 1977). There are two distinct interpretations of assessment as described by Biswas and Pradhan (2002):

- I. Interpretation of assessment in terms of the routine tasks that students undertake in order to receive feedback on their learning and a mark or grade signifying their achievement.

II. Assessment as applied to processes at the institutional level, for example, programs evaluation.

Assessment plays an important role in the teaching and learning process (Danili & Reid, 2005). Assessments cast important effects on individuals' life prospects not only immediately after graduation but also in later years of their life. Boud (1995) is of the view that the effects of bad assessment are far more dangerous than any other aspect of teaching. So assessments should be based on balanced tests taken from a wide range of curriculum and its objectives, valid and reliable, and covering all levels of Bloom's taxonomy. And the marking should also be consistent and having multiple assessors that make assessment highly valid and reliable (Gielen & Wever, 2015). Generally, assessment tests include objective items, subjective items short/restricted response questions and long questions or descriptive response questions. Presentations, students' profiles, and interviews can also be used for assessment (Bearman & Ajjawi, 2019). Kirkpatrick, Mistry, and Leshikar (2020) also found that presentations can assess the interpersonal skills of a person.

Teachers develop their tests locally to assess the achievement of their students through formative assessment. Summative assessment, on the other hand, requires that tests should be meeting maximum levels of validity and reliability. The tests that ensure maximum levels of validity and reliability are called standardized tests.

Standardized Tests

Standardized tests contain test items developed by teams of experts ensuring that these are valid in all forms. These test items are administered to a large number of examinees for a number of times to see their reliability. These tests can also be used to evaluate programs and projects. Standardized testing evolved through the ages. Han dynasty of 650 AD in China employed some kind of standardized tests for selecting civil officers. It laid the foundation to assess an individual based on some standardized criteria (GGI Insight, 2024). In the First World War, the US Army used Alpha Army Tests and Beta Army Tests for selecting officers. Later on, it was realized that standardized tests could also be used for the bureaucracy recruitment process. Gradually, these tests were adopted for educational assessments on a large scale. In the beginning, these tests were based on assessing the factual knowledge of students in various subjects, especially in English and Mathematics. By the mid-20th century, the educational researchers diverted their focus towards assessing higher-order thinking skills.

Evolution of Standardized Testing by Various Personalities

Binet and Simons (1916) pioneered the standardized intelligence testing in the early 20th century. The College Entrance Examination Board commenced in the 1920s in the USA, later on renamed as the Scholastic Aptitude Test, or SAT. Bloom (1956) contributed through his “Bloom’s Taxonomy of Educational Objectives (1956)”. His taxonomy is a framework for assessing different and complex levels of cognition. Following this, Thorndike et al. (1964) developed multiple-choice tests. Coleman (1966) pointed out disparities in educational outcomes based on economic status and stressed the need to address these disparities through standardized testing by raising awareness for educational equity.

Standardized Testing: Its Process and Procedures

Standardized testing is a cumbersome process; it needs careful planning and execution. The development of test items requires subject as well as assessment expertise. They collaborate to develop test items that should have all types of validity; validity and content validity. The test should assess truly the intended learning outcomes and could provide tangible feedback. The test items are then administered on a very large number of subjects time and again to ensure reliability and added as part of a standardized test. Once the test is finalized, its administration is done through careful planning and smart organization, i.e., selecting the appropriate testing environment, ensuring security and secrecy, and giving clear instructions to the candidates taking these tests.

Pros and Cons of Standardized Testing

There is a difference of opinion among various stakeholders in education. One of the protagonists, Smith (2018), argues that standardized testing ensures objective and unbiased measurement of students’ performance. It allows identifying areas for improvement through comparisons. Grodsky, Warren, and Felts (2008) are of the opinion that standardized tests do not create social stratification. The antagonists such as Meier and Knoester (2017), on the other hand, apprehend the limitation and undesirable consequences of standardized testing. They point out the narrowness of high-stakes testing that may ignore holistic education. Moreover, students from different social statuses may feel discriminated against while taking these tests because of the input of fewer resources and exposure in education. These disparities may affect their performance in the tests.

Alternative Assessment Models and Their Effectiveness

Keeping in view the criticism against standardized tests for its limitations, researchers ventured to explore new avenues for balanced and better assessments. They proposed alternative models for assessment to have a comprehensive view of students' abilities. These models include portfolios, projects, and performance-based tasks, where students could show their knowledge and skills in real-life situations (Wiggins, 1998). Alternative models of assessment are beneficial to see deep insights into students' performance and promote higher-order thinking skills and creativity. So these are better in comparison to standardized tests for assessing a broader range of students' abilities. In recent years, alternative models for assessment are gaining popularity.

Computer-Based Assessment

Computer-based standardized testing has gained popularity by replacing conventional paper-pencil tests in recent times with the inventions of computers and information technology. It enables adaptation where various difficulty level question items can be selected based on the student's performance to precisely measure the individual's ability. It also provides automated scoring, thus reducing subjectivity. Standardized testing has proceeded further through the use of the latest technologies such as virtual reality (VR), augmented reality (AR), and artificial intelligence (AI) (Gandolfi, 2018). AI-driven systems are powerful tools for providing immediate feedback, enabling students to track their individual progress and better learning. These have revolutionized standardized testing through the formation of question item banks (QIBs). Quellmalz and Pellegrino (2009) stress the need and importance of computer-based tests used for administering these tests either in standalone offline or through an internet configuration on the World Wide Web (www).

Paper Setting Process in Pakistan

Education boards in Pakistan adopt the traditional way of paper setting. Some training sessions are arranged for paper setters by the boards regarding assessment, its process, and procedures. They are trained on how to prepare question papers in their relevant subjects. BISE calls paper setters in various subjects in groups and conducts refresher courses for them for two days. Afterwards, they prepare question papers as per instructions given by the board. All the education boards such as the BISE Rawalpindi calendar (1997) - part II chapter 4 section III clause 22-38- contains rules related to the appointment of paper setters under the title appointment of examiners. Some of these rules require that:

- I. None of the questions should be set by him (the paper setter) for any other examining body.
- II. He should either write or type the paper by himself for secrecy.
- III. No copy of the paper should be kept for him; even carbon paper if used should be destroyed immediately.
- IV. He should check and ensure that the question paper is strictly as per the prescribed syllabus.
- V. The question paper should be error, obscene, and vulgar-free in content and language.
- VI. No relative of him should be a candidate for the exams for which he is setting the paper.
- VII. He should not be teaching the subject to the classes for the previous two years for which he is setting the question paper.

