A COMPARATIVE ANALYSIS OF NATIONAL TESTING AGENCIES: STRUCTURE, FUNCTIONS, AND EDUCATIONAL IMPLICATIONS

Zainab Qamar[*1], Iffat Khalil[2]

*[*1]Research Associate, Faculty of Social Sciences and Humanities, Minhaj University Lahore, Pakistan
PhD Scholar, Institute of Education and Research, University of the Punjab, Lahore, Pakistan
[2]PhD Scholar, Institute of Education and Research, University of the Punjab, Lahore, Pakistan*

[*1]zainab.ra@mul.edu.pk; [*1]zainabqamar10@gmail.com; [2]iffatkhalil25@gmail.com

Corresponding Author: *
Zainab Qamar

Abstract
The paper examines the structure, functions and challenges of testing agencies in Pakistan and also places them within international comparative framework. International assessments are psychometrically sound, transparent and centralized practices can be found in such agencies as the Educational Testing Service (ETS) in United States and the Universities and Colleges Admissions Service (UCAS) in the United Kingdom. In Pakistan, however, Higher Education commission, Education testing Council (ETC), National Testing Service (NTS), provincial boards have been dysfunctional, over reliance on rote based multiple choice testing, crippled with bad governance and poor technological application. This paper, with a qualitative method of document analysis, discloses the shared problems of ineffectiveness, duplication and transparency that undermine the confidence of the population and the fairness among the students. Comparative insights of the study indicate that there is a need to introduce the reforms centered on the psychometric rigor, digital innovation, and central governance. Such findings suggest that a re-branding of Pakistani testing agencies in the light of the international patterns can assist in many ways in enhancing transparency, justice, and international recognition of the national qualification.

INTRODUCTION

The education assessment agencies play a key role in ensuring equity, quality and accountability within the education systems of the world. More industrialized environments, such as in the United States and the United Kingdom, have come up with stringent measures employed in evaluating student learning and also regulating entry into tertiary education. Using Educational Testing Service (ETS) in the U.S as an example, the administration of globally recognized exams such as GRE and TOEFL is entrusted to it, but in the U.K. Universities and Colleges Admissions Service (UCAS) is in charge of administering centralized admissions through standard evaluation systems (Altbach, 2015; Braun & Kanjee, 2006). Not only will these systems provide valid tools of assessing academic readiness, but will also provide transparency and meritocracy in the admissions process. Even such emerging economies like Asia (India or China) have managed to create strong national testing institutions since they realize the necessity to standardize

education and ensure that different groups of people receive the same opportunities (Gupta, 2020).

The assessment agencies come in especially in developing countries where access, disparity and quality of education remains a clear picture. Such agencies are very significant in the context of Pakistan, in as far as balancing differences between institutions of various academic standards are concerned. They not only do anticipate the ability of students but also guarantee that they are benchmarked at the same level across the institutions, either privately or owned by the state (Shah, 2019). With the increasing number of students who want to get higher education and competitive jobs, testing agencies make sure that college and university admissions are not based on social or economic standing, but on merit. In such a way, the effectiveness, integrity, and equity of these agencies directly affect the personal career path and the development of the human resource nationally (Khan & Riaz, 2021).

Although they are important, the assessment agencies of Pakistan are encountered by numerous challenges, which bring down their efficiency. The absence of transparency, minimal technological adoption, overlapping, and multiple accusations of malpractice have reduced trust among people (Ahmed & Batool, 2017). The students are frequently reporting inconsistent standards of testing, and the universities have been raising concerns about the predictive validity of the admission tests in predicting academic success (Saeed, 2020). Additionally, the best practices of computers like testing, psychometric validation, and adaptive testing models are not very popular in Pakistani systems. The research problem is that current testing agencies are unable to meet international standards while also addressing the specific needs of local education systems.

