LARGE SCALE ASSESSMENT IN PAKISTAN: A CRITICAL ANALYSIS OF DEVELOPMENT AND APPLICATION OF NATIONAL ACHIEVEMENT TEST

BY

Saima Abbas

NATIONAL UNIVERSITY OF MODERN LANGUAGES

ISLAMABAD

July, 2025

LARGE SCALE ASSESSMENT IN PAKISTAN: A CRITICAL ANALYSIS OF DEVELOPMENT AND APPLICATION OF NATIONAL ACHIEVEMENT TEST

By

SAIMA ABBAS

PhD. Education, National University of Modern Languages, Islamabad

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF THE
REQUIREMENTS FOR THE DEGREE OF

DOCTORATE OF PHILOSOPHY IN EDUCATION

To

DEPARTMENT OF EDUCATION
FACULTY OF SOCIAL SCIENCES

NATIONAL UNIVERSITY OF MODERN LANGUAGES, ISLAMABAD

July, 2025

NATIONAL UNIVERSITY OF MODERN LANGUAGES FACULTY OF SOCIAL SCIENCES

THESIS AND DEFENSE APPROVAL FORM

The undersigned certify that they have read the following thesis, examined the defense, are satisfied with the overall exam performance, and recommend the thesis to the Faculty of Social Sciences for acceptance.

Thesis Title: Large Scale Assessment in Pakistan: A Critical Analysis of Development and Application of National Achievement Test

Submitted by: Saima Abbas **Registration #:** PD-F18-120

Degree: Doctor of Philosophy

Discipline: Education

Dr. Aisha Bibi ________________________
Signature of Research Supervisor

Co- Supervisor ________________________
Signature of Co-Supervisor (If any)

Prof. Dr. Muhammad Riaz Shad ________________________
Signature of Dean (FSS)

Maj Gen Shahid Mahmood Kayani HI(M), Retd ________________________
Signature of the Rector

AUTHOR’S DECLARATION

I **Saima Abbas**

Daughter of **Sher Abbas**

Registration # PD-F18-120

Discipline <u>Education</u>

Candidate of **<u>Doctorate of Philosophy</u>** at the National University of Modern Languages do hereby declare that the thesis ‘**Large Scale Assessment in Pakistan: A Critical Analysis of Development and Application of National Achievement Test”** submitted by me in partial fulfillment of PhD degree, is my original work, and has not been submitted or published earlier. I also solemnly declare that it shall not, in future, be submitted by me for obtaining any other degree from this or any other university or institution.

I also understand that if evidence of plagiarism is found in my thesis/dissertation at any stage, even after the award of a degree, the work may be cancelled, and the degree revoked.

____

Signature of Candidate

____

Name of Candidate

____

Date

PLAGIARISM UNDERTAKING

I solemnly declare that research work presented in the thesis titled "**Large Scale Assessment in Pakistan: A Critical Analysis of Development and Application of National Achievement Test**" is solely my research work with no significant contribution from any other person. Small contribution/help wherever taken has been duly acknowledged and that complete thesis has been written by me.

I understand the zero-tolerance policy of the HEC and NATIONAL UNIVERSITY OF MODERN LANGUAGES towards plagiarism. Therefore, I as an Author of the above titled thesis declare that no portion of my thesis has been plagiarized and any material used as reference is properly referred/cited.

I undertake that if I am found guilty of any formal plagiarism in the above titled thesis even after award of PhD degree, the University reserves the rights to withdraw/revoke my PhD degree and that HEC and the University has the right to publish my name on the HEC/University Website on which names of students are placed who submitted plagiarized thesis.

Student / Author Signature: ______________________

Name: ____

ACKNOWLEDGEMENT

I am deeply grateful to Allah Almighty for granting me the opportunity to acquire knowledge and engage in this research. It is through the grace of the All-Knowing that we have been blessed with the guidance of His Holy Prophet (peace be upon him), who has provided us with foundational knowledge and will continue to be the source of wisdom in the future. I would like to express my sincere appreciation to Prof Dr. Nasir Mahmood, Vice Chancellor (AIOU), for his valuable judgment in allowing me to undertake this research thesis.

Furthermore, I extend my profound gratitude to my supervisor, Dr. Aisha Bibi, for her exceptional leadership, dedication, and unwavering guidance. Her love, support, and insightful feedback have been invaluable throughout the completion of this study. I am truly thankful for her patience and understanding, especially during times when I faced challenges or missed deadlines. Her vast knowledge, calm demeanor, and constructive criticism continuously inspired me to strive for improvement and lifelong learning. I feel deeply honored to have been her student and to learn under her mentorship.

I would also like to express my sincere appreciation to my former supervisor, Mr. Muhammad Rafique Tahir, whose support in providing the study permit (NOC) made this academic journey possible. I am indebted to my colleagues and friends for their valuable advice, collaboration, and encouragement, which helped me stay focused and motivated. I am especially grateful to my family for their constant support. My heartfelt thanks go to my parents and my husband for their moral support, motivation, prayers, and unwavering faith in my ability to achieve this milestone.

Lastly, I would like to extend special thanks to my coworkers—Ms. Abida, Ms. Mansoorah, and Ms. Zehra—for their consistent support, prayers, and kind wishes.

In conclusion, I am sincerely grateful to all those who supported me and kept me in their prayers throughout this significant phase of my life.

Saima Abbas

DEDICATION

I dedicate this thesis to my parents for their boundless love, unwavering support, and constant encouragement. I am deeply grateful for their presence and strength throughout this journey.

I also dedicate this work to my two angels, Meeno and Zoha, who have endured many hardships during this time. Their resilience, love, and patience have been a continuous source of inspiration and motivation for me.

Above all, I am forever thankful for the love and support of my parents, husband, and children—without whom this achievement would not have been possible.

ABSTRACT

Title: Large-Scale Assessment in Pakistan: A Critical Analysis of the Development and Application of the National Achievement Test

Assessment in education plays a crucial role as it systematically measures students' knowledge. This study critically analyzes the development and application of the National Achievement Test (NAT) in Pakistan and investigates teachers' perceptions of the test and its implementation. Utilizing an observational approach, this study retrospectively examines pre-existing data from the NAT administered in 2016 and 2019 across public and private schools in Pakistan. Specifically focusing on Grades 4 and 8 Mathematics and Science, the study aims to evaluate the consistency and effectiveness of the NAT assessment instruments in measuring intended criteria. Additionally, the research incorporates a correlational analysis to understand the relationship between NAT scores and teacher perceptions. The survey includes a combination of open-ended and closed-ended questions to gather both quantitative and qualitative insights. Furthermore, a checklist is utilized to assess the reporting of NAT results, enhancing the rigor and comprehensiveness of the analysis. The reliability of the scale was assessed through a pilot study. The study's population included Grade 4 Mathematics and Grade 8 Science teachers and students from public schools participating in NAT 2016 and 2019 in Islamabad and AJK, further stratified into urban and rural areas. A stratified sampling technique was employed to select samples. A mixed-methods approach was employed: item response analysis examined test development, while surveys gauged teachers' perceptions of NAT and its application, utilizing a standardized tool by Don Klinger adapted for this purpose. Pre-existing NAT 2016 and 2019 data were analyzed using statistical techniques such as descriptive analysis with IBM-SPSS and a percentage method for teacher perceptions. Checklist and Document analysis were used for reporting and dissemination. Findings from item analysis highlighted gaps in difficulty and discrimination index in both NAT tests. The tool for teacher perception emphasized the importance of NAT but showed negative results regarding its implementation. The checklist revealed gaps in reporting and dissemination of NAT. Recommendations include conducting pre and post-analyses to ensure fair assessments and effective communication among education stakeholders, along with proper reporting and dissemination for its effective utilization.

Saima Abbas, 2024

TABLE OF CONTENTS

2. LITERATURE REVIEW

3. RESEARCH METHODOLOGY

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

1PL	One Logic Parameter
2PL	Two Logistic Models
3PL	Three Logistic Models
AJK	Azad Jammu & Kashmir
BISE	Board for Secondary and Intermediate Examination
CTT	Classical Test Theory
DE	Distractor Efficiency
DFID	Department for International Development
DI	Discrimination Index
DIF	Difficulty Index
EQAO	Educational Quality and Accountability Office
IEA	International Association for Evaluation of Educational Achievement
ILSA	International Large-Scale Assessment
IPSLE	International Primary School Examination
IRT	Item Response Theory
LLECE	Latin American Laboratory for the Assessment Quality of Education
MLA	Monitoring Learning Achievement
MOFE&PT	Ministry of Federal Education and Professional Training
NAEP	National Institute for Educational Progress
NAPLAN	National Assessment Program, Literacy and Numerate
NEAS	National Education Assessment System
NEP	National Education Policy
OECD	Organization for Economic Cooperation and Development.
PEACs	Provincial/Local Education Assessment Institutions
PEITF	Physical Education Teachers Federation

PIE	Pakistan Institute of Education
PISA	Program for International Student Achievement
SACMEQ	Southern African Consortium for Monitoring Education Quality
TIMSS	Trends in International Mathematics and Science Study
UNESCO	United Nations Educational, Scientific and Cultural Organization
UNICEF	United Nations Children's Fund
WAT	National Achievement Test

.

CHAPTER: 1

INTRODUCTION

Assessment in education plays a crucial role as it systematically measures students' knowledge. It provides essential feedback throughout the learning process, making it a fundamental source for enhancing educational standards at all stages. Mohan (2023) describe assessment as the technique of collecting facts and data both qualitative and quantitative to track a student progress helping in planning future educational course of action. Ukwuije (2013) considers assessment as an all-encompassing term, where teachers measure certain aspects of student learning and evaluate the effectiveness of their teaching methods. Assessment encompasses cognitive, psychomotor, and affective skills.

Scott, and Webber (2015) discuss about Assessment in education, it encompasses the systematic evaluation of student learning and progress within centralized curriculum frameworks. It emphasizes both subject-specific knowledge and cross-curricular skills, while promoting continuous, teacher-led diagnostic and formative assessment practices. These assessments are often standardized nationally, ensuring accountability and informing instructional decisions. Additionally, assessment includes formal certification processes at lower levels of education. Rust (2002) adds that assessment involves judging student performance and identifying their strengths and weaknesses in specific subjects, serving as the primary method to measure learning progress.

In the classroom, assessment serves as a tool for driving change, helping students by recognize the gap between their current performance and targeted success. Oyebola (2013) emphasizes that the success of any assessment depends on its comprehensive measurement of individuals' access to information from a global perspective.

Summative, formative and diagnostic assessments are critical components of the education system. Summative assessment provides an overall performance of the student at the end of academic session. It focuses on learning outcomes and provides data on student achievements. Final exams serve as an example of summative assessment, allowing students to demonstrate their knowledge and growth in specific areas (Maki, 2002).

Diagnostic assessments serve as valuable tools for both teachers and students, offering insights into current abilities and identifying specific areas for future improvement. Ideally, these assessments should be conducted prior to the start of the course to determine a student's suitability. Based on the results, it may be necessary to consider placing the student in a lower-level course that aligns better with their existing skills (Machin & Richardson, 2016).