The paper setters are required to hand over the question papers (at least three different copies) in sealed envelopes to the authorized officer confidential press. The process of paper setting has not shown positive results in terms of the quality of question papers. Question papers are set from a limited content range and do not cover all the cognitive levels described in Bloom's taxonomy. These often met with criticism for low quality of assessment.

The Inter Board Committee of Chairmen (IBCC) in 167th, 168th, and 169th meetings (2020-21) deliberated for the establishment of a question item bank. The reference was made to the instructions of the Ministry of Federal Education and Professional Training vide letter No. F. 2(11)/2020-Coord, dated May 18, 2020, that a meeting held in the Prime Minister Secretariat on the subject of Review Performance Agreement of Financial Year 2019-20 has decided that IBCC may be directed to formulate a Question Bank for SSC and HSSC levels in order to standardize question papers for examinations for all the Boards of the country. It was argued that question papers are not balanced. Portions of understanding and application as per Bloom's Taxonomy are often missing. This is one of the reasons for marks inflation. It is pointed out that whenever there is a reference about Question Item Bank it is meant for curriculum-based examinations whereas

boards in Pakistan conduct examinations on the basis of content available in syllabus books. There is a need for assessing the students through SLOs-based examinations.

A resolution was also adopted in the 169th meeting of IBCC (2021) for constituting and to work out and speed up the process of establishing a Question item bank in coordination with the IBCC Secretariat and progress in this regard be shared at the next Forum Meeting. The IBCC Secretariat would look after the entire process.

Question Item Bank

In recent years, the assessment process is done through the use of a question item bank (QIB) by developing and storing large numbers of items in the form of a bank. Tests are devised by selecting items from this vast variety of stored items from the QIB. A QIB is a planned library of test items pooled through cooperative efforts under the protection and support of an institution for the use of evaluators, academics, and students in partial fulfillment of the requirements of the teaching-learning process (Biswas & Pradhan, 2002). A QIB is a very large collection (probably thousands) of question items in all forms; objective/MCQs, SRQs, and DRQs stored under all levels of Bloom's taxonomy cognitive domain, with varying difficulty levels and are properly indexed. These items should be secure and can be easily accessible for use in question paper form for assessment. Ghio et al. (2021) recommended CAT item banks to rapidly generate indicators of how students acquire knowledge so that strategies could be devised to enhance their academic performance. QIBs are useful for assessing large numbers of students through on-campus or online assessment. Akuadi (2010) opined to have QIB and employed through computer-based assessment programs in developing countries to cope with the practices in vogue in the developed world.

Establishing a Question Item Bank

The establishment of a question item bank starts with proper and careful planning. The objectives should be clearly defined to enhance the measurement accuracy and value of assessment for evaluation purposes. The storage of QIB should be safe and secure from trespassing and hacking. Items should be prepared with due care as per Bloom's taxonomy of cognitive domains and difficulty level (Prasad & Xavier, 2006). Items should cover the whole curriculum and learning outcomes. All types of items - objective/MCQs, SRQs, and DRQs - should be included in the bank. Items once written should be reviewed by other subject specialists. They should ensure that the

written item is authentic in content for a specific grade level. The expert in assessment should ensure that the given objective or outcome can be tested by the item. The second round of review of items should be done by the assessment expert to ensure that the format of the item is appropriate for testing a specific learning outcome. After essential improvements, these may be stored in the form of a QIB. The question paper meant for assessment can be designed based on: a) learning outcomes, b) types of questions, c) content areas in the given course, and d) difficulty level (Akuadi, 2010).

Pros and Cons of Question Item Bank

Question item banks are quite beneficial for assessment as they ensure more validity and reliability. Some of the advantages of QIB are given below:

- I. QIB is developed with the collaboration of many participants benefiting from one another's experiences and expertise, ensuring quality.
- II. Locally developed items make authors feel proud of their ownership. But even if items are taken from already developed sources, these save time, energy, and money.
- III. Items are categorized as per Bloom's taxonomy and difficulty levels ensuring a balanced question paper.
- IV. QIB being a large pool of items ensures security features making it difficult to remember answers to thousands of items, thus reducing the chances of cheating in the examinations.
- V. Use of question items for online testing can save a lot of time and resources for administering the test, calculating the statistics for improvement of tests by retaining stronger and reliable items and deleting the weaker and less reliable ones.

Despite these and many other advantages, QIBs have some issues too. Poor and developing countries might face economic difficulties at the initial stages. Hiscox (1983) was of the view that it was not easy to implement QIBs; securing a sound and useful collection of items, maintaining it by experts, and using the items appropriately and effectively. Some other difficulties are also pointed out such as estimating norms (when each student has a different test), and the measurement error emanating from violations of assumptions (when banks have items calibrated using item

response theory). Another major threat to the QIB is its safety and security when kept online, because of hacking threats.

Developing Question Item Bank

A question item bank can be developed either manually; developing items on paper, reviewing these by experts for validation, administering on a large sample multiple times to see their reliability indexing, and recording them on paper in the form of a library. The other and more effective way is to use system software for the whole process. Where a number of participants can collaborate for developing, validation, and storing items in soft forms. These storages can be saved offline or online for generating tests for future use. A number of software applications and repositories are available online for generating tests on a payment basis. The experts have suggested various steps and strategies for the development of question item banks. The Table 1 below shows important considerations while developing QIB:

Items Must:	Items May:	Items May Not:	Items Must Not:
Protected and secure	Be used to supplement faculty created exams (stored in the exams software)	Be edited except for titles and tagging categories (may be changed to align with programs’ tagging nomenclature)	Be edited except for:
Faculty be participating	Used to supplement faculty created formats (stored in exams software)		i. An original item; only item titles and tagging categories be changed.
No pics or transcriptions	Used to create muddiest points quizzes (stored in exams software)		ii. To make any other edits, make a copy first.
Be within the exams software/platform			iii. Original items and their copies should remain in their respective folders

Table 1: Considerations for the development of QIB. Adapted from: Best practices for Building your own Question Bank for Clinical Judgement exams. Retrieved from Support. examsoft .com /hc