In Pakistan, the Higher Education Commission (HEC) regulates the standard of education in the tertiary level and has even initiated programs like the Education testing Council (ETC) to administer standardized exams. Nonetheless, ETC is a relatively new concept and not one that is broadly accepted (HEC, 2018). The National Testing Service (NTS), which was formerly a major establishment has come under fire over its governance, fee framework and credibility (Malik, 2016). The provincial boards of intermediate and secondary education have a role of assessment as well but are mostly concentrated on school-level examinations with a little innovation. As a result, the Pakistani ecosystem of assessment seems to be disjointed, and national and provincial agencies lack adequate coordination (Shah, 2019).

Even though a lot of work of international agencies has been recorded on the role of assessment bodies in attaining fairness and academic quality, there is paucity of literature on how the role of the agencies in Pakistan fits or does not fit in the practices. Earlier research has primarily dealt with one of the entities, including NTS, or provincial boards without providing a comparative overview (Ahmed & Batool, 2017; Saeed, 2020). The integration of technology, validity of assessment frameworks and governance issues in the testing agencies in Pakistan have also not been critically questioned. The comparison with other global models (ETS or UCAS) is poorly studied. The current research attempts to fill this gap by placing the testing agencies in Pakistan in a global context, and pointing out the strengths and weaknesses.

Objectives of the Study

1. Discuss the structure, functions and challenges of educational testing agencies in Pakistan.
2. Compare Pakistani testing agencies with the international testing agencies.
3. Identify gaps in efficiency, transparency, and innovation within Pakistan's assessment ecosystem.
4. Propose reforms that aid in aligning the national agencies with international standards and local education demands.

Research Questions

1. What are the structures, functions and challenges of educational testing agencies in Pakistan?
2. What are the similarities and differences between Pakistani testing agencies and international testing agencies?
3. What are the existing gaps when it comes to efficiency, transparency, technological integration, and innovation on the Pakistani assessment ecosystem?
4. To what degree can the reforms in national testing agencies make them meet the

international standards in terms of responding to the local needs in education?

Literature Review

The educational testing agencies play a pivotal role in providing justice, responsibility and quality to the education systems. The current state of education considers testing not only as a means of grading but a means of measuring the results of learning, determining personal skills, and standardizing access to advanced levels of education or work. Researchers note that assessment is a more general term that covers formative and summative processes whereas testing is a more specific term as it involves systematic procedures that aim at quantifying the learning and skills. The growing need of transparent and merit-based systems in higher education across the world has increased the relevance of standardized testing agencies. These agencies are necessary in the societies with mass education as well as international mobility of people to ensure that there is an equal opportunity as well as adapting the local educational systems to international standards.

In Pakistan, the history of testing agencies depicts the struggle of the state on the whole to establish a balance between equity and quality in education. The formation of Higher Education Commission (HEC) marked a step towards the regulation of higher education and a process that had been disjointed. HEC has presented a national-scale system of testing, through its Education Testing Council (ETC), to allow a decreasing number of inconsistencies in university-specific systems of admission testing and to guarantee a meritocratic process of entry to higher education. Simultaneously with HEC, the National Testing Service (NTS) was also started in 2002, which was the first independent testing service in the country on the example of the Educational Testing Service (ETS) in the United States. It was mandated to carry out entry examinations, professional licensing examinations and recruitment examinations on behalf of government and non-government agencies. Although NTS has established a new culture of standardized testing in Pakistan, it has frequently been criticized because of lack of transparency, excessive charges and lack of demonstration of psychometric rigor.

In addition to the federal, provincial governments in Pakistan have also set up their assessment and testing bodies. In Punjab, Sindh and Khyber Pakhtunkhwa, as an example, provincial testing agencies have been established in order to administer testing of recruitment and university entrance and certain certification. The logic to these bodies is to decentralize testing and make it more accessible locally. Nevertheless, multiple agencies have resulted in duplication of activities, inconsistency in the quality of tests and confusion among applicants as they can be subjected to overlapping or repeated testing. Compared to the centralized system of other countries, the parallel structures in Pakistan add to the fragmentation that undermines the standardization and introduces the disparities between the educational opportunities in the provinces. Such redundancy does not only cause financial strain on the student but also destroys the confidence of the population in the efficiency of the assessment systems.