On the other hand, formative assessment occurs throughout the learning process, evaluating students' progress over time. It provides ongoing feedback and helps teachers identify areas for improvement. The formative assessment is valuable in supporting students' learning journeys and fostering their academic growth.

Assessments can also categorized based on their significance and impact. High-stakes assessments, such as grades, promotions, graduations, or admissions, have critical consequences for students. In contrast, low-stakes assessments are used to evaluate school and program effectiveness without severe life-changing consequences. The low-stakes assessments emphasize accountability and system development, providing positive support rather than negative pressure (Lee & Kang, 2019).

Large-Scale Assessment (LSA) refers to a type of educational evaluation that involves the systematic and standardized measurement of students' knowledge, skills, and competencies on a large, often national or international, scale. These assessments are conducted to gather data on a wide range of students, typically from different regions or countries, and offering valuable perspectives on the overall efficiency and performance of education systems.

The principal objective of LSA is to assess the performance of education systems, identify strengths and weaknesses, and inform policy decisions aimed at improving educational outcomes. These assessments are designed to measure the attainment of educational goals and learning objectives set by educational authorities at national, regional, or international levels (Schleicher, 2019).

In the context of international comparisons, International Large-Scale Assessments (ILSAs) play a crucial role in benchmarking education systems and identifying best practices across countries. These assessments include well-known international studies which allows policymakers in education to understand how their country's educational performance compares to that of other nations, providing valuable insights for policy improvements (Kirsch et al., 2013; Baird et al., 2016; Meyer & Benavot, 2013).

Regional assessments are designed for developing countries with similar economic and social conditions. Participation in regional assessments allows education ministers to gauge the performance of their education programs and students in comparison to others in the region.National assessments, whether conducted within a single country or across nations, aim to provide systemic data on learning outcomes set by the state or education department. These assessments often accompany policy-level targets, serving as a means to evaluate curriculum appropriateness and support decision-making at the policy level (Kellaghan & Postlethwaite, 2008; Benavot & Kamens, 2011).

In Pakistan, the National Assessment Wing (NAW), currently under the auspices of the Pakistan Institute of Education (PIE), conducts the National Achievement Test (NAT) to gauge student achievement levels and identify factors influencing student learning (Ahmed et al., 2019; Shehzad et al., 2018). This assessment tests serve as indispensable tools for teachers and assessment bodies. It play a vital role in the assessment process, helping educators gauge students' understanding and knowledge of the subject matter. However, it is of utmost importance to ensure that these tests possess high reliability and validity, particularly when using norm-referenced assessments for instructional purposes (Burgess, 2020). The quality of the test has a substantial influence on the overall assessment process. Therefore, it is crucial for teachers to possess a strong grasp of testing techniques, enabling them to effectively fulfill their responsibilities in the testing and evaluation process (Köhler, Robitzsch, & Hartig 2020). Mostly in these type of assessment, the multiple-choice format, are widely used in assessments at various levels today. Evaluating test items through item analysis is essential to determine their validity and improve test construction. Item analysis helps identify the difficulty level of each item and allows for adjustments to be made to enhance the quality of the test (Anastasi & Urbina, 2007). According to Rehman, Aslam and Hassan (2018), lack of item analysis can have a negative effect on entire test, it will not only cast an adverse impact on the grade of pupil but in the long run on their career and life ahead as well (Reichert, 2011). Studies carried out by Nedeau-Cato, Laughlin, and Rus (2013) and Hijji (2017) concluded that more than 85% of items have at least one or more flaws which strengthen the theory that conducting item analysis is essential for every exam to make it more purposeful and error free.

Item analysis involves the examination of four key components: Difficulty Index (DI), Discrimination Index (DI), Distractor Efficiency (DE), and the assessment of reliability through the application of the Kuder-Richardson Formula 20 (KR20). Different techniques were used to identify and maintain the quality of the achievement test. Among item analysis, one another important source to identify the test quality is teacher. Different educational principles and theories support that the teacher's views and beliefs are essential to any assessment process as they influence different teaching and learning processes. Teachers have an important role to play in identifying what is needed or what is best for their students as they focus on their day-to-day teaching. It not only affects teaching habits and classroom behavior but also, is related to student achievement (Ringstaff, & Sykes, 1985).

The dissemination of result is another important challenge related to achievement test that is how to effectively utilize assessment results to inform stakeholders and drive improvements. Sohail and Sadia (2022) reveals the serious lack of inclusion of NAT results in educational policy and classroom practices, limiting their potential to inform instructional improvement.It is crucial to leverage results from International, regional, and national assessments to inform policy decisions and empower stakeholders, including parents, teachers, and school administrators, to enhance their practices. Moreover, comprehensive dissemination of learning assessment results to the general public enables valuable feedback for policymakers.

The study aims to comprehensively evaluate the National Achievement Test (NAT) administered in Pakistan in 2016 and 2019. Specifically, it seeks to critically analyze the test's evolution and application over these years, examining its psychometric properties and identifying any areas for improvement. Additionally, the research aims to delve into teachers' perceptions regarding the administration and utility of the NAT. By understanding teachers' perspectives, the study aims to shed light on potential challenges and opportunities associated with the test's implementation. Furthermore, the research endeavors to propose a robust model for the reporting and dissemination of NAT results, aiming to enhance the effectiveness and transparency of the assessment process

1.1 Statement of the Problem

The use of Achievement tests is crucial for assessing student performance and identifying areas for curriculum improvement. Research in testing highlights the decline in the examination system, its ineffectiveness, and unreliability in diagnosing students‟ weaknesses and in assessing their abilities (Khan, 2006). However, concerns persist regarding the lack of proper test analysis and dissemination mechanisms, which affect the reliability and utilization of its results (Imran et al., 2020). Butta and Rizvi (2022) report that low performance was observed in NAT results compared to school assessments, indicating a discrepancy between national-level testing and internal evaluation practices. Despite its significance, there remains a gap in educational research regarding the comprehensive evaluation and utilization of NAT. This study aims to bridge this gap by critically analyzing the development and application of NAT, with a specific focus on the 2016 and 2019 NAT. Key objectives include identifying the difficulty and discrimination indices of the Science and Mathematics achievement tests within NAT, teacher perception regarding NAT and identifying the current practice related to the dissemination and utilization of NAT results. By proposing a phased approach for enhancing the reporting and dissemination of NAT outcomes, this research seeks to contribute to the overall enhancement of the assessment system in Pakistan.

1.2 Rationale of the Study

Assessment is a fundamental aspect of education worldwide, influencing policies and practices to improve student learning outcomes. In Pakistan, the National Assessment Wing (NAW) conducts the National Achievement Test (NAT) to assess student achievement levels and inform educational decision-making. Previous studies have highlighted the shortcomings of the NAT, emphasizing the need for evidence-based insights and practical recommendations to address these issues (Imran et al., 2020). However, concerns persist regarding the reliability, validity, and dissemination of NAT results, prompting the need for a comprehensive analysis of these areas.

This research study aims to address the gaps in current assessment practices in Pakistan, focusing on the development and application of the NAT. By critically examining the reliability and validity of the NAT, as well as proposing a model for more effective reporting and dissemination of NAT results, this study seeks to enhance the overall assessment system in Pakistan.

The rationale for this research stems from the urgency to improve the quality and effectiveness of assessment practices in Pakistan's education system. Through critical examination of the NAT's reliability and validity, along with proposing a phased approach for more effective reporting and dissemination of NAT results, this study aims to address previously identified research gaps and enhance the overall assessment system in Pakistan. By emphasizing the reasons for conducting this particular research study and identifying gaps in assessment practices, this study seeks to contribute to the existing body of knowledge and offer recommendations for improving assessment practices in Pakistan's education system.

1.3 Theoretical Framework

Classical test theory (CTT) is a theory with a simple mathematical model that shows the relationship between observed scores, actual scores, and measurement error (Mardapi, 2012). CTT is applied in estimating reliability, level of difficulty, discrimination index, distractor function, and measurement error (Retnawati, 2017). CTT is considered widely used because it does not require a large number of respondents (more than 100) and is easy to understand and apply (Argianti & Retnawati, 2020).

CTT focuses on three fundamental components: the true score, measurement error, and the observed score. The formula/equation for Classical Test Theory is as follows:

$X = T + E$

Where: X = the recorded test score T = The actual test score E = The error score

Description of Classical Test Theory:

1. The observed test score represents the actual score obtained by a test taker on a particular test or assessment. It is the score that is directly measured and recorded (Brennan, 2001).

2. The true test score reflects the individual's actual level of knowledge, ability, or trait being measured by the test. It represents the hypothetical "pure" score without any measurement error.

3. The error score represents the discrepancy between the genuine score with the recorded score. It is the variability caused by factors such as test administration conditions, guessing, or rater bias (Aksay & Ünal, 2016).

Reliability, a key aspect emphasized by CTT, assesses the consistency and stability of test scores commonly employed reliability coefficients, such as Cronbach's alpha and test-retest correlations are usually used to measure the reliability of test scores. Additionally, CTT provides methods for item analysis, allowing for the evaluation of individual test items (Aksay & Ünal, 2016). Item analysis facilitates the Identifying problematic items, evaluating item difficulty and discrimination, and assessing the overall quality of items.

While CTT primarily focuses on reliability, the consideration of test validity is crucial. Validity pertains to the extent to which a test accurately evaluates the intended construct (Ayre & Scally, 2014). Although CTT does not directly address validity, it serves as a foundational framework for subsequent validity analyses.

CTT is based on several assumptions, including the parallel forms assumption, tau-equivalence assumption, and local independence assumption. The parallel forms assumption suggests that different but equivalent test forms yield similar results. The tau-equivalence assumption posits that items with the same true score exhibit equivalent measurement properties. The local independence assumption assumes that items are independent of each other when controlling for the underlying construct. While CTT offers valuable insights into test analysis, it has limitations. CTT assumes a linear relationship between observed and true scores and does not account for specific item characteristics or individual examinee characteristics (Brennan, 2001). Furthermore, CTT assumes that measurement error is random and unrelated to the true score.

Classical Test Theory provides a robust framework for assessing and gauging the reliability and validity of test scores in educational measurement and psychometrics. Despite its limitations, CTT's emphasis on the true score, measurement error, and observed score provides a solid foundation for assessing the quality of test scores.

To supplement the statistical evaluation offered by Classical Test Theory, this study also incorporates Nagy’s Large-Scale Assessment (LSA) Framework to explore the practical utility and perceived validity of NAT in classroom settings. Nagy’s LSA framework extends beyond psychometric analysis by emphasizing the educators’ perspective—how teachers perceive, interpret, and utilize large-scale assessment results in practice. The framework considers the alignment of assessments with curriculum goals, fairness and transparency of test items, and the instructional relevance of assessment feedback. It supports evaluating whether such assessments provide meaningful and actionable insights for improving teaching and learning.

By combining CTT’s theoretical rigor with Nagy’s practice-oriented lens, this research provides a holistic assessment of NAT—analyzing both its psychometric properties and the way it functions as a classroom evaluation tool. The dual-framework approach strengthens the study’s capacity to make informed recommendations regarding test design, interpretation, and application within the educational system.