Available QIB Software and Their Features

Certain software is available that can be used for assessment. Most of this software has their own developed question item banks for different levels and grades. One can have access to these through the internet either on a one-time payment or paying for each and every connectivity. Some of these and their salient features are given below:

ExamSoft (by Turnitin): Developed by the famous Turnitin organization (2024), it is offered for government assessment. This software has certain features such as:

- I. Multiple types of assessments available that can be chosen as per the needs of the organization,
- II. Graded exams options are available that can be customized,
- III. Tests can be conducted through this software 24/7 anywhere without compromising security,
- IV. Performance analysis can be done by collecting data for individuals or groups,
- V. It is compatible with iPad and other applications,
- VI. Can be used offline with a standalone computer system,
- VII. Items are randomized to prevent memorization,
- VIII. Screen shots are disabled to ensure more security.

eSystem Question Bank: This software has been developed by Dory (n.d) from the Faculty of Medicine, McGill University. The main features of this software are its flexibility in developing question banks with multiple groups, working 24/7 from anywhere, matching SLOs with the assessment naming it as ‘blueprinting’, statistical analysis of items and performance of examinees, and it can be stored in the cloud or kept individually.

In Pakistan, the Federal Board of Intermediate and Secondary Education Islamabad and some other boards in KPK and Sindh provinces are using software provided by **REDMarker systems**; a private firm in the field. These boards are using software for paper generation and on-screen marking of scripts. REDMarker (2019) provides various products such as rTestGen and rMarker for paper generation and on-screen marking respectively. These products were developed during the COVID-19 period. rTestGen is an item bank software tool for generating question papers for

students' assessment. Its key features include developing items in all forms; i.e., MCQs, short response questions, and descriptive response questions in a multilingual format and according to Bloom's taxonomy cognitive levels. It has options for item tagging, similarity check, and multiple paper generation. Whereas rMarker is a software used for on-screen marking of papers by scanning the scripts and sending them item-wise to various examiners for quick assessment and compiling of results. In the beginning, the users feel some difficulties but later on their pace and performance are increased.

Many other question item bank software are available online on payment. The following software is considered among the top ten popular software for question item banks:

- Online ExamMaker
- Quizlet
- ClassMarker
- ExamView
- TestGenius
- Testbase
- Respondus
- QuestionWriter
- Quia
- TeacherMade

The features of these software are described briefly below:

- **Online ExamMaker:** This provides options for users to import items in bulk using Microsoft Word or Excel, and duplicate items can be deleted easily. It supports mathematical equations and can add ten types of question items. But it has no option for multilingual script and advanced features cost more.
- **Quizlet:** It is a popular study app that provides options for studying through flashcards learning activities. Teachers can use them for keeping students' learning history. It offers multiple question types and an easy-to-use format. It has limited analytical tools and offers more learning options than testing.
- **ClassMarker:** It offers fast analytical and reporting tools. It is easy to use for managing tests and analyzing students' performance. Tests can be customized as per local requirements of the class. But it has limited options for integrating with other learning management systems. Moreover, its contents do not meet defined standards.

- **ExamView:** This software offers to write, edit and save assignments to be assessed and can get immediate responses. It has flexible assessment and secure delivery options. But the learners may feel difficulty and their performance might slow down in the beginning.
- **TestGenius:** It has the options for customized assessment with diverse question types, analysis of performance and efficient reporting. But it has limited options for integration with other learning management systems.
- **Testbase:** It is meant for online test generation and assessment for quick creation, customization, delivery of tests and quizzes, analysis of results, and fast reporting. It has quite a large number of items in the library, but limited when compared with other competitors in the market. It also has limited custom branding.
- **Respondus:** It is widely used software by educators and institutions for conducting tests and assessments. The users can import items from various sources on this platform. There are options for compatibility lockdown for enhanced security measures. But it has limited advanced reporting capabilities.
- **QuestionWriter:** It is an online platform for educators to create, format and print tests for quick delivery. It offers collaborative test planning, execution, and powerful reporting features. Having many features, this might be costly for small groups and institutions.
- **Quia:** This software offers online games, surveys, quizzes, and tools for teachers to engage learners in creating their own assessments. Being a comprehensive module, it includes a question bank pool too. But it has limited analytical and reporting tools. Its options for customizations of assessment are also limited.
- **TeacherMade:** It is an online assessment creation platform where teachers can create, assign, and score tests in aligned standards. They can also report easily. But it has a limited number of items as compared to other competitors in the field.

All these software have some common features and certain unique features to each of them.

The table 2 below shows a comparison of features of the top ten popular software:

Name of Software	Question Item Type	Option for Adding Multimedia Audio/Video/Pic	Reviews
Online ExamMaker	10	Limited	G2. 4.5/5 (http://www.g2.com/products/onlineexammaker/reviews)
Quizlet	5	No	G2. 4.5/5 (http://www.g2.com/products/Quizlet/reviews)
ClassMarker	6	Yes	Capeterra 4.7/5 (http://www.capeterra.com/p/2021)
ExamView	6	Limited	G2. 4.2/5 (http://www.g2.com/products/examview)
TestGenius	4	Limited	Capeterra.com. 5.0/5 (http://www.capeterra.com/p/2154)
Testbase	9	Yes	
Respondus	5	No	G2. 2.4/5 (http://www.g2.com/promonitor/reviews)
QuestionWriter	10	Yes	
Quia	10	Limited	Common Sense, 3.0/ (http://www.commonsense.org/eduweb)
TeacherMade	13	Yes	

Table 2: Features of top ten web-based QIBs Adapted from: <https://onlineexammaker.com/kb/top-10-question-bank-software-for-teachers-students/>, 2023/en-us/articles/1267959881 -Best-practices-for-B, retrieved May 23, 2024, 21:27 PST.

Method

This research can be categorized as research and development. The relevant literature was studied thoroughly and reviewed accordingly. Some of the literature was studied and reviewed under the following subheadings: Assessment, Standardized tests, its history, benefits and drawbacks of using standardized tests, their evolution, process and procedures for developing standardized tests, computer-based assessment, paper setting process in Pakistan, question item bank, its establishment and development, pros and cons of question item bank. Available question bank software and their features were also discussed.