The functions of testing agencies are wide in scale and involve test design, test administration, scoring and reporting. They participate in the assurance of content validity, reliability and fairness of examination, determination of test specifications and development of scoring rubrics. Companies such as ETS and the National Testing Agency (NTA) in India have been conducting numerous research and capacity-building works within the psychometric sphere to make sure that its tests are robust and valid internationally. In Pakistan in Pakistan, testing agencies have been slow, however, in adopting advanced psychometric technology, or in conducting systematic studies of the fairness and validity of their tests. Despite the fact that ETC and NTS do test at the scale of large, there remains the concern about quality of test construction, overuse of multiple-choice question forms with little use of higher-order thinking skills. These weaknesses lead one to wonder whether or not such testing environments is a measure of the whole student competencies or merely a measure of rote learning which is the standard with the schooling system.

Pakistan can be compared to other countries of the international models of testing. The Educational testing Service (ETS) in the USA is an illustrious global pioneer in the formulation of examinations

such as GRE, TOEFL, and SAT that are executed with extremely high levels of psychometric rigor and global acceptability. Another approach offered by the United Kingdom Universities and Colleges Admissions Service (UCAS) focuses on centralized processing of admissions, rather than the standardization of tests, which simplifies the admissions route to students and offers transparency. One similar example of such a large-scale centralized testing is already present in South Asia, with the National Testing Agency (NTA) established in 2017 in India already handling millions of candidates annually with digitized and computer-based platforms. These models appear superior to the Pakistan system in the sense of its fragmentation and absence of technology that could be substantially advanced.

Evaluation and testing theoretical models center on application of measurement in standardization of learning results and fairness. Classical Test Theory (CTT) and the Item Response Theory (IRT) are two famous theories that are used to design a valid and reliable test (Embretson & Reise, 2000). CTT is concerned with observed scores as a measure of true ability and error whereas IRT describes responses of test takers in terms of latent trait and item difficulty. These systems are the ones that guarantee psychometric validity and equity in high volume testing systems. Constructivist perspectives on assessment also highlight formative and authentic assessments as a means of capturing higher-order skills beyond rote memorization (Black & Wiliam, 2009). Together, these theoretical models provide a foundation for evaluating the effectiveness of testing agencies.

Past studies in developed contexts highlight the success of centralized testing agencies in promoting meritocracy. For instance, Braun and Kanjee (2006) argue that standardized assessments ensure accountability and comparability across institutions. Altbach (2015) further notes that global higher education increasingly relies on admission tests such as GRE, GMAT, and SAT to benchmark student quality. In contrast, developing countries face challenges such as limited capacity, corruption, and resource constraints, which hinder the credibility of testing systems (Gupta, 2020). The study conducted in India and China demonstrates how the national agencies have used technological solutions in the form of computer-based tests and adaptive exams to support efficiency and transparency despite a massive population (Zhao, 2018). In Pakistan, research has found that testing agencies have serious flaws. Ahmed and Batool (2017) also point out the problem of duplication, absence of transparency, and inconsistent governance in such agencies as the National Testing Service (NTS). According to Malik (2016), there are governance and accountability breaches that diminish the confidence that people have in such institutions. Shah (2019) addresses the factor of greater fragmentation, in which provincial boards, HEC, and private agencies are working with overlapping mandates without much coordination. Saeed (2020) also doubts the predictive validity of Pakistani admission tests, observing that the student scores do not always correlate with the results in higher education. Such results suggest that Pakistan is underdeveloped in the cases of psychometric models and technological innovations adopted on the global scale.

The world-famous agencies can be compared and offer their helpful models. Institutions, including Educational Testing Service (ETS) in the United States, have certification tests such as GRE and TOEFL that are highly regarded as reliable and valid (ETS, 2020). Likewise in the United Kingdom, UCAS also provides centralized and transparent admission processing, minimizing duplication and bias of admissions (UCAS, 2021). Standardized tests such as the GMAT and SAT are created with the assistance of IRT, and they are not biased amongst different populations or circumstances (Kuncel & Hezlett, 2010). The validity of these agencies lies in the fact that they integrate psychometric theories, online testing systems, and constant research into test construction. HEC in Pakistan developed Education Testing Council (ETC) to centralize and standardize the university admissions. ETC, nevertheless, is still in its initial phases; it is not accepted by universities and students entirely (HEC, 2018). Once powerful, the National Testing Service (NTS) suffers a loss of credibility and falling trust because of the mismanagement accusations (Malik, 2016). The provincial boards pay much attention to the examinations at the secondary level, yet their work is criticized as being outdated, too reliant on the rote learning and prone to malpractice (Shah, 2019). The