1.4 Conceptual Framework

The conceptual framework of this study is divided in two phases first the development phase in which CTT was used and the second application phase which consist of Nagy's framework for large scale assessment (LSA), which comprises four critical assessment tasks: accountability, gatekeeping, instructional identification, and monitoring of student progress (Nagy, 2000; DeLuca, Klinger, & Miller, 2008).

Originally, LSAs primarily served as gatekeeping mechanisms, influencing admission or graduation decisions (Nagy, 2000). However, the role of LSAs has evolved to encompass broader responsibilities, including system-wide accountability (Earl, 2010), feedback for teaching diagnostics, and monitoring student progress.

In this study, Nagy's framework serves as the foundation to address the specific research objective, which focuses on obtaining the objective and accurate information about the application of the National Achievement Test (NAT) by considering the four functions of Large Scale Assessment (LSA), namely accountability, gatekeeping, instructional identification, and monitoring of student progress through teacher's perception. By adopting Nagy's framework, this study aims to assess the perceptions, experiences, and challenges faced by teachers in the utilization of National Achievement Test (NAT) results and identify effective strategies for applying NAT data to improve student achievement.

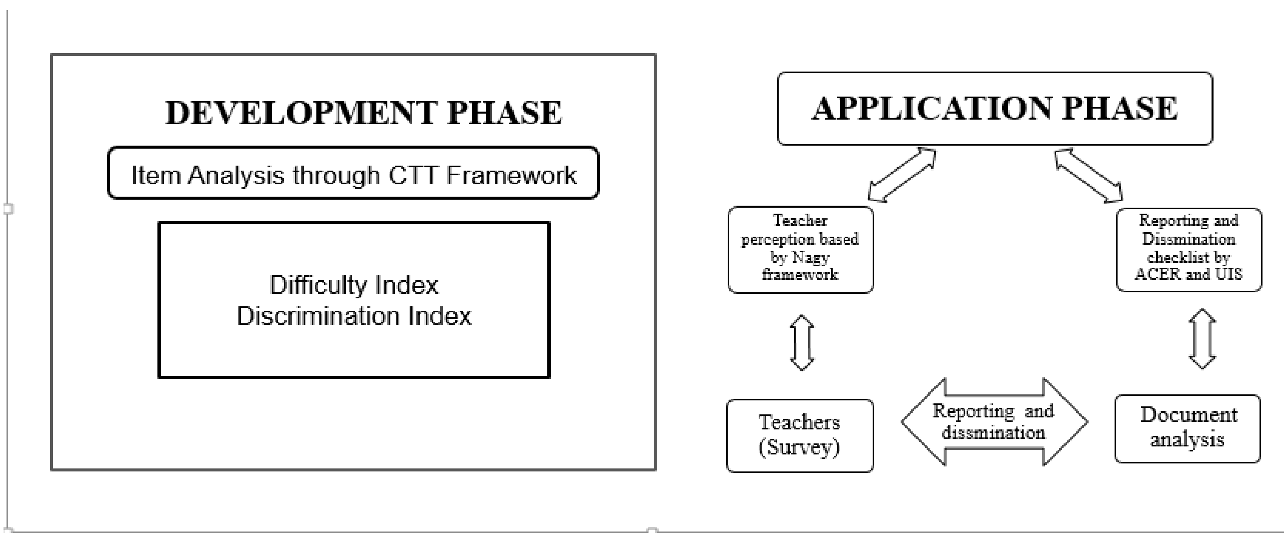


Fig 1.1 Conceptual Framework for Large Scale Assessment

The integration of Classical Test Theory (CTT) within the conceptual framework enriches the analysis of large scale assessment practices, particularly concerning the National Achievement Test (NAT). CTT principles provide a foundational framework for understanding the reliability and validity of NAT scores, aligning with the assessment tasks outlined in Nagy's framework for Large Scale Assessment (LSA). Through the assessment of reliability and validity, guided by CTT principles, this study aims to ensure the consistency and accuracy of NAT scores, thereby contributing to the accountability, gatekeeping, instructional identification, and monitoring of student progress functions within the broader objectives of LSA.

1.5 Objective of this Study

The objectives of this study were:

1. To assess the psychometric characteristics of the National Achievement Test (NAT) and assess its effectiveness as an assessment tool by using the Classical Test Theory (CTT) Framework
2. To obtain objective and accurate information about the application of the National Achievement Test (NAT) by considering the four functions of Large Scale Assessment (LSA), namely accountability, gatekeeping, instructional identification, and monitoring of student progress through teacher’s perception.
3. To assess the perception of teachers in the utilization of National Achievement Test (NAT) results and identify effective strategies for applying NAT data to improve student achievement.
4. To propose a Phased approach for effectively disseminating NAT results to various stakeholders, facilitating their access to and understanding of the assessment outcomes.

1.6 Research Question (s)

The study aims to answer the following research questions

Research Question 1:How appropriate are the difficulty indices and discrimination indices of the test items in the NAT 2016 and 2019?

Research Question 2: How do teachers perceive the application of the National Achievement Test (NAT) in relation to the four functions of Large-Scale Assessment (LSA) - accountability, gatekeeping, instructional identification, and monitoring of student progress?

Research Question 3: How do teachers perceive, utilize, and experience challenges in applying National Achievement Test (NAT) results for improving instructional practices, student learning, and classroom effectiveness?

Research Question 4: What mechanisms exist for the dissemination of NAT results, and how should results be effectively reported to different stakeholders?

1.7 Significance of Study

Teachers' views and convictions about assessment hold significance importance in a ways that it influencing diverse aspects of the teaching and learning dynamics. Numerous educational principles and theories highlight the significance of teacher perception, as it directly influences instructional practices, classroom behavior, and ultimately student achievement.

It is essential to acknowledge that teachers are in a unique position to observe and understand the needs of their students on a daily basis. Their insights and understanding of the assessment process can significantly impact instructional decisions, classroom strategies, and student learning outcomes. The beliefs and perceptions teachers hold about assessment can shape their teaching approaches and strategies, enabling them to effectively address the diverse learning needs of their students.

By examining teachers' perceptions of the National Achievement Test (NAT) and its application, this study aims to shed light on the crucial role of teacher beliefs in the context of assessment. Understanding the perspectives and attitudes of teachers towards NAT can provide valuable insights into the effectiveness and relevance of the test in the Pakistani education system. Moreover, it can inform educational policymakers and stakeholders about the potential areas for improvement in the implementation and dissemination of NAT. The significance of this study lies in its potential to bridge the gap between assessment practices and teacher perceptions. By acknowledging and considering teachers' viewpoints, educational institutions can work towards developing assessment strategies that align with the needs and aspirations of teachers and students alike. This research aims to contribute to the existing body of knowledge by exploring the relationship between teacher perception, assessment practices, and student achievement in the context of NAT in Pakistan.

In addition to teacher perceptions, item analysis is another crucial tool for evaluating the quality of assessments. It involves analyzing both the answers and the questions themselves to assess the overall test quality. In this study, item analysis techniques are employed to critically analyze the NAT 2016 and 2019 achievement test. The findings from this analysis will contribute to the improvement of test development and instrument design, creating a stronger connection between the assessment system and other educational departments such as the curriculum wing, textbook development centers, teacher professional development centers, and examination bodies. This, in turn, will enable more informed and strategic utilization of assessment results to enhance student learning.

The significance of this study can be highlighted in the following points:

Firstly, it aims to provide valuable feedback to policymakers and assessment experts, enabling them to make necessary improvements to the assessment system. By identifying strengths and weaknesses through item analysis, measures can be proposed to enhance examination and evaluation practices, monitoring and supervision protocols, an overall accountability of schools, teachers' effectiveness, and students' progress in achieving national educational goals.

Secondly, the study seeks to establish a solid foundation for improving student learning outcomes by emphasizing student achievement and consolidating their intellectual abilities. By analyzing student performance within different instructional contexts and considering various factors that may impact achievement, the study can provide insights into learning achievements and inform strategies for improvement. Moreover, the assessment of learning achievements through item analysis allows for the evaluation of teacher performance and the overall effectiveness of educational institutions. This study aims to contribute to the judgment of student learning achievements and provide a basis for assessing the performance of teachers and educational institutes. Overall, this research holds significance in informing educational policies, promoting accountability, and enhancing the quality of assessment practices. By evaluating and improving the assessment system, it has the potential to positively impact student learning and contribute to the achievement of national educational goals.

1.8 Research Design

This study adopts an observational design, using retrospective data from the 2016 and 2019 National Achievement Test (NAT) conducted in public and private schools across Pakistan. It focuses on Grade 4 and 8 Mathematics and Science to assess the consistency and effectiveness of the NAT in measuring learning outcomes. A correlational analysis is also included to explore the relationship between NAT scores and teacher perceptions. Data were collected through a mixed-format survey (open- and closed-ended questions) to capture both quantitative and qualitative insights. Additionally, a checklist and document review were used to evaluate how NAT results are reported, ensuring a comprehensive analysis.

1.9 Research Methodology

Guided by a pragmatic paradigm, the study emphasizes the practical application of findings within real-world educational settings. An ex-post facto design was used to analyze NAT data from 2016 and 2019 without researcher manipulation. Classical Test Theory (CTT) informed the quantitative phase, examining psychometric attributes such as item difficulty and discrimination indices to evaluate test validity and reliability. Qualitative elements focused on the perceived impact of NAT on teachers, students, and policymakers. The survey began with structured (closed-ended) questions to gather measurable insights on teacher perceptions, followed by open-ended items to explore deeper views on the test's influence on teaching practices. Document analysis further supported the investigation by examining how NAT outcomes are communicated and utilized in educational planning. This integrated methodology offers a well-rounded view of NAT's role in shaping instructional and policy decisions.

1.10 Population

National achievement test was administered in all province and areas of Pakistan. Due to limited resources and time constraints. The researcher select the population of Islamabad and Azad Jammu and Kashmir (AJK). The population was consist of teacher and students of Mathematics of Grade 4th & 8^{th} and Science of Grade 4th & 8^{th} and of public school which was the part of previous NAT 2016 and 2019 in Islamabad and AJK. The population was divided urban and rural wise.

Table 1.1:

The Total Numbers of Teacher Participated in NAT 2016

Grade	Area	Subject	Location	Gender	Teacher
Grade 4	AJK		Rural:37	Female:27	44
		Mathematics	Urban:07	Male:17	
	ICT		Rural:09	Female:18	28
			Urban:19	Male:12	
Grade 8	AJK	Science	Rural:40	Female:30	48
			Urban:08	Male:18	
	ICT		Rural:11	Female:24	34
			Urban:23	Male:10	

Table 1.2:

The Total Numbers of Teacher Participated in NAT 2019

Grade	Area	Subject	Location	Gender	Teacher
Grade 4	AJK	Science	Rural:10 Urban:04	Female:07 Male:07	14
	ICT		Rural:02 Urban:06	Female:08 Male:08	16
Grade 8	AJK	Mathematics	Rural:09 Urban:05	Female:07 Male:07	14
	ICT		Rural:02 Urban:06	Female:04 Male:04	08

Table 1.3

Numbers of Students Participated In NAT 2016

Grade	Area	Subject	No of Students	Total School
Grade 4	AJK		880	59
	ICT	Mathematics	560	28
Grade 8	AJK	Science	960	48
	ICT		680	34

Table 1.4

Numbers of Students Participated In NAT 2019

Grade	Area	Subject	Location	No of Students	Total School
Grade 4	AJK	Mathematics	Rural:10 Urban:04	78	14
	ICT		Rural:02 Urban:06	516	08
Grade 8	AJK	Science	Rural:10 Urban:04	113	14
	ICT		Rural:02 Urban:06	242	08

1.11 Sampling

Regarding the research study's sample size, there seems to be a lack of consensus in the literature regarding the methodology for determining the suitable number of participants (Thabane, 2010). However, Hertzog (2008) emphasized that if the pilot test is to evaluate the internal consistency of the instrument small sample size is inadequate. Keeping in view this researcher took the 40 teacher out of 52 teachers from NAT 2019 and 60 teacher out of 158 from NAT 2016 & 2019 as a sample. As far as the student is concerned the overall population is considered for this study. One important thing which was in the consideration within the selection of the schools is that it was those schools that participated in the National Achievement Test 2016 and 2019.