The literature review provided necessary information and insight to the researchers. They thought on the plan on how to define certain protocol schemata for the development of question item bank software. The main concern with the prevailing and available software is the access to their repositories on payment. Anyone can have access to these banks on payment. So it is not advisable for examining agencies in Pakistan especially, the boards, universities and public service

commissions to use these repositories for generating question papers for their tests and examinations. They need such system software banks where items should be safe from hacking. They should have the proprietary rights for these question item banks. Moreover many other security and safety features are provided in this proposed protocol schemata. The details of proposed protocol schemata for developing web-based question item bank software for summative assessment are given in the fourth coming paragraphs.

Research and Development

The proposed software for developing a web-based question item bank ensures all the safety, security, and confidential measures and protocols. It would provide complete solutions for the examining agencies to use it for the development of items on the web by engaging a large number of item writers, reviewers, and selectors simultaneously. The commencement of the process is detailed in the following paragraphs. Figure 1 below provides details regarding participants who can contribute to developing a web-based question item bank:



Fig. 1: Participants in developing QIB

The software portal would need four types of participants for the development and storage of question items: I. Item Writer, II. Item Reviewer (referred to as reviewer), III. Item Selector (referred to as selector), and IV. Software Controller (referred to as administrator). Their qualifications and roles are described below:

I. Item Writer: The item writer should be qualified for teaching specific subject content and be an expert in item development at a certain level/grade in the specific subject. They should be able to develop items as per Bloom's Taxonomy cognitive levels and various difficulty levels.

II. Reviewer: The reviewer should have knowledge and teaching experience in the specific subject and in assessment.

III. Selector: The selector should be an expert in assessment and item development as per cognitive and difficulty levels and have some teaching experience in the specific subject.

IV. Administrator: The main role of the administrator is to ensure security, safety, and confidentiality measures. He should be an IT expert who would handle the software and manage traffic among the other three participants. He would troubleshoot any problems related to software functioning. He would monitor and control the traffic but would be unable to see the content of items. In case more pendency lies on the part of any item writer, reviewer, or selector, the administrator could divert the task toward other participants with similar roles.

The above-mentioned participants would be anonymous to one another while working on the QIB software portal.

Training Module: The process of software functioning for question item development would start with the training of the above-mentioned participants. The training modules would also be available on the portal on a large scale in the main areas of: I. Subject-specific item development, II. Item development as per Bloom's taxonomy cognitive, and difficulty levels, III. Assessment, and IV. Software application usage. Training modules would utilize all the tools and resources; i.e., PPTs, documents, PDFs, tutorials, and would follow a pre-test post-test design for assessment to see the level of achievement of trainees. On completion of successful training through auto and self-assessment, the participants would get certified. They would submit that they have fairly been trained in software handling and any one of the above-mentioned areas for item development.

Certification would follow mandatory registration and providing an affidavit that they would ensure the secrecy and safety of the items developed and all sorts of leakage should be plugged in. The registration process would need three main types of information in the profile of participants: personal information, professional information, and bank account information for payment purposes. The figure 2 below explains the profile and registration process:

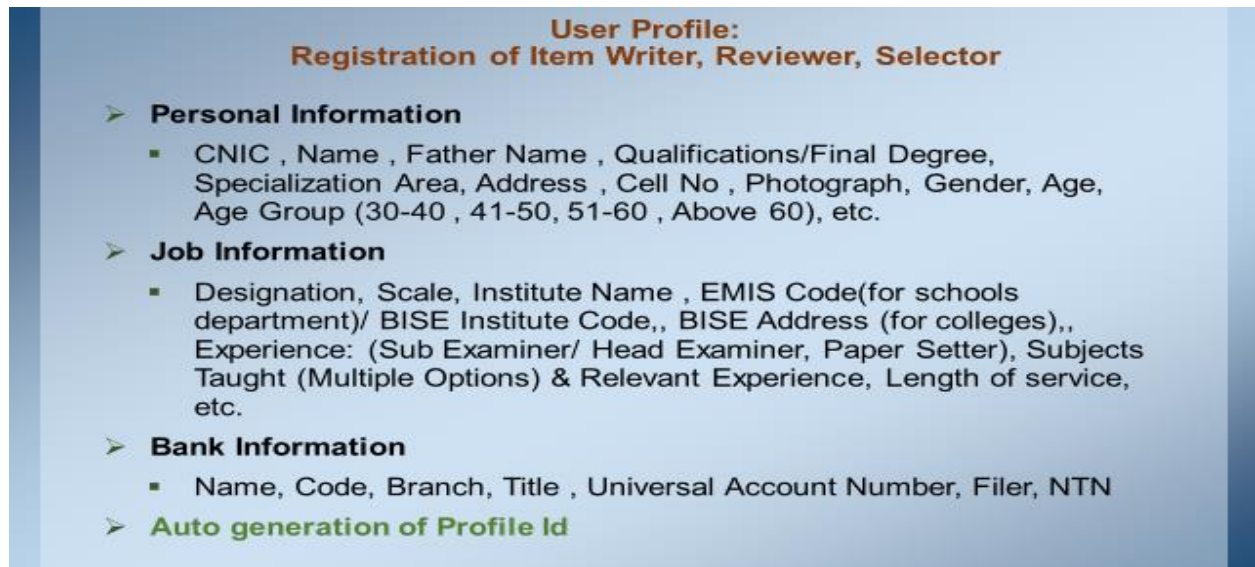


Fig 2: User profile of participants in developing QIB

The registration process of the item writer, reviewer, and selector contains three types of information in their profile: I. Personal information, II. Job-related information, and III. Bank account information. Profile of item writer, reviewer, and selector contains the details as given below:

I. Personal Information: They would provide details regarding their name: father's name; CNIC number, qualifications with a special area of studies; gender; age; address; contact number, and age group.

II. Job-Related Information: This contains, designation; pay scale; relevant experience posting place; EMIS code of the institution; institution code allotted by the examining body (board); experience as item writer/ paper setter/ examiner; subject taught.