discontinuity of the Pakistani system is in contrast to integrated systems such as ETS and UCAS that focus on transparency, fairness and efficiency. Comparative analysis reveals than there is a big disparity between Pakistan and the global best practices. To illustrate, when ETS and UCAS operate using computerized adaptive testing and using centralized admission processing, Pakistani agencies continue to rely on paper-based tests with low psychometric validation (Ahmed and Batool, 2017; Saeed, 2020). In addition, the question of governance and accountability is also still relevant in Pakistan which negatively affects the credibility of findings. Practices in other countries have indicated that confidence in assessment systems is founded on transparency, rigorous quality assurance and incorporation of current technologies (Braun & Kanjee, 2006). Pakistan will continue to run the risk of increasing the distance between its educational practices, assessment-wise, and the international standards without such similar reforms. Transparency and accountability are one of the most significant issues that have been repeatedly noted in the Pakistani testing agency literature. Concerns about leaks of papers, no external auditing and no published information regarding test reliability and validity have been raised on several occasions. These loopholes cast doubt on the reliability of testing results and create the risk of being viewed as being corrupt or biased when recruiting and admitting to school. In comparison, global organizations like ETS frequently release their studies on fairness, psychometric validation and item analysis hence establishing credibility and trust among stakeholders. Transparency has been promoted even by the NTA of India who have published test specifications and sample questions in detail. Such practices can be used to restore the legitimacy in the case of Pakistani agencies to align the agencies with the global standards.

As a rule, the literature suggests that testing agencies would be critical instruments of equity and quality control but they are hindered by various challenges in various circumstances. The case of the international agencies proves that by means of application of psychometric theories, the development of technologies it is possible to increase the reliability and fairness; and Pakistan with its chaotic and uncontrolled system is still struggling with credibility.

The review emphasizes that complete reforms including the restructuring of governance, adoption and alignment of technology to the international best practices are required in the Pakistani testing agencies.

Research Design and Methodology

The current study was qualitative in nature and adopted document analysis and comparative review. As secondary data, official reports of Higher Education Commission (HEC), National Testing Service (NTS), Education Testing Council (ETC) and provincial boards as well as international testing agencies such as Educational Testing Service (ETS) and Universities and Colleges Admissions Service (UCAS) were utilized. Relevant peer-reviewed publications and policy materials were also analyzed. The thematic content analysis was performed to outline the main issues regarding the governance, transparency, validity, technological integration and compliance with international standards. Comparative perspectives were made to show similarities and differences drawn between Pakistani and global assessment practices. This methodology yielded a logical and valid knowledge on the strengths, weaknesses, and reform requirements of testing agencies in Pakistan.

Results and Discussion

The first theme was the organization and operations of educational testing agencies. In Pakistan, evaluation tasks are shared among a number of institutions such as National Testing Service (NTS), Education Testing Council (ETC) of the Higher Education Commission and provincial examination boards (Ahmed & Batool, 2017). Conversely, test agencies such as Educational Testing Service (ETS) in the United States are highly specialized and centralized organizations with a very definite task to design and administer such tests like GRE, TOEFL, and Praxis (ETS, 2020). On the same note, the United Kingdom UCAS facilitates the admissions to the universities within one unified authority, which promotes transparency and effectiveness (UCAS, 2021). It can be seen in the comparison that where the system in Pakistan lacks, the international agencies live on centralization and specialization.