Table 1.5

Sample Size Selected from NAT 2016 & 2019

		2016		2019	
Grade	**Area**	**Students**	**Teachers**	Students	**Teachers**
Grade 4	AJK	78	20	78	11
	ICT	516	10	516	13
Grade 8	AJK	113	20	113	11
	ICT	242	10	242	05

1.12 Methods and Tools of Data Collection

The aim of this research was to analyze the development of National Achievement Test and to identify teacher perceptions which they hold regarding NAT and its application. Teachers 'perceptions of NAT were based on Nagy's (2000) framework for LSAs which based on the four functions for LSA: accountability, gatekeeping, instructional identification and monitoring.

To analyze the development of National Achievement Test the item analysis method was adopted.

1.12.1 Item analysis

The Achievement Test 2016 & 2019 of Science and Mathematics **(Annex-1)** were used as a research instrument to identify the psychometric properties of the test items. The reason why this achievement test is selected for this study is that it the only national achievement test conducted in Pakistan. Therefore, the findings will be applicable at the national level. The achievement test which items was used in this study contained objective items, in which items related to science and mathematics. For item analysis, the correct response was coded as 1 and 0 as the wrong response.

1.12.2 Teacher Perception regarding NAT

To identify the teacher perception the standardized tool used **(Annex-II)**. The instrument contained three sections: First section related to demographics information (6 items), Second sections deals with NAT perception related items (24 items) which further subdivided into the following accountability (9 items), monitoring (5), Instructional methodology (6) and gatekeeping related item(4 items).

Figure 1.2 ***Items Division***

Dimension	Items
Accountability	9
Monitoring	5
Instructional methodology &	6
Gatekeeping	4
Total number of items	24
Rating scale	Ordinal frequency Five-point

The third section comprised of 10 Open-ended questions The questionnaire was distributed among a subject teacher of Math's, of Grade 4th and Grade 8th Science which were the part of previous NAT 2016 & 2019 to identify their perception about the national achievement test and its application in their instructional decision-making.

1.12.3 Reporting and Dissemination of NAT

In this endeavor to conduct a thorough analysis of the reporting and dissemination process of Large-Scale Assessments (LSA), the researcher adopted a comprehensive checklist (ACER-GEM & UIS, 2017) **(Annex III)** as a systematic tool. This checklist is instrumental in guiding our evaluation and assessment of pertinent policy and practice issues, as well as the diverse information needs of stakeholders engaged in the LSA program.

1.13 Procedure of Data Collection

Following ethical research standards outlined by Bryman (2012), the researcher sought formal approval to ensure participant protection, informed consent, and data privacy. A detailed request titled *"Request for Research Data from NEAS"* was submitted to the Ministry of Federal Education and Professional Training (MOFE&PT), outlining research objectives, methodology, tools, and data protocols. Approval was granted by senior authorities (Annexure IV). Additionally, permission to use the teacher perception tool was obtained via email from its original author **(Annexure V).**

1.14 Data Analysis

The data analysis in this study involved utilizing the achievement tests from 2016 and 2019 in Science and Mathematics as research instruments to evaluate the psychometric characteristics of the test items, such as item difficulty and item discrimination. The data analysis in this study involved the utilization of SPSS. Through this psychometric analysis**,** items that appeared ambiguous were identified. The use of psychometric analysis is justified as it provides an objective and statistical foundation to assess the quality, reliability, and validity of test items. By analyzing how test items function in terms of difficulty and discrimination, the study identifies strengths and weaknesses in the assessment design. This ensures that future National Achievement Tests are more accurately aligned with student abilities, curriculum standards, and national education goals. Psychometric techniques allow for evidence-based decision-making, essential for improving the effectiveness and fairness of large-scale assessments. Additionally, the analysis of the teacher perception tool was conducted using the mean and percentage. This allowed for a comprehensive examination of the teachers' perceptions regarding the National Achievement Test. For reporting and dissemination of NAT, the checklist was used which was analyzed through document analysis.

1.15 Limitations

This study is limited to schools in Islamabad and AJK, which may restrict the generalizability of the findings. The selected teacher and student samples may not fully reflect the broader population involved in NAT. Due to time constraints, only data from 2016 and 2019 were analyzed, excluding more recent trends. The focus was limited to Grade 4 and 8 Mathematics and Science, excluding other grades and subjects.

Using an ex post facto approach meant no control over independent variables, limiting the ability to establish causality. This approach is justified because the events and data under study had already occurred prior to the research. The ex post facto design is appropriate for educational evaluation studies where variables cannot be manipulated due to ethical, practical, or procedural reasons. It allows the researcher to analyze real-world data and outcomes, such as test performance and teacher feedback, without experimental intervention. Although it limits causal inference, it enables the investigation of associations and patterns in a natural educational context. The reliance on descriptive data analysis also restricts deeper exploration of factors influencing NAT performance.

1.16 Delimitations

The study is confined to NAT assessments from 2016 and 2019, excluding other national or international evaluations. It centers on teacher perceptions, not considering input from students, parents, or policymakers. Item analysis was the sole method used to evaluate test quality. Only schools that participated in NAT 2016 and 2019 were included. An adapted instrument by Dr. Don A. Klinger, originally developed for a different context, was used to assess perceptions of large-scale assessment.

1.17 Operational Definitions

In the context of this research, the following operational definitions are provided for the terms "development," "application," and "achievement test":

1.17.1 Development: In the context of this study, "development" is operationally defined as the systematic process of creating educational materials, assessments, and resources in alignment with specific learning objectives or standards. This process includes activities such as curriculum design, item writing, pilot testing, and refinement based on feedback and evaluation. It aims to produce instructional materials and assessments that facilitate student learning and measure their progress (Goodwin & Leech, 2003).

1.17.2 Application: Within the framework of this research, "application" is operationally defined as the practical implementation and utilization of educational materials, strategies, and assessments in real-world educational settings. It involves integrating instructional resources and assessment tools into teaching practices to enable students to apply acquired knowledge and skills in authentic contexts. Application also encompasses the effective use of pedagogical approaches, instructional technologies, and assessment strategies to promote meaningful learning experiences and transfer of knowledge to real-life situations.

1.17.3 Achievement Test: In this study, an **achievement test** is operationally defined as a standardized tool used to measure the extent to which students have acquired the knowledge, skills, and competencies outlined in a prescribed curriculum or set of learning outcomes. It reflects students' actual learning performance in specific subject areas and is typically administered under controlled conditions to ensure reliability and fairness. Achievement tests are designed to provide objective, quantifiable data that can be used to evaluate educational effectiveness, monitor student progress, and guide policy or instructional decisions (adapted from the National Council on Measurement in Education [NCME], 2014).

1.17.4 Assessment: In the context of this study, **assessment** refers to the systematic process of gathering, analyzing, and interpreting evidence about student learning from a variety of sources and methods. It includes both formal and informal strategies aimed at understanding what students know, understand, and are able to do. The purpose of assessment is to evaluate learning progress, provide feedback for improvement, support instructional planning, and enhance the overall learning experience. It emphasizes the development of each individual learner by identifying strengths, addressing gaps, and informing future teaching and learning strategies.

1.17.5 Large Scale Assessment: Large-scale assessment are system-level assessment conducted in Pakistan by Pakistan Institute of Education, that provide a snapshot of learning achievement for a given group of learners in a given year and in a limited number of domains.

CHAPTER SUMMARY

This chapter provides an overview of the research topic, establishes the research problem and objectives, and introduces the theoretical and conceptual frameworks that guide the study. It sets the foundation for the subsequent chapters, which will delve deeper into the analysis and findings related to the development and application of the NAT in Pakistan.

.

CHAPTER: 2

LITERATURE REVIEW

Various studies and articles were reviewed by the researcher to gain insight into a particular field of study. This research was meticulously designed to review the development and application of the large scale assessment in Pakistan.

2.1 Assessment

Assessment holds a key position in teaching-learning process, it not only enables the teachers to judge students' performance but also make them aware of the strong and weak areas of their students. Thus, they are in a better position to guide them as to how to overcome their weak areas and further polish their strong areas, along with keeping them motivated to perform their best for the future (Harlen, 2006). Assessment, as described by Tosuncuoglu (2018), is a comprehensive and extended process that involves gathering information and data about student progress. It plays a crucial role in the learning process, enabling teachers to understand the extent to which students can comprehend the materials and solve problems presented through tests and assignments. Consequently, assessment stands as an essential procedure for effective teaching and learning. The assessment provides an overview of the whole learning process and makes you aware of how much students have learned a specific topic and which part is still confusing them (Alkharusi, 2008).

Student assessment plays a vital role in the teaching and learning process, which aims to bring about positive changes in student behaviour. Success tests serve as valuable tools for measuring the extent of learning and behavioral changes in students. While some teachers rely on formal assessment instruments, others base their evaluations on subjective impressions formed through daily interactions with students. Tests are widely used across the education system, with teacher-developed tests being commonly implemented in classrooms to address specific student needs. In contrast, standardized tests are employed for purposes such as certification, quality assurance, and benchmarking at regional or national levels. Reliance on test scores has become prevalent in educational institutions, influencing decision-making and setting standards. The importance of test quality is paramount, focusing on various aspects such as test design, analytical techniques, and score interpretation. Quality test design considers elements such as format, length, administrative procedures, validity, and assessment methods. Ensuring the effectiveness and efficiency of tests in measuring ability is essential because the information gathered from achievement tests determines educational development and educational efforts (Olatunji & Owolabi, 2009).

Summative and formative assessments are two important forms of assessment commonly used in our education system. Summative assessment is used at the end of the course and the main aim of this assessment is to assess the overall learning of the student. Another objective of this type of assessment is to produce a measure that “sums up” student learning. Its style is all-encompassing, and it is focused on learning outcomes. Moreover, it also facilitates providing data and evidence about the pattern of student’s achievements and allowing them to learn from their mistake and work hard in improving their weak areas (Maki, 2002). Final exams are one of the examples of summative assessment.