III. Bank Account Information: This is also required as account title; account number/ universal account number; bank and branch; NTN; filer/ non-filer. Upon filling in all the information one can get registered oneself for a specific role. An auto profile ID would be generated on completion and submission of the profile. The interface for the profile would be as given in figure 3 below:

Participant's Profile

A Web Page

<https://bisesececy.com/personalinfo>

Teacher Profile

Personal Info

Related Attachments **+**

Crno no: Name: Father Name:

Mobile: Address:

City: Gender: ☐ Male ☐ Female Age:

Subject Taught: English Mathematics **+**

Job Info

Emis Code: Scale: Designation:

Institute (code - name): -

Address:

Sub-Examiner Experience: No Of Year Experience:

Bank Info

Bank: Branch Code: Branch:

U. Account No.: Filer: ☐ Yes ☐ No NTN Number:

Login Info

Login: Password:

SAVE

Fig.3 Participant's Profile; Dashboard view

Similarly, figure 4 below shows the interface for login details accordingly:

Login Form

A Web Page

<https://bisesececy.com>

username:

password:

Login

Fig 4: Login form portal view

Software features for developing a question item bank: The question item bank (QIB) software has certain features. The features of QIB software are depicted in figure 5 below:

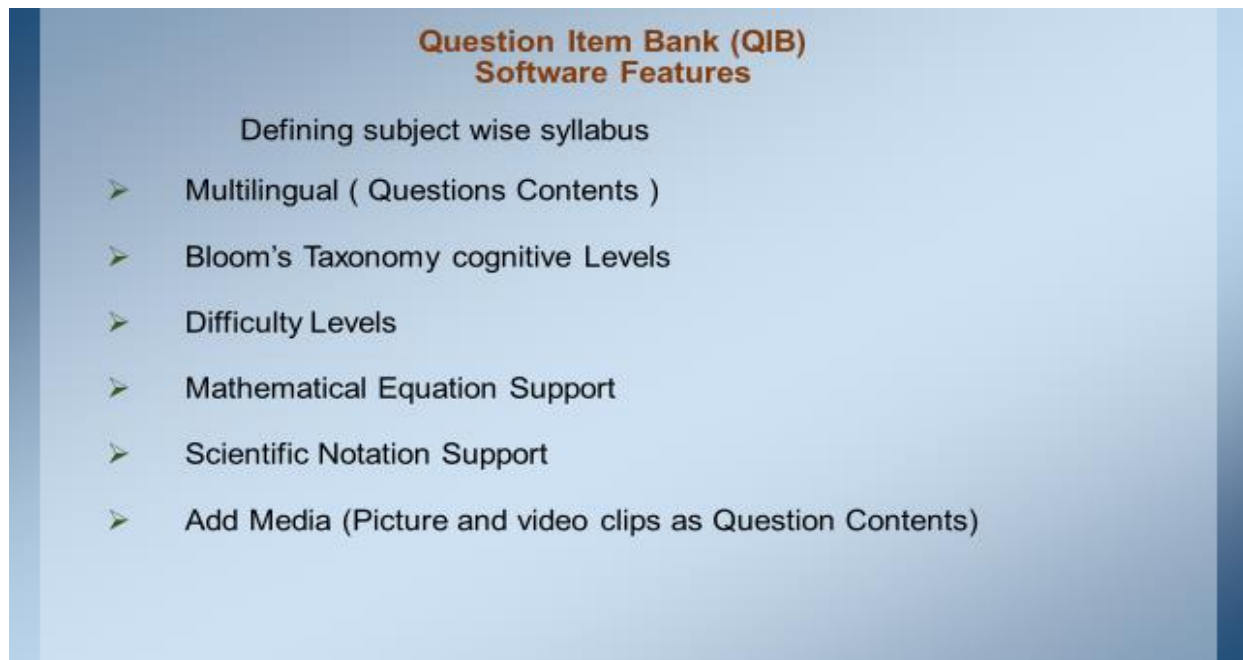


Fig.5: QIB Software Features

- I. **Defining Subject-wise Syllabus:** The curriculum and syllabus contents are digitally added to the software for use by the item writers. Whenever they develop an item from the content and refer chapter and page number, it would be tagged with the content automatically. The other way could be providing all details if the syllabus is not added digitally in the software.
- II. **Multilingual (Questions Item Contents):** Items in any language could be developed or translated. In the Pakistani context, it would support English and Arabic script.
- III. **Bloom's Taxonomy Cognitive Levels:** Items can be defined and developed as per cognitive levels in Bloom's taxonomy.
- IV. **Difficulty Levels:** Items can also be defined and developed for various difficulty levels.

The Bloom's taxonomy and difficulty levels would initially be decided mutually by the item writer, reviewer, and selector. However, for standardized tests, these levels would be decided through large-scale administration of items and employing statistical tools.

V. Mathematical Equations Support: The software interface would support Mathematical equations and notations for developing items in subjects having numerical values.

VI. Scientific Notation Support: The software interface would also support scientific notations for developing items in specific science subject areas.

VII. Add Media (Picture and Video Clips as Question Contents): One could add pictures and audio-video clips for developing items (in case online testing is adopted).

The process of QIB development starts by adding the curriculum digitally to the software for digital tagging. It will be added on the following lines: class; subject name and code; chapter number; SLO; topic; sub-topic; page number; and valid year of syllabus. Adding/Saving Question Items by Item Writer: The item writer will start by adding necessary information for the question item for tagging. They will add information regarding class; subject; chapter number; SLO; topic; sub-topic; page number; cognitive level, such as (knowledge, understanding, application, analysis, synthesis, evaluation/creativity); difficulty level (difl); marks per item, and answering time (if to be used for online tests, the items would appear as per difficulty level from easy to difficult in a sequence until a candidate replies incorrectly. Otherwise, an easy item with lesser marks would appear); entry date and time (these would be auto-generated); entered by (auto-generated from item writer/reviewer/selector's ID assigned to them) and questions item unique SNO/Index (auto-generated).

Each item writer whenever logs in on the portal for work, after retrieving the data, he would be able to add, delete, edit, save and submit the question item. But once the item is submitted, no one would be able to see it again, until returned to him for improvement. For example, if he wants to edit the MCQ, he would either change the stem of the item or the distractors, or correct response be changed. The reviewer can review submitted items by the item writer and can modify, approve, or reject. Similarly, the selector can review approved items submitted by the item reviewer. The selector can finally approve or reject question items. All of them would not be able to take print of any item or screen shot at any stage of the process. The proposed software portal would function, and the item writer would fill in all the fields/boxes before adding MCQ as described below in Figure 6:

Fig. 6: QIB MCQ Entry Form

Each of the items would have a barcode such as taken from syllabus content:

- a) For example, combination of:

Subject, Chapter #, Page #, Topic, Sub Topic & Index

1001 01 02 100 07 01

Barcode: 1001011010011101



- b) But if the item contains only SLO, then the SNO would be as:

For example, Combination of: Subject, SLO, Page #, & Index

1001 1005 0027 02

Barcode: 100111111011011101110



The item writer would also develop a short response question (SRQ) item by adding all the details in the boxes before writing the item. He would also translate the item into another

language as required. Moreover, the correct response would also be added as a key for the assessment process. Figure 7 depicts the short response question (SRQ) form below:

Fig. 7: QIB Short Response Question Form

Similarly, a descriptive response question (DRQ) item would also be added by the item writer by filling in the boxes with necessary details and writing item content, translating it, and providing an answer key and rubrics for marking. Figure 8 below depicts the details:

Fig. 8: QIB Descriptive Response Question Form

The dashboard of the item writer, reviewer, and selector would be presented as given in Figure 9 below:

Fig. 9: Dashboard Display for QIB for the Participants

As shown in Figure 9 above, the item writer, reviewer, and selector can view their performance on the dashboard. They could edit their personal profile, add or edit MCQs, Short Response Question items (SRQ), Descriptive Response Question items (DRQ), pending list, and rejected list by making necessary changes suggested and required. They could see their performance summary regarding the total number of MCQs, SRQs, DRQs, and subjects submitted on the software portal too. The question item bank software would work as shown in Figure 10 below:

bank software would work as shown in Figure 10 below:

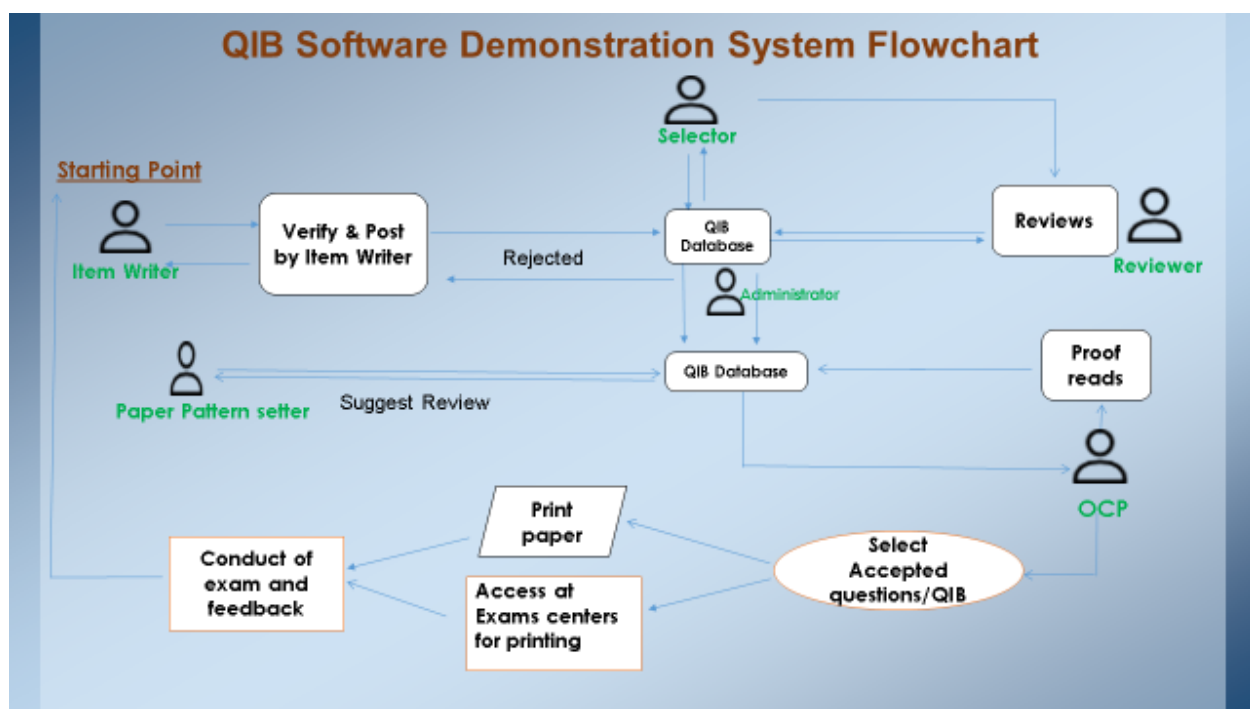


Fig. 10: Conceptual Map/Flow Chart for QIB and Auto Generation of Question Paper

The proposed protocol schemata for developing a web-based question item bank ensures all the safety, security, and confidential measures. It would provide complete solutions for the examining agencies to use it for the development of items on the web by engaging a large number of item writers, reviewers, and selectors simultaneously.

The item writer would start writing question items in either MCQ, SRQ, or DRQ form, specifying the cognitive or difficulty levels. After making necessary corrections and editing and verification, he would submit the item to the system software. It would travel to the item reviewer automatically through the system administrator. The reviewer would review it and, if it is appropriate in content and cognitive and difficulty levels, would send it to the item selector through

the administrator for final selection. In case any improvements are required, the item would be referred back to the item writer with comments and suggestions. When the item is finally selected, it is stored in the software database.

The software for the generation of question paper would generate by accessing the database. Instead of engaging paper setters, the examining bodies would engage a paper pattern setter. The paper pattern setter would instruct the software to select each question item as per cognitive and difficulty levels, either from the domain of defined content from a range of chapters/page numbers or from SLOs. When finalized, the Officer of Confidential Press (OCP) would request the software through secure and time-bound access for generating the question paper. The software would generate a single question paper and would be locked automatically. The next time, the same process would have to be followed to generate another question paper. The OCP, after proofreading and necessary editing, would either print question papers in bulk quantity or give limited access to the examination centers for on-the-spot printing on the examination day. When the tests are administered, the feedback would come from students, teachers, and parents, leading to the improvement of items. The items would be improved accordingly.