The second theme was on governance and transparency. Pakistani testers are regularly accused of anomalies, such as leaked paper and grading

discrepancies (Malik, 2016). Public trust in NTS, once dominant, has eroded significantly due to accountability issues. Provincial boards also suffer from poor monitoring and susceptibility to malpractice (Shah, 2019). Conversely, ETS and UCAS ensure transparency through strict security protocols, psychometric validation, and independent oversight committees (Kuncel & Hezlett, 2010). India's National Testing Agency (NTA) provides another instructive example, having introduced digital exam monitoring systems to enhance accountability (Gupta, 2020). The contrast illustrates that while international agencies strengthen credibility through governance, Pakistan's weak accountability structures undermine stakeholder confidence.

The third theme addressed psychometric soundness and test validity. Most Pakistani assessments rely on rote-based multiple-choice questions with limited evidence of construct validity (Saeed, 2020). Tests administered by NTS and provincial boards rarely incorporate modern psychometric frameworks such as Item Response Theory (IRT). By comparison, ETS, GMAT, and SAT routinely apply IRT to ensure fairness across diverse populations and contexts (Embretson & Reise, 2000). Moreover, international agencies continuously conduct research to refine item pools and validate predictive reliability, something largely absent in Pakistan's system. The gap reflects a fundamental weakness in Pakistan's approach to ensuring fairness and accuracy in assessment.

The fourth theme considered technological integration. Pakistani agencies continue to depend on paper-based testing, with only limited and experimental use of online formats. Logistical inefficiencies and accessibility concerns have hindered adoption of computer-based testing. In contrast, ETS pioneered computer-based GRE and TOEFL decades ago, and the NTA in India administers large-scale online exams such as NEET and JEE using adaptive platforms (Zhao, 2018). UCAS also utilizes digital platforms to facilitate transparent university admissions processing. Comparative evidence suggests that while international agencies embrace innovation, Pakistan lags behind, creating inefficiencies and opportunities for malpractice.

Finally, the analysis examined reform potential and alignment with global standards. Pakistan has recognized the need for change by establishing ETC under HEC, yet its limited adoption reflects resistance from universities and a lack of infrastructural support (HEC, 2018). International models show that credibility and efficiency are enhanced when agencies are centralized, technologically advanced, and guided by psychometric principles. Pakistan could adopt a hybrid approach by integrating digital testing systems, strengthening governance mechanisms, and aligning assessment content with higher-order cognitive skills as recommended by formative assessment theories (Black & Wiliam, 2009). The comparative review therefore highlights both the challenges facing Pakistan and the opportunities available through global best practices.

Taken together, this comparative analysis demonstrates that while testing agencies worldwide function as pillars of quality assurance, Pakistan's fragmented and outdated system struggles to achieve credibility. International examples show that centralization, transparency, psychometric rigor, and technological adoption are essential for building trust and fairness. Unless similar reforms are pursued, Pakistan risks widening the gap between its educational practices and international benchmarks.

Tabel 1
Comparative Insights of the Study

Dimension	Pakistan (NTS, ETC, Boards)	International Models (ETS, UCAS, NTA, etc.)
Governance Structure	& Fragmented, multiple overlapping agencies; lack of coordination (Ahmed & Batool, 2017).	Centralized, specialized bodies with clear mandates (ETS, UCAS).
Transparency & Credibility	& Frequent allegations of leaks, mismanagement, weak accountability (Malik, 2016; Shah, 2019).	Strong monitoring, psychometric validation, strict security (Kuncel & Hezlett, 2010).
Psychometric Rigor	Reliance on rote-based MCQs; limited validity and reliability checks (Saeed, 2020).	Use of Item Response Theory (IRT) and continuous test validation (Embretson & Reise, 2000).
Technological Integration	Mostly paper-based exams; limited online adoption.	Computer-based, adaptive, AI-supported testing (ETS GRE, India's NTA digital exams).
Admission Processing	Universities rely on multiple agencies, leading to duplication.	Centralized admission services (UCAS), reducing duplication and bias.
Reform Potential	ETC created but faces resistance; outdated infrastructure.	Continuous reforms, digital systems, alignment with equity and higher-order skills.

This table consolidate the comparative insights from the thematic analysis. It becomes evident that Pakistan's system is hindered by fragmentation, weak accountability, and outdated practices, whereas international agencies succeed through centralization, psychometric rigor, and technological innovation. The model demonstrates that without reforms, Pakistani testing agencies risk perpetuating inefficiencies and inequities, while adopting international practices could enhance transparency, fairness, and global credibility.