Formative assessment is another type of assessment where student's performance is evaluated step by step and continues throughout the learning process. The major objective of this kind of assessment is to identify the level of achievement and to use the same in designing and modifying the learning process so that students can learn to their maximum capability. While interpreting students' level of achievement with formative assessment and, instructors help students by sharing the results with them to highlights their strengths and weakness and to guide them that how they improve further in their remaining studies" (Maki, 2002). Formative assessment is rooted in classroom as a diagnostic approach it is considered as heart of the learning process as it create a more active learning environment and it support student to learn to being an independent learners. An independent learner is critical to building a resilient education system and is a crucial area of development under the New Deal Document (2018), Under the title "Developing Skills and Knowledge: Every child, irrespective any discrimination will learn appropriate skills and knowledge that help him or her prepare for future including the ability to learn and then apply learning in changing circumstances.

According to the Times Educational Supplement (2019), formative assessment is described as an approach to assessment known as "assessment for learning." In this method, assessment occurs during the learning process itself, as opposed to being conducted solely at the conclusion of a topic or a series of lessons, as is the case with summative assessment. Formative assessment does not involve assigning grades or making comparisons among learners; instead, it focuses on evaluating individual progress in relation to the curriculum. Booth (2017), comments that giving marks can be counter-productive as learners focus on the outcome and compare themselves with others rather than consider how they can improve. The formative assessment goes beyond the collection of evidence. The key essence of it entails educators using the information they gather to enhance future teaching and the overall learning experience. This view also supported by Black and William (2009) states that "Formative classroom practice involves gathering and interpreting evidence of student achievement, guiding teachers, students, or peers to make informed decisions for more effective instruction, surpassing decisions made without such evidence".

2.2 Different Assessment Models

Over the past century, influenced by various assessment movements, numerous assessment models have surfaced, with scholars like Serafini (2001) putting forth frameworks to delineate unique educational research paradigms (Giordano, 2005).

Liang (2006), examined the prevalence of the scientific assessment model, which strives to offer an impartial gauge of student accomplishment. They outlined two assessment models: (1) The Scientific Measurement Model and (2) The Judgmental Model. These models differ in their metaphysical assumptions about humans, epistemological perspectives, and assumptions about the measurement process, assessment principles, and assessment practices.

The Scientific Measurement Model primarily quantifies achievement, focusing on structured problems with defined answers, single correct responses, and controlled test conditions. In contrast, the Judgmental Model concentrates on cognitive processes, involving unstructured problems, varied interpretations, diverse types of work (e.g., group projects, portfolios), and collaboration between students and teachers in assessment task design.

Serafini (2001) introduced three assessment paradigms: (1) Measurement-Based Assessment, (2) Procedure-Based Assessment, and (3) Inquiry-Based Assessment. Measurement-Based

Assessment emphasizes standardized testing, fixed scoring approaches, and multiple-choice items to assess products of achievement. 'Assessment as Procedure' extends this model by incorporating various data collection methods. The 'Assessment as Inquiry' model, closely linked to performance assessments, aims to improve individual learning, assess processes, and actively involves students and teachers in assessment development. However, it is more resource-intensive due to its focus on learning processes. While these models provide tangible elements of assessment, they offer limited accounts of student learning, their differences, historical context, and practical application.

Numerous fundamental views emerge from the these models (1) Product-Oriented, which shares similarities with measurement and procedure models, and (2) Process-Oriented, bearing a resemblance to the inquiry model. These perspectives are not mutually exclusive but rather contingent on the objectives of the assessment. Test developers should explicitly define which perspective elements underlie their assessments and how they align with their goals.

Assessment goals profoundly influence assessment frameworks (e.g., development, administration, reporting). Common practices and adherence to guidelines, such as the Principles for Fair Student Assessment Practices in Canadian Education (1993) play a crucial role. This includes clearly articulating the assessment's purpose, theoretical basis, content, procedure, and score derivation to ensure fairness and transparency.

2.3 Large Scale Assessment

The term Large-scale assessment (LSA) originates from the concept of general assessment and generally refers to programs conducted by different bodies e,g public and private in which the evaluation of an individual's abilities and performance at a specific age and educational level was conducted through the utilization of tests or similar assessment tools. The results of these assessment are used to inform ideas on learning quality, program, and resource allocation. With the growing practice of using test to improve educational outcomes, has added accountability as a major consideration. This led to a discourse founded in the early 20th century that included student evaluation and decisions about school performance (Nagy, 2000). Popham (2001), advocates for the regular gathering of student achievement data on an annual basis. This practice enables the general public and education policymakers to assess whether teachers are meeting performance expectations. Lewis (2000), describes LSA functions as a "feedback loop" by integrating fresh, dependable, and precise data into the policy-making process to enhance programs and enable the monitoring and development of measures. For Almond (2002), LSA assistance paves a path for decision-making that promotes development, efficiency, and accountability. The presence of LSA results highlights the flaws and strengths of the education system and specific teaching methods. Today, these assessment programs adapt and make changes to policy curriculum and academic performance, especially in the area of data-driven decision-making (Clarke, 2012). Student assessment in these types of assessments is used for a variety of purposes, such as the comparison of educational systems, facilitating accountability measures, and evaluating achievement at a broad curriculum level (The Council of Ministers of Education, Canada [CMEC], 2008).

In education, now a days large-scale assessments are highly discussed phenomena due to their outstanding outcomes and influence on educational policies in many jurisdictions worldwide. Researchers identify that assessment results can greatly help policymakers in making informed policies that significantly impact the school system (Johansson, 2016). Large-scale educational assessments, according to de Lange (2007), involve administering tests to significant numbers of individuals for various educational purposes, including placement, graduation, admissions, and evaluating school accountability. These assessments are not limited to any particular region but have a global presence, with countries like the United Kingdom, Germany, France, Northern European nations, and Japan employing them. Kloosterman and Burkhardt (2017), mandate the creation of state-specific large-scale assessments for primary school students, particularly in reading and mathematics. The primary objectives of these assessments are to gauge progress and ensure that all students attain a specified proficiency level. While large-scale assessments face scrutiny and critique, their importance in accountability reporting remains significant. They gain popularity and see increasing use, especially in many countries. Despite debates surrounding their effectiveness, large-scale assessments continue to be integral to educational evaluation and decision-making.

2.3.1 History of LSAs

In the early 20th century, scientific inquiry played a vital role in science, shaping experiments with rigorous steps to ensure unbiased outcomes. This influenced the mental measurement movement, where psychologists sought ways to quantify intelligence, notably gaining traction after World War I (Cherry, 2020). In the1980s and 1990s era saw a more focused approach by the World Bank towards importance of assessment and comparison in our education system. "Examine assess and compare" this was their slogan to highlights the above mentioned aspect. They gave a simple strategy that is to examine your system, assess what is delivered and then compare your results with other countries. This trend of assessment and comparison though started back in 1960 but got a strong hold 1990s. Comparative studies of students achievements the outcome different testing methods etc. became a field of interest, it all started when a non-government organization IEA created a team of experts and scholars, with the task of designing standardized assessments of student learning (Keeves, 1995). With this the first wave of cross-national comparisons of subject-specific learning outcomes started. However this remained mostly in developed countries and only few developing countries contributed in these types of evaluations (Benavot &Tanner, 2007). The evolution of educational assessments has led to three distinct types: international assessments for cross-national comparisons, national assessments, and regional assessments. These assessments, varying in quality and scope, have become increasingly prevalent since the mid-1990s, particularly in middle and low income countries e.g. LLECE in Latin America, SACMEQ in Anglophone Africa, and PASEC in Francophone Africa (Benavot & Tanner, 2007).

2.3.2 Low and High Stake Assessment

When considering the impact of LSA assessment on students, it is important to analyze both high-stakes and low-stakes testing Tests often serve as motivators for learners, as they are associated with positive outcomes such as course placements, job opportunities, certificates, grades, and recognition from teachers (Green, 2013). However, if students perceive no tangible benefits or consequences from low-stakes tests, their motivation to perform well or even participate in the exams may diminish. To address this issue, Finney (2011), suggested increasing the consequences of low-stakes tests to enhance student effort and thereby improve the validity of scores. Examining student motivation within low-stakes testing contexts is crucial for ensuring score validity, as the lack of student motivation may result in scores that do not accurately reflect their true abilities (Swerdzewski et al., 2011).

However, ethical concerns arise when stakeholders increase the importance of tests or offer rewards for passing, especially when students perceive the tests as low-stakes while teachers view them as high-stakes due to implications for course funding (Finn, 2015). Additionally, there can be negative repercussions associated with increasing the significance of tests. Stiggins, 2010 argues that while high-stakes tests may be motivating and challenging for confident and capable students, those who believe success is beyond their reach often experience demotivation and failure.

Furthermore, if students repeatedly fail exams or face multiple resist, they may become demotivated and even abandon their language learning altogether. This can occur in both low-stakes and high-stakes testing situations since the perceived importance of the exams is determined by the stakeholders themselves. It is crucial to consider test-takers' perceptions of the exams' significance, not only regarding scores but also in terms of long-term effects. Weir (2005) introduces the concept of consequential validity within a socio-cognitive framework, emphasizing the need to incorporate test-takers' voices and perspectives into test development and evaluation.

Shohamy (2001) argues for the importance of listening to test-takers' experiences and taking them into account when designing tests. One illustrative example involves a man who, after failing a secondary school test, faced rejection from his family, disengaged from education, and turned to drugs. While this is an isolated case, it highlights the potential long-term impact of failure. Acknowledging the potential consequences of exam results is essential in understanding their true significance.

To clarify the concept of low-stakes assessments, researcher refer to Gambell and Hunter (2004), who characterize such assessments as those that offer “Provincial, national, or international trends and profiles, but no individual marks for promotion or purposes of graduation”. Low ratings are usually understood as ratings that do not directly affect teachers in terms of consequences such as sanctions or salary recommendations. They are also generally perceived as putting less pressure on both teachers and students (Klinger & Rogers, 2011).

The influence of extensive-scale assessments on teaching methods is a complex matter marked by various dimensions (Cimbricz, 2002). Generally, research suggests that these assessments can lead to several consequences, including curriculum narrowing, reduced instructional time for test preparation, a shift towards more test-focused teaching practices, and an increased potential for academic dishonesty (Klinger & Rogers, 2011). Large-scale assessments present a dual perspective, as they are often criticized for promoting "teaching to the test," diverting valuable time away from non-tested subjects (Volante, 2004), and providing information with limited utility for enhancing classroom practices (Stiggins, 2002). However, they also carry the potential to recognize effective teaching approaches and skilled educators, stimulate teachers' reflective and critical thinking in instructional planning (Goldberg & Roswell, 2000), and improve teachers' assessment and instructional strategies, particularly when they participate in the scoring process (Volante, 2006). Some teachers have reported allocating a disproportionate amount of time to the subjects covered by these tests and adopting inappropriate test-focused preparation methods. These test preparation practices tend to emphasize teaching specific skills necessary for the exam, sometimes at the detriment of the hands-on, exploratory, and investigative aspects of the curriculum, as exemplified in Grade 9 mathematics (Volante, 2007).