The process of developing the question item bank is neither linear nor static; rather it is cyclic. Figure 11 below illustrates:

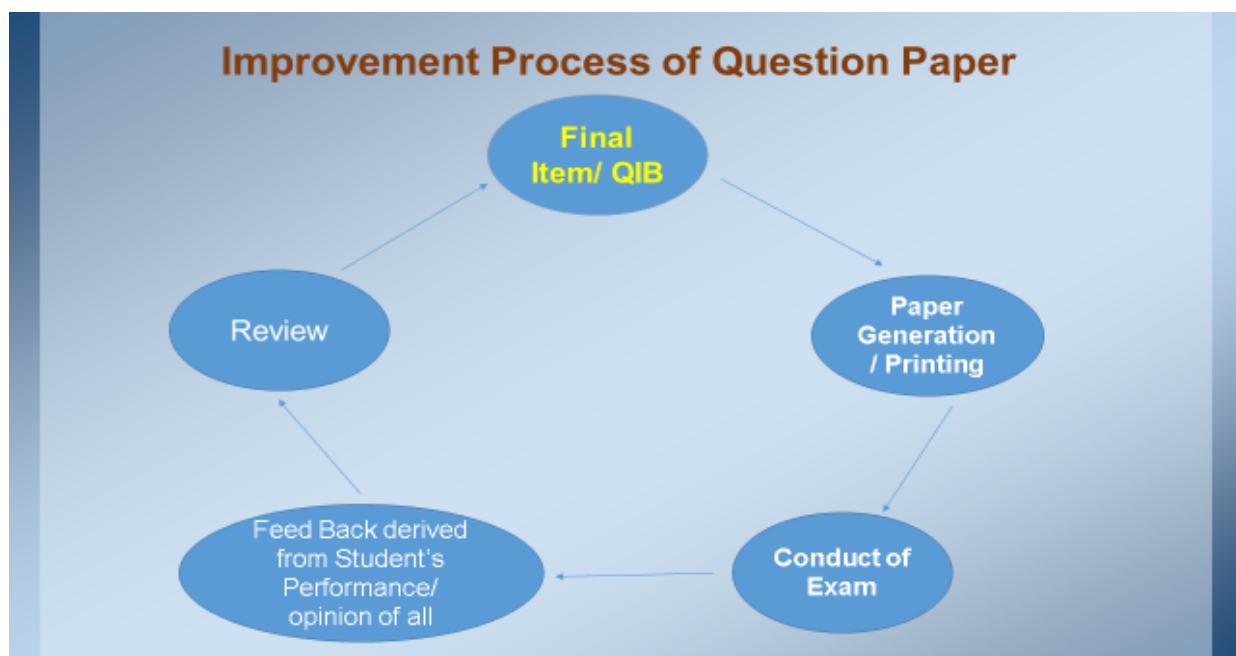


Fig. 11: Improvement Process for Question Paper Generation Using QIB Software

As given in Figure 11 above, the improvement process of the question item bank follows a cyclic process. The question paper is generated by accessing the QIB data bank and printed accordingly. The question paper would be administered in the exam. The examinees, the teachers, the parents, educationists, and even the public would provide feedback regarding the question items. The feedback data could also be obtained to examine the items statistically to find validity and reliability. This feedback would lead to a review of items for improvement. In this way, the examining bodies may develop standardized tests through validity and reliability examinations.

Discussion

The proposed protocol schemata for developing a web-based question item bank ensures all the security, safety, and ease of access measures. Some of these are given below:

- a) The participants - item writers, item reviewers, and item selectors - can work online 24/7 from anywhere. It ensures ease of access and working as per their availability of time. Akuadi (2010) supports the use of computer software for developing and storage of QIB. She stresses the need for using computer-based assessment programs in developing countries (such as Pakistan) to cope with the practices employed in the developed world.
- b) All the participants would be anonymous to one another. They would log in through their profile IDs using passwords and work accordingly. This would ensure that they could not form any group outside the portal and thus could not work or copy the items for marketing purposes.
- c) The items' development would go through the process of item development, review by subject and assessment experts, and final selection by assessment experts to ensure that the items are properly designed and valid.
- d) The QIB software would have the options for multilingual scripts, all types of question items, mathematical equations, and scientific notations supported. Online ExamMaker software also supports mathematical equations and scientific notations. Media can also be added for online testing facility.
- e) They would not be able to get print or even a screenshot of any item. This would ensure the safety and security of items.
- f) The items would be developed as per Bloom's taxonomy cognitive levels and difficulty levels. Prasad and Xavier (2006) also advocated that items should be prepared with due care

as per cognitive domains in Bloom's taxonomy and difficulty level. Moreover, they stressed that QIB storage should be safe from trespassing and hacking.

- g) The question item bank would be kept offline, except when tests are required to be generated. On request from an authentic and authorized person from some other examining agency, the custodian officer confidential press (OCP) shall send him a one-time password to give directions and instructions for generation of question paper as per specifications decided from the bank's repository. The system shall automatically lock once the command is executed. The available online software for QIB has some of these features discussed below: multiple type assessment items, and 24/7 usage from anywhere (Examssoft by Turnitin, eSystem Question Bank by Dory (n.d), rMarker by REDMarker (2019), TestGenius, and option for offline usage of the software.

The proposed protocol schemata for developing web-based question item bank for summative assessment provides a comprehensive, safe, and secure solution for establishing a question item bank for use by the examining bodies in Pakistan.

Conclusion

The aim of this study was to propose a protocol schemata for developing a web-based question item bank for conducting summative assessments. Question item banks can be developed by a large number of subject specialists in collaboration with assessment experts in various subjects across the curriculum. The objective of this study was to propose a protocol schemata for the development of software for web-based development of question item banks and auto-generation of question papers for summative assessment.

A thorough review of relevant literature has revealed that the trend of developing question item banks has been in vogue, and recent developments in question item banks applying web-based applications are attracting the attention of researchers across the globe. Many QIB software tools are available online having many features for ease and security. Examining bodies in Pakistan, such as education boards need some secure and safe QIB software with exclusive proprietary rights to make their assessments authentic. Hence, efforts are in progress to establish question item banks at various education boards. They a need for some QIB software having comprehensive constructs and interface options where auto-generation of question papers could be possible for administering summative assessments. The proposed protocol schemata can address all the requirements for developing and storing question items that could generate question papers as per the requirements

of the test as per difficulty level and Bloom's taxonomy of cognitive domains. It would also ensure secrecy and security protocols and periodical revision and updating of question items. Some of the salient features of the proposed protocol schemata include:

- a) The participants—item writers, item reviewers, and item selectors—can work online 24/7 from anywhere. It ensures ease of access and working as per their availability of time.
- b) All the participants would be anonymous to one another. This would bar them to form any group outside the portal for working or copying the items for marketing purposes.
- c) The item development and review process would ensure the items are proper design and validity as per blooms taxonomy cognitive domain and difficulty levels.
- d) The QIB software would have the options for multilingual scripts, all types of question items, mathematical equations, and scientific notations supported and adding media options.
- e) Screen shots or prints would be disabled ensuring security of items.
- f) The question item bank would be kept offline, except when tests are required to be generated.