Challenges

The national testing agencies in Pakistan have a number of problems that restrict their usefulness and trust. One of the greatest problems is excessive use of multiple-choice questions, which can be graded with ease and do not help assess higher-order competencies, including critical thinking, creativity, and problem-solving, as is the case with international models that combine performance-based and analytical tasks. There are also inconsistencies in standardization across the provinces with variances in

testing requirements and grading practices becoming a challenge to the students. The poor technological infrastructure also renders the system weak since internet-based testing systems are less advanced and not available to majority citizens particularly in the rural areas. The issue of corruption, paper leaks, and accountability remain the factors that hinder the trust of the population, and the lack of proper orientation with the 21st-century skills and global standards makes Pakistani students less competitive. In addition, the lack of research, innovation and cooperation with international agencies such as ETS or UCAS points at the inherent flaws in the system, and reforms are required to enhance the level of transparency, standardization and international adaptability.

Conclusion

Testing agencies are involved in the achievement of quality, equity, and standardization in education and as are the gatekeepers to higher education and employment possibilities. In Pakistan, the HEC, NTS, provincial boards, and ETC have played a role in the

assessment process, although systemic problems of fragmentation, inadequate transparency and technological innovation have impeded their success. Pakistani agencies are yet to adopt psychometric validation, performance-based measures, and highly novel digital systems fully, as compared to the global models such as ETS and UCAS, and this has placed students in a destitute position of competitiveness and international recognition. This study shows that there is an urgent need to have systemic changes which could fill the gaps within the governance system, enhance accountability, and align the testing system with the modern needs of education. In the absence of these reforms, the validity of national testing systems will be doubted and the academic and professional mobility of students will still be limited.

Recommendations

1. Enhance psychometric validation of tests to promote fairness and credibility to various groups of students.
2. Concentrate testing operations to minimize duplication, enhance economies of scale and maximize financial resources.
3. Add digital testing platforms, remote proctoring technologies to increase access and minimize malpractice.
4. Increase transparency by releasing test reliability, validity indices, and test evaluation methods to the external arena.
5. Adjust the national testing systems to the international one, i.e. GRE, SAT, UCAS, etc. to enhance the international acceptance of Pakistani qualifications.

REFERENCES

Ahmed, R., & Batool, S. (2017). Challenges in the standardization of testing services in Pakistan. *Pakistan Journal of Education, 34*(2), 45–60.

Altbach, P. G. (2015). *Global perspectives on higher education*. Johns Hopkins University Press.

Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability, 21*(1), 5–31.

Braun, H., & Kanjee, A. (2006). National assessment of educational achievement. *Education Policy Series, UNESCO*, 1–37.

Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. Psychology Press.

ETS. (2020). *About ETS*. Retrieved from https://www.ets.org

Gupta, A. (2020). Standardized testing and higher education in India: A critical review. *Asian Education Review, 11*(3), 201–215.

Higher Education Commission (HEC). (2018). *Annual report*. Islamabad: HEC Publications.

Khan, M. A., & Riaz, T. (2021). The role of testing agencies in promoting meritocracy in Pakistan. *Journal of Social Sciences, 8*(1), 112–130.

Kuncel, N. R., & Hezlett, S. A. (2010). Standardized tests predict graduate students' success. *Science, 7*(3), 1081–1082.

Malik, S. (2016). Accountability and governance issues in Pakistan's National Testing Service. *South Asian Studies, 3*(1), 213–228.

Saeed, M. (2020). Predictive validity of standardized admission tests in Pakistani universities. *International Journal of Educational Research, 9(2)*, 101–112.

Shah, I. H. (2019). Educational assessment in Pakistan: Issues and prospects. *Bulletin of Education and Research, 41*(2), 79–96.

UCAS. (2021). *About UCAS*. Retrieved from https://www.ucas.com

Zhao, X. (2018). Innovations in standardized testing in China. *International Journal of Testing, 18*(2), 123–139.