2.3.3 Perception about Large Scale Assessment

Perception, according to Ryan (2008), refers to the attitude or understanding derived from observations or thoughts. It involves the ability to notice or discern things that go unnoticed by most individuals. This definition suggests that perception plays a significant role in shaping individual and societal differences in attitudes and understandings. Rather than being inherent in objects, perception is the interpretation that our minds give to sensory or brain information. In essence, the perception of a group of people represents their reality, encompassing their understanding and attitude toward a particular subject of study. Therefore, any innovative program should be based on the people's perception (Ryan, 2008).

Sensation and perception are closely connected yet serve distinct roles in our understanding of the world. Sensation can be defined as the process of sensing our surrounding though touch, taste, sight smell and sound. After sensing the information's transmitted to the brain from where the role of perception started. Perception is the process in which we interpret these sensations and make the sense of our surroundings (Heffner, 2014)

According to Mcleod (2007), perception heavily relies on prior knowledge and past experiences. Sensation can be described as the passive process of receiving information from the external world and transmitting it to the brain. This information remains intact and can be reproduced. Sensation does not require conscious engagement in a "sensing" process. On the other hand, perception is an active process that involves selecting, organizing, and interpreting the sensory information received by the brain. It is an active and pattern-seeking process closely associated with thinking (Moscovich, 2006).

Generally, LSA assessments conducted for numerous objectives or purposes for example positioning, admissions, and accountability of the school. These assessments in numerous countries are quite often identified with promotion examinations based on curriculum. There are numerous other purpose of LSA, but central to them is three: accountability, choice, and comparison. Responsive LSAs are assessments aimed at holding academic professionals responsible for their students' educational progress.

2.3.3.1 Arguments against Large-Scale Assessment: There are different perceptions about large scale assessment. The main arguments against large-scale assessment are that this testing:

• endorses policy-decisions that have not been properly examined and are based upon those procedures which have important psychometric limitations (Popham, 2001).

• It creates the feeling of lower self-efficiency by considering teacher responsible for poor results (American Educational Research Association, 2000).

• exacerbates the disparity among students by overlooking significant factors recognized to influence student performance, such as their socioeconomic status, language proficiency, and both physical and emotional well-being (Valencia & Villarreal, 2003).

• penalizing creative thinking, innovation, and intellectual effort diminishes the educational standard (Anderson & Postl, 2001).

• promoting a "teaching to the test" approach restricts and distorts the curriculum (Smith & Fey, 2000).

• provides data that is of relatively used little in improving the classroom (Stiggins, 2001).

• affects the children's confidence level which leads to learner disengagement (Meaghan & Casas, 2001; Miller, & Tovey, 1996).

2.3.3.2 Supporter of Large-Scale Assessment

Advocates for large-scale assessments contend that those who critique them have not presented empirical data to substantiate their perspective regarding the drawback of large-scale assessment (Covaleskie, 2002). They highlighted that large-scale test or assessment have positive outcomes for all the stakeholder. Testing has been found through research to incentivize increased effort from students, teachers, and educational institutions (Anderson, 1990). Through this, they perform better in different assessments such as international, regional, or national, etc. (Bishop, 2001). Some other points in favor of large-scale assessment are given below:

- Highlighting the best teaching practices and generate proficient teacher (Cizek, 2001),
- Develops the reflective and critical thinking among teacher especially while they are planning instruction (Goldberg & Roswell, 2000),
- Improve educators' understanding of testing and related concerns (Earl & Torrance, 2000).
- Enhance educators' assessment and teaching methods, especially when they are engaged in evaluating such assessments (Gambell & Hunter, 2004).
- Through action research, facilitating the basis for school improvement (Wideman, 2002),
- Accountability a key function of large scale assessments guide teachers' focus and attention towards principal who then in return divert the attention of their student and their parents. LSAs by comparison hold the attention of political leaders and leaders of the education system, who have the power to make important policy changes, which are also recognized by students and teachers.

- With the attention they can give, LSAs have great potential. The primary and desirable goal for LSAs often involves impacting all facets of education and influencing educational procedures and operations. When doing well in the LSA becomes a major goal, anything that measures what sets the agenda and learning, determines the allocation of time for both teachers and students, and has an impact on the community and parents investing. Therefore, if the LSA can weigh all the educational requirements and priorities of life, it can serve as an excellent tool.

2.4 Different Large Scale Assessment System Worldwide

In the realm of Large-Scale Assessments (LSAs), it's important to acknowledge that quantifying every facet essential to education, individual achievement, and societal well-being presents a formidable challenge. Several factors contribute to this complexity. Firstly, there exists ongoing debate regarding the evolving nature of knowledge, skills, and personal attributes that define success in a world marked by constant societal transformations (Duckworth & Yeager, 2015). Secondly, there's a dearth of valid and reliable techniques for assessing critical human qualities like innovation, entrepreneurship, and robust socio-emotional well-being on a grand scale (Zhao, 2016). Thirdly, the very context in which these assessments operate complicates the comprehensive measurement of all crucial dimensions, often pitting them against each other due to inherent conflicts or competition (Zhao, 2018). Lastly, the substantial costs involved pose a significant hurdle to the development and consistent administration of an infinite number of LSAs. Consequently, LSAs have primarily focused on evaluating skills and knowledge in specific domains, such as mathematics, language, and science. Their assessments typically center on perceptual abilities within these areas as a means to enhance educational practices. Numerous assessments carried out worldwide on a different scale.

2.4.1 International Large-Scale Assessment

International Large-Scale Assessments (ILSAs) is a relatively recent international measure of student access to education and learning (Kirsch et al. 2013). It has the potential to play a key role in driving, supporting, and shaping the education policy and transformation efforts promoted by governments at the national, regional, and local education levels (Baird et al. 2016).Trends in International Mathematics and Science Study (TIMSS) Program for International Student Achievement (PISA) and Progress in International Reading Literacy Study (PIRLS) are some examples of International large-scale assessment which is carried out to identify the student achievement in different areas. For example, TIMSS measure student achievement in mathematics and science subjects. International Large Scale Assessments (ILSAs) are evaluations of academic achievement or literacy conducted on a global scale using uniform testing instruments and methodologies. These assessments have existed for more than five decades, serve the dual purpose of comparing educational outcomes, curricula, educational strategies, and learning approaches across nations while attempting to unravel how educational policies differ internationally—such as structure, administration. Variations in student achievement and other educational outcomes are influenced by a combination of factors including the legal framework, economic considerations, and policy environments that shape both national and subnational education systems.

These ratings have evolved over the years with several notable studies:

- The First International Mathematics Study (FIMS): Conducted in 1964 by the International Association for the Evaluation of Educational Outcomes (IEA), it assessed students at age

13 as well as students at a grade level where most students were around age 13 and pre-university students. The emphasis was on mathematics.

- First International Science Study (FISS): Administered in 1970/71 by the IEA, this study assessed students aged 10 and 14, as well as students in the final year of secondary education majoring in science.
- Second International Mathematics Study (SIMS): In 1980, the IEA conducted this study targeting students in the 8th grade and final year of secondary education majoring in mathematics.
- Second International Science Study (SISS): In 1984, the IEA conducted the SISS, which assessed students in the 4th, 8th and final years of secondary education majoring in natural sciences.
- Third International Mathematics and Science Study (TIMSS): In 1995, the IEA conducted the TIMSS, which involved students in years 3, 4, 7, 8, and final secondary education who assessed both mathematics and science.
- Third International Mathematics and Science Study-Retest (TIMSS-R): This follow-up study was conducted in 1999 and focused on 8th grade students and continued to assess mathematics and science.
- Trends in International Mathematics and Science Study (TIMSS): Since 2003, TIMSS has been conducted every four years by the IEA and assesses students in Years 4 and 8 in mathematics and science subjects.
- Progress in International Reading Literacy Study (PIRLS): Started in 2001, PIRLS is conducted every five years by the IEA and assesses the reading skills of Year 4 students.

- Program for International Student Assessment (PISA): Beginning in 2000 and repeated every three years, PISA is conducted by the Organization for Economic Co-operation and Development (OECD). It assesses 15-year-old students in reading, math and science.
- International Adult Literacy Study (IALS): Conducted in 1994 by Statistics Canada, the IALS assessed individuals aged 16-65 on literacy and quantitative skills.
- Adult Literacy and Life Skills Survey (ALL): Administered by Statistics Canada in 2003, the ALL assessed the literacy and numeracy skills of individuals aged 16-65.

2.4.1.1 Implications for Learning and Teaching of International Large Scale Assessment

While (inter)national large-scale studies offer advantages such as representativeness and long-term perspectives, they also pose significant challenges. These challenges include the intricacies of data analysis and the limitations of information, such as issues related to classification, missing data, uncovered constructs, and attrition in longitudinal studies. Furthermore, these studies often encounter difficulties in effectively communicating their results to non-scientific audiences, including teachers, principals, and policymakers, due to their inherent complexity. Therefore, it becomes crucial to bridge the gap between research findings in Learning and Student Achievement (LSA) and their practical application, ensuring that the information is presented in an accessible manner without oversimplifying the results.

One notable advantage of international large-scale studies is their capacity to challenge stereotypes and address preconceived notions that national studies might struggle to overcome. For instance, these studies can examine gender differences in subjects like mathematics and science. In many Western contexts, there tends to be a stereotype favoring boys in these fields. However, international studies, like PISA, have revealed that several countries exhibit mathematics scores favoring girls and nearly equal distribution in science scores (OECD, 2015). This international perspective allows for an analysis of how various phenomena evolve over time and across different countries, offering valuable insights into educational system changes. It is worth noting that education fundamentally aims at individual development, making longitudinal follow-up surveys and cohort analyses particularly beneficial for understanding the impacts on learning and teaching. IEA developed a curriculum model that encompasses three components: intended curriculum, implemented curriculum, and attained curriculum. This model serves to capture the expected learning outcomes specified in educational policy documents and standards, the actual implementation of the curriculum in classrooms, and the tangible learning outcomes achieved by students after receiving instruction (Travers & Westbury, 1988).

2.4.1.2 Different International Large Scale Assessment

TIMSS was introduced by the IEA in 1995, assessing both science and mathematics content. This initiative has since established a consistent assessment schedule, spanning four years, with a specific focus on 4^{th} and 8^{th} Grade students' performance in mathematics and science across participating countries. This design enables the examination of cohort data longitudinally, as the fourth-grade cohort progresses to the eighth grade in subsequent TIMSS cycles. TIMSS has also developed a framework for learning context questionnaires, which encompasses five levels: National and community contexts, home contexts, school contexts, classroom contexts, and student characteristics and attitudes toward learning. This framework offers a comprehensive lens to explore learning outcomes in relation to specific learning contexts, facilitating a multidimensional analysis for a more nuanced interpretation of test scores.

OECD-PISA (Organization for Economic Cooperation and Development - Program for International Student Assessment) was initiated by the OECD in 1997 as an international large-scale assessment program. PISA aims to address the question of what knowledge and skills citizens require to thrive in their personal, social, and professional lives. Over seven triennial testing cycles, starting in 2000, PISA has evolved into an international assessment program encompassing three primary subject matter domains: science, reading, and mathematics. Each test primarily focuses on one of these domains, accounting for a significant portion of the testing time, while the remaining subject domains are assessed during the remaining time. PISA also incorporates an assessment of financial literacy and has introduced Innovative domain assessments, including problem solving, collaborative problem solving, and global competence, with assessments of creativity planned for 2022 and learning in the digital world for 2025. PISA's primary objective is to provide policy-relevant indicators, particularly to OECD member countries, to support the improvement of their education systems, considering the perceived literacy needs and challenges of the 21st century (OECD).Instances of extensive educational evaluations encompass a variety of programs, e.g. PIRLS, PISA, PIAAC, SACMEC, PASEC, and ERCE.