On request from an authentic and authorized person from another examining agency, the custodian officer confidential press (OCP) shall send a one-time password to give directions and instructions for the generation of question papers as per specifications from the QIB's repository. The system shall automatically lock once the command is executed.

Briefly speaking, both objectives of the study were achieved by a thorough review of relevant literature to become aware of concepts of assessment, standardized tests, question item bank, its development, and pros and cons. The proposed protocol schemata for developing a web-based question item bank for summative assessment provides a comprehensive, safe, and secure solution for establishing a question item bank for use by the examining bodies in Pakistan.

Recommendation:

The proposed protocol schemata for developing web-based question item bank for summative assessment, being comprehensive, secure and having multiple features, is recommended for use by the education boards and examining bodies in Pakistan

References

- Akuadi, C. O. (2010). Sustainable assessment and evaluation strategies for open and distance learning. *Turkish Online Journal of Distance Education-TOJDE*, 11(4), 121-129.
<https://doi.org/10.17718/tojde.1302-6488>

- Bearman, M., & Ajjawi, R. (2019). Students' motivation for rubric use in the EFL classroom assessment environment. *Frontiers in Education*. Retrieved June 8, 2014, from <https://www.frontiersin.org/articles/10.3389/feduc.2019.00003/full>
- Binet, A., & Simon, T. (1916). *The development of intelligence in children (The Binet-Simon Scale)*. Williams & Wilkins Company.
- Biswas, P., & Pradham, B. (2002). Assessment and evaluation in distance education.
- Board of Intermediate & Secondary Education. (1997). *The calendar of the Board of Intermediate & Secondary Education Rawalpindi Volume II (Rules)* (pp. 31-39). BISE Rawalpindi.
- Boud, D. (1995). Assessment and learning: Contradictory or complementary? In P. Knight (Ed.), *Assessment for learning in higher education* (pp. 35-48). Kogan Page Ltd.
- Danili, E., & Reid, N. (2005). Assessment formats: Do they make a difference? *Chemistry Education Research and Practice*, 6(4), 204-218. <https://doi.org/10.1039/B5RP90018F>
- Dory, V. (n.d.). *Speedwell software. eSystem Question Bank*. Retrieved January 22, 2024, from <https://www.speedwellsoftware.com/exam-software/question-bank-software/>
- Economic Adviser's Wing. (2023). *Pakistan Economic Survey 2022-23*.
- Erwin, T. D. (1994). Assessing learners in higher education (S. Brown & S. Knight, Eds.). Kogan Page.
- Gandolfi, E. (2018). Virtual Reality and Augmented Reality. In K. Kennedy & R. E. Ferdig (Eds.), *Handbook of research on K-12 online and blended learning* (2nd ed.). Retrieved June 8, 2024, from https://www.researchgate.net/publication/329161707_Virtual_Reality_and_Augmented_Reality
- GGI Insight. (2024). *(Gray Group International)*. Retrieved January 18, 2024, from <http://graygroupintl.com>
- Ghio, F. B., Bruzzone, M., Roja, L., & Cupani, M. (2022). Preliminary development of an item bank and an adaptive test in mathematical knowledge for university students. *European Journal of Science and Mathematical Education*, 10(3), 352-365. <https://doi.org/10.30935/scimath/11449>
- Gielen, S., & de Wever, B. (2015). Integrating peer assessment in higher education: The need for a revised approach. *Frontiers in Education*. <https://doi.org/10.3389/feduc.2015.00003>

- Grodsky, E., Warren, J. R., & Felts, E. (2008). Testing and social stratification in American education. *Annual Review of Sociology*, 34(1), 385-404.
<https://doi.org/10.1146/annurev.soc.34.040507.134711>
- Hiscox, M. D. (1983). A balance for educational item banking. Paper presented at the annual meeting of the National Council for Measurement in Education, Montreal, Quebec, 12-14 April. <https://eric.ed.gov/?id=ED232033>
- Inter Board Committee of Chairmen (IBCC). (2021). Establishment of question item bank jointly by IBCC and all member boards. *Proceedings of the IBCC 169th meeting, agenda item No. 6*. Islamabad, 24-25 March.
- Coleman, J. S. (1966). *Equality of educational opportunity*. U.S. Government Printing Office.
- Kirkpatrick, D. R., Mistry, K. P., & Leshikar, G. A. (2020). Assessing interpersonal and communication skills. *Journal of Graduate Medical Education*, 12(3), 249-260.
<https://doi.org/10.4300/JGME-D-20-00090.1>
- Meier, D., & Arther, J. A. (1984). Issues in item banking. In J. Millman & J. A. Arther (Eds.), *Seven assessments of students and schools more effective than standardized tests* (pp. 34-56). Teachers College Press.
- Millman, J., & Arther, J. A. (1984). Issues in item banking. *Journal of Educational Measurement*, 21(4), 315-330. <https://doi.org/10.1111/j.1745-3984.1984.tb01030.x>
- Prasad, S. K., & Xavier, G. (2006). Use of ICT in evaluation: On-demand examination system of National Institute of Open Schooling (NIOS), India. Paper presented at the workshop on Demand Examination System, National Open University of Nigeria (NOUN), NOUN Headquarter Lagos, 6-17 November.
- Quellmalz, E., & Pellegrino, J. W. (2009). Technology and testing. *Science*, 323(5910), 75-79.
<https://doi.org/10.1126/science.1168046>
- REDMarker. (2019). *Cutting-edge online assessment tools*. Retrieved May 27, 2024, from <https://redmarker.io/rtestgen-exam-paper-generator>
- Smith, J. A. (2018). The role of standardized testing in educational assessment. *Assessment in Education: Principles, Policy & Practice*, 25(3), 215-230.
<https://doi.org/10.1080/0969594X.2018.1451234>
- Turnitin. (2024). *Examsoft*. Retrieved February 22, 2024, from <https://examsoft.com/>

- Thorndikw, R. L., Cunningham, G.K., Thorndike, R.L., & Hagen, E. (1964). *Measurement and Evaluation in Psychology and Eductaion*. NewYork: Wiley.
- Wiggins, G. (1998). *Educative Assessment: Designing Assessments to Inform and Improve Student Performance*. San Francisco, CA: Jossey-Bass.