The PIRLS survey, conducted by the International Association for the Evaluation of Educational Outcomes (IEA), assesses the reading comprehension skills of fourth graders, usually between the ages of nine and ten. This study has been conducted every five years since 2001, with a third cycle in 2011 involving more than 50 education systems worldwide. The PIRLS assesses two primary reading purposes: 1) reading for literary enjoyment and 2) reading for acquiring and applying information. An equal number of tasks are devoted to each of these reading goals

(Mullis, Martin, Kennedy, Trong, & Sainsbury, 2009).The PIRLS also measures four basic reading comprehension processes that readers use in and out of school

- Targeting and obtaining explicitly stated information
- Making direct inferences.
- Interpreting and integrating ideas and information.
- Evaluation and criticism of content and textual elements.

PIRLS takes a curriculum-oriented approach, taking into account the curricula of different countries when creating reading passages and corresponding tasks. In addition to academic achievement, PIRLS provides valuable information about students' sociodemographic background, dispositional characteristics, home environment, and teaching and learning conditions in classrooms and schools (Mullis et al., 2009). This wealth of information can benefit policymakers and researchers interested in a variety of areas such as reading research, motivational beliefs, educational disparities, and comparative education, among others.

International Civic and Citizenship Education Study (ICCS) stands as a paramount international endeavor, singular in its dedication to the examination of civic and citizenship education worldwide. It serves as an invaluable wellspring of knowledge, enriching our comprehension of how different nations prepare their youth for active citizenship within the ever-evolving contexts of democracy and civic engagement. As our world grapples with new and emerging challenges, ICCS endeavors to deepen our understanding of critical issues such as students' roles in global citizenship, environmental sustainability, social dynamics within schools, the use of new social media for civic involvement, digital citizenship, and the complex themes of migration and diversity. By participating in ICCS, countries gain access to reliable comparative data, which, in turn, assists in shaping educational policies that address these vital areas. Furthermore, ICCS participation aids countries in monitoring their progress towards achieving the United

Nations' Sustainable Development Goals, particularly Target 4.7, which emphasizes quality education for all.

2.4.2 Regional Assessment

Regional assessments are specially designed for developing countries with the same conditions of economic and social development. Participation in regional courses enables education ministers to determine how well their education programs are performing on time and to measure how well their students are performing compared to others in the region. Following are some of the other important regional assessment:

2.4.2.1 The Programme for the Analysis of Education Systems (PASEC) is managed by CONFEMEN and has been in place since 1993. It primarily assesses student abilities in mathematics and reading in Grades 2 and 6, measuring progress over the course of a school year. PASEC has been administered in Francophone West Africa, with some participation from Cambodia, Lao PDR, and Viet Nam in 2011/2012.

2.4.2.2 The Pacific Islands Literacy and Numeracy Assessment (PILNA) is an evaluation designed to gauge the literacy and numeracy proficiency of students in Grade/Year 4 and Grade/Year 6. This assessment is conducted in specific schools across 15 Pacific Island Countries and is a component of the Educational Quality and Assessment Program (EQAP) implemented by the Secretariat of the Pacific Community (SPC). The results of PILNA are shared with the Ministries of Education in these 15 countries to assist in enhancing student literacy and numeracy performance.

2.4.2.3 The South Pacific Form Seven Certificate (SPFSC) is a regional qualification overseen by EQAP, a program operating under the Secretariat of the Pacific Community (SPC). It has been accessible since 2004 and is conducted in Kiribati, Samoa, Solomon Islands, Tuvalu, and Vanuatu. SPFSC functions as a Grade 13 certification, encompassing various subjects such as accounting, agriculture, English, physics, and tourism & hospitality etc. This certificate serves multiple purposes, including acting as a secondary school leaving certificate, an entry qualification for universities and tertiary institutions, and as a basis for employment selection.

2.4.2.4 The Central Asian Program for Student Assessment (CAPSA) is a regional initiative supported by the German Society for International Cooperation (GIZ) and involves Kazakhstan, Kyrgyzstan, Tajikistan, and Turkmenistan. Its objectives include supporting education reform, establishing networks between educational institutions and experts, and building capacity. CAPSA assesses the mathematical and reading comprehension skills of Grade 4 students.

2.4.2.5 The Southeast Asia Primary Learning Metrics (SEA-PLM) is a collaborative regional project aimed at advancing educational evaluation in primary education within Southeast Asia. The initial data collection occurred in 2019, with the participation of six countries. This unique assessment evaluates the reading, math, writing, and global citizenship skills of fifth-grade students.

2.4.3 National Assessment

National assessments are characterized as "Routine and methodical measurement activities aimed at ascertaining the knowledge students have acquired from their educational journeys" (UNESCO, 2000).A national assessment is a survey carried out in schools to gather evidence about pupils' achievement in the core curriculum. They provide information about all or part of the educational process, such as classes or age groups. There are two types of national assessments: National Assessment (NA) and Public Examinations (PE). National assessments typically pose minimal risk to individual students and serve the purpose of monitoring a country's education system while offering valuable feedback for school enhancements. Physical education, on the other hand, poses a high risk to students and is regulated and sometimes blamed by schools for change, selection and recognition in schools (HO, 2012).Most countries have NA and PE, but some countries only have one or the other. Public examination results are often used for selection or certification, and are sometimes also used to evaluate schools. National bodies within government or independent of government are responsible for setting standards and operations for NA. The principal purpose of NA is to assess student achievement, improve curriculum and practice, and inform educational policy and interventions (Postlethwaite & Kellaghan, 2008). NA approaches and methods vary from country to country, but student goals and assessed subjects are common. NA focuses mainly on students from secondary to completed basic education and at the end of secondary education. Subjects assessed include First Language, Second Language, Mathematics and Natural and Social Sciences. However, there is a tendency to value reading and mathametical skills, and other subjects such as science are also important to foster innovation and international competitiveness (Fromin, 2007).

The Public exams are more important for students and schools; it plays an important role in guiding students' learning, determining their future careers and evaluating school work performance. However, there has been controversy about the negative aspects of testing, such as promoting "teaching to the test" and narrowing the curriculum and learning objectives (Ho, 2012; UNESCO, 2018). It also refers to assessments aimed at providing systemic data on the achievement of learning outcomes set by the government department or education department. Apart from that, it also organized data refers to how students work as a group rather than as individuals for example The Educational Quality and Accountability Office (EQAO) assessment, which is compulsory in Ontario, Canada, delivers annual standardized outcomes across multiple subjects in alignment with the skills and knowledge prescribed by the provincial curriculum (EQAO, 2009).The national assessment is usually in line with the policy standard at an integrated level, for example, a planned target outcome may be that all schools should achieve at least 50% as a minimum point mark in Grade 6. National assessments provide statistical data on limited student performance by informing policy makers, assisting with curriculum development and delivery, and better resource management (Smithers, 2004).The Dakar Framework for Action in 2000 stressed on the need for an assessment system that accurately assessed, and government worked towards ensuring quality education (Kellaghan & Postlethwaite 2008).

To identify the achievement of students in a curriculum by keeping in view the age and level appropriateness is one of the needs of the key purposes of the National Assessment Test. It facilitates the policy decision-maker about the key areas of the system by providing valuable information or data.

2.4.4 School Based Assessment

School assessments are a major element of the process of instruction and learning. Since they help to increase student knowledge and the efficiency of instruction. Implementing school assessment can demonstrate what has been accomplished during the teaching and learning process (Hammond &Wentworth, 2010).

The term School-based Assessment (SBA) has been used to refer to formative assessment on all subjects for the secondary grades in schools. Through SBA, a student's achievement of all the outcomes specified by the national curriculum can be measured, ensuring that even those competencies that cannot be measured easily on a paper and pen examination, i.e. the GCE examinations, will give proper attention in each child's schooling (Asian Development Bank, 2019).

For teachers and schools to be effective in creating a harmonious and balanced community, school evaluation is necessary. In order to maximize each learner's potential, school evaluation is a continuous process that needs the dedication and direction of instructors and schools.

In addition, school assessment features the following traits (Williamson, 2017).

a) Holistic - which means being able to give all the information about the knowledge and skills that the students have acquired.
b) Persistent − assessment activities throughout the teaching and learning process.
c) Flexible – various evaluation techniques as necessary and student readiness
d) Standardized --Refers to performance standards developed by the standard curriculum.

2.5 Different Approaches for Administration of LSA

The first approach involves linking state-administered tests to a national benchmark like the National Assessment of Educational Progress (NAEP). This approach aims to make state test results comparable to a common standard, akin to comparing different currencies against the U.S. dollar. However, this approach faces significant challenges, including variations in content, format, and student samples among state tests. Moreover, it requires extensive overhauls in state assessment programs, standardization of administration conditions, and the development of new test forms each year, making it unlikely to gain widespread consensus or feasibility in many educational systems.

The second approach, known as the Voluntary National Test (VNT), proposes the administration of a single, common test across states, which would be designed to align closely with a widely accepted national standard like NAEP. This approach, while potentially providing meaningful national and international comparisons, raises concerns about federal government control over curriculum and the narrowing of instruction to focus on preparing students for one test. Additionally, the VNT's limited coverage of only fourth graders in reading and eighth graders in math may overlook other important subjects and grade levels. Test security and the substantial costs associated with large-scale testing also present formidable challenges in implementing this approach. These considerations emphasize the need for a balanced and comprehensive evaluation of the benefits and drawbacks of both strategies when considering the establishment of a national measure of achievement in any educational system.

2.6 Purpose of Different Assessment at Different Level

Different forms of learning assessment are used in education institution**.** At the global level, international assessments (IAs) such as PISA, PIRLS, TIMSS, and ICILS provide data for assessing student outcomes and allow for cross-country comparisons. The main purposes of participating in IAs include education policy review and reform, monitoring and evaluating education quality, curriculum review and reform, cross-national comparisons, and policy or program evaluations. However, participation in IAs is limited, with only a few countries taking part.

At the regional level, some countries participate in regional assessments (RAs) such as CAPSA, PILNA, PASEC, and SPFSC. Regional assessment (RAs) are employed to oversee and assess educational quality, scrutinize and revise curricula and educational policies, facilitate international comparisons, and gauge disparities in education.

Education at the national level uses National Assessments (NA) and Public Examinations (PE) tme asure student learning. NAs are designed to measure student achievement in specific subjects and p rovide an overall assessment of academic achievement across the curriculum. These assessments s erve many purposes, such as monitoring and evaluating education quality, informing education pol icy refom, reviewing and strengthening curricula, facilitating change comparison within the countr y, and identifying educational disparities. However, NAs are rarely used for student and teacher res ponsibility or to assess children's abilities outside of the academic curriculum. NA mainly focuses on assessing students' core skills, knowledge gained from the course, and their ability to apply knowledge in a real environment.The most commonly assessed subjects in NA incl ude Mathematics, Native Languages, English, Science, Research and Mathematics. These tests are offered at different levels, from Grade 1 to Grade 12, specific to primary or secondary school.
In contrast, the primary purpose of public examination is to assess students' knowledge gained in t he classroom and their ability to apply that knowledge in the real world. Public examination genera lly does not measure a student's interest or attitude towards a subject, nor does it measure non.intell ectual abilities. Subjects frequently assessed in Public examination include mathematics, English, l ocal languages and other school subjects.

Public examination.is usually completed in secondary education, especially in year 12. However, som e countries also offer Public examination at the primary level, such as the 5th and 6th grades.

2.7 History of National Assessment

It started in 1990 when developing countries started focusing on national achievements programs and by 2006 half of the developing countries were carrying out national assessment test. Furthermore, in a more recent survey, it was found that over two thirds countries which is almost 151 have participated during 1960-2008 period in one or other assessment for example international, regional or national assessment (Benavot, & Kamens, 2011).

2.7.1 Common Features of National Assessment

All over the world, the National Assessment System has some common features for example all deals with the assessment of students' language or literacy and students' mathematics abilities or numeracy. Additionally, a prevalent characteristic is that many educational systems evaluate students' performance in subjects like a second language, science, or social studies. Moreover, while the majority of countries worldwide conduct national assessments at the primary-school level, there are regions where national assessments are also conducted during secondary school.

2.7.2 Different National Assessment Worldwide

Various versions of national assessment are used worldwide, for example in the U.S., The National Institute for Educational Progress (NAEP) compiles data from a nationally representative sample (NAEP, 2006). In France, they employ diagnostic assessments for students at advanced levels to pinpoint individual needs and tackle performance challenges (Jennings et al., 1999). The Australia NAPLAN administers on an annual basis to students in grades 3, 5, 7, and 9.Meanwhile, Canada has its own Pan-Canadian Assessment Program., established by the Department of Education, designed to evaluate the literacy, science, and mathematics skills of 13 to 15-year-old students (Canadian Council of Education Ministers, 2007). Singapore adopts the International Primary School Examination (IPSLE), a national assessment undertaken at the conclusion of primary education, assessing students' suitability for tertiary education and determining their placement in courses that align with their learning pace, capabilities, and inclinations. Results categorize candidates into four distinct streams: special, explicit, general (educational), and general (technical) (Singapore Department of Education, 2007).While most of these tests are not mandatory, some areas in government-sponsored countries also have additional local comparison tests.

Mostly National achievement looking for responses to one or several of the following inquiries

- To what extend student learning is good in the education system. Concerning the aims and objective of education, general expectation, and keeping in view the life-based education.
- Does evidence suggest weak and strong areas in the students' knowledge and skills?
- Do differences exist among male and female, urban, and rural area students or students from different backgrounds and regions of the country? Do subgroups in the population perform badly?
- Student achievement is linked with what elements? To what extent qualities of the learning environment or the surrounding of the impact on the achievement of the student?
- In the provision of resources are government standards being met?
- Is there any difference occurs in the achievements of overtime? To answer this question different assessments, need to be carried out at a different point that generates comparable data (Greaney & Kellaghan 2004).

- Provide factual information to assist in policy making, especially student learning quality data; making informed decisions about the curriculum, resource allocation, and teacher training strategies.
- Gather general information about student performance to help improve standards.
- Introducing practical standards by encourage a realistic perception of suitable achievement levels is crucial. Unrealistic expectations can lead to lower performance levels.
- Identifying the correlations of successes and outcomes by linking the scope of factors and outcomes to identify strengths and weaknesses in the system.
- Provide teacher guidance to enhance student achievement by aligning teaching and learning practices with national assessment goals and prescribed outcomes.
- Encouraging accountability and utilize assessments to justify the distribution of resources.
- Raising public awareness and publish national surveys to attract greater media attention, thereby raising public awareness of educational issues.
- Informative political debate: direct statistical evidence on issues arising from the education system (Greaney & Kellaghan 2008).

2.7.3 Key Consideration for the Development of National Achievement Tests

a. Development of assessment framework

- Creating a clear definition of the construct: Clearly define the specific areas of knowledge being assessed, such as reading or mathematics, and understand the processes and characteristics associated with the tasks.

- Purpose and Stakeholder Engagement: Determine the purpose of the assessment and engage education stakeholders, including curriculum experts, policy makers and education managers, to ensure that the assessment is aligned with education policy objectives.
- Test plan: Create a test plan or table of specifications that outlines the data to be collected, the length of the test, and the proportion of items for each curriculum area, the assessment of different skills and cognitive processes, and the balance between multiple-choice and open-ended items.
- Validation: Work with subject matter specialists and curriculum experts to ensure that test content accurately matches the curriculum or construct being measured. Consider whether the test provides sufficient evidence of student achievement in that area.
- Test Language: Clearly define and justify the language(s) to be used in the assessment. Consider the teaching method and the need for translation and ensure that the translated versions are equivalent for comparison purposes.
- Item Format: Determine appropriate formats for student responses, such as multiple choice, closed response, open-ended short response, or essay/extended response, based on assessment objectives.
- Target population: Define the target population for the assessment (eg grade) and justify the selection. Consider including or excluding specific subpopulations, such as students with special educational needs or students from remote areas.

- Reporting of results: Decide how the results will be reported, with respect to subscales, proficiency levels, or other forms of reporting. Plan to publish technical reports, summary reports for specific audiences, and press releases.
- Contextual Information: Consider gathering contextual information related to home and school factors that may affect student learning. This information can provide insight into differences in student achievement and guide the development of questionnaires.

b. Characteristics of good test items:

- Clarity and relevance: Test items should be clear, relevant to the curriculum and focused on a specific aspect of learning.
- Fairness: Test items should be fair to students from diverse backgrounds and avoid penalties or advantages based on irrelevant life experiences. Good test items should relate to a specific learning area, be constructive and meaningful, be consistent with document plans, and focus on the main issues presented in the stimulus material.

c. Utilization of sample items and provision of suitable models:

Item writers can benefit from examining high-quality sample items published by reputable testing organizations and international reviews. However, it is important to ensure the appropriateness and suitability of these items. Developing expertise in item writing requires a common understanding of the terminology and purpose of the items. Concept items should be classified based on item format, text type, grade level, learning outcome, and cognitive process.

d. Determining Item Difficulty:

Determining item difficulty can be challenging, and item writers often need to create items that are easier than the intended curriculum. Pretesting items on a sample of students similar to the target population is essential to obtain objective data on difficulty levels and to avoid items that are excessively difficult. Factors such as unknown processes and item wording can affect item difficulty.

e. Avoiding Item Misrepresentation:

To ensure fairness, item bias must be avoided. Authors of items should be aware of various forms of bias, and reviewers should be vigilant in identifying biased or offensive items. Stimulus material plays a key role in providing context for the items and should be factual, interesting, challenging, self-contained, and factually accurate. Clear and simple language should be used when writing items, avoiding difficult vocabulary, convoluted logic, ambiguity and inconsistency.

f. Importance of pre-testing

Pretesting items in test development, especially in the context of national assessments, is essential. Pretesting involves administering a test to a sample of students with characteristics similar to the target population. It helps assess the appropriateness and quality of test items and allows fine-tuning of test administration procedures.

g. Using multiple test booklets

National assessments often use multiple test booklets to ensure sufficient coverage of the subject matter without overwhelming individual students. The design and administration of such tests can be complex and require careful attention to item overlap, comparisons between subsamples, and appropriate distribution of booklets.

h. Conducting preliminary testing

Pre-testing should ideally be done a year in advance in the same timeframe as the final test. It is recommended that schools of different sizes, locations and socio-economic backgrounds be involved. Preliminary examination forms should be similar in length to final examination forms and should be labeled alphabetically by grade. Multiple pre-test forms are required and should be randomly distributed in each class to ensure fairness.

i. Pre-Test Forms Link

Linking pretest forms is essential for data collection and comparison. Common items, known as link items, are included in various forms to establish connections. Horizontal linkage occurs within the same terrain level, while vertical linkage is necessary to test multiple terrain levels. Reference items should be moderate in difficulty so that students of average ability can answer them correctly.

j. Minimum Pre-Test Requirements

Each pretest item should be attempted by a minimum of 200 grade-level students, with at least 150 responses collected. The pretest should reflect the conditions of the final test, including the time allotted for completion. It is recommended that all pre-testing be completed within two to three weeks.

k. Draft final test

The design of the final test involves consideration of multiple factors to ensure its quality and suitability for analysis. The involvement of a data analyst or statistician is essential to meet specific requirements. These include ensuring an appropriate format for recording student background information, selecting an appropriate method for recording item responses, creating appropriate scoring guides, and implementing statistically reliable horizontal and vertical links to prior test data.

l. Front Cover of Test Booklets

The front page of the test booklets should contain relevant background information for students for the purposes of the test and reporting of test data. This information usually consists of the name of the school, student.

2.7.4 General Key Feature of High Quality Items

- High-quality items prioritize core concepts and skills within disciplines like ELA, math, science, and history. For example, ELA prompts emphasize using evidence in writing, while math questions assess crucial ideas without unnecessary complexity.
- Top-notch assessment items go beyond basic comprehension, integrating knowledge, understanding, and skills. In science, this means combining scientific knowledge and practices. In history, it involves assessing knowledge within its historical context.
- Items are designed with clear language and engaging contexts to support diverse students, including English learners. Tasks avoid unnecessary complexity and provide various entry points, like visuals or technology aids in math.

- Constructed and extended-response items encourage multiple student responses, focusing on reasoning skills rather than a single right answer. Math allows multiple solution strategies, while ELA and history prompts remain unbiased.
- Whenever possible, items replicate authentic discipline-specific work. History tasks might involve analyzing artifacts, science items could use simulations, and ELA questions reflect core concepts.
- Technology-enhanced items should enhance complexity or accessibility. They should offer scaffolding without reducing cognitive demand or increase item complexity, such as using simulations for authentic science assessment (Walkington, 2013).

2.8 General Principles for Designing Science Assessments

In crafting science assessments, several fundamental principles should guide the development process to ensure their effectiveness and alignment with modern educational objectives. Below are four key principles to consider:

- Assessments should align with the multifaceted nature of scientific understanding. Science education often comprises disciplinary content, scientific practices, and overarching concepts. Therefore, assessments should evaluate students' abilities to not only recall facts but also apply their knowledge to engage in scientific practices, such as experimentation, analysis, and problem-solving. A comprehensive assessment should probe each of these dimensions to provide a holistic view of a student's scientific proficiency.
- Science assessments should prioritize fundamental concepts and big ideas over minutiae. Rather than focusing on rote memorization of facts, assessments should evaluate students'

comprehension of overarching principles. This ensures that students grasp the core concepts that underpin scientific knowledge and can apply them in a variety of contexts.

- Assessments should encompass a wide range of scientific and engineering practices. Modern science education includes a diverse set of practices that extend beyond traditional modes of assessment. Students should have opportunities to demonstrate their ability to not only conduct experiments but also design investigations, analyze data, and communicate their findings effectively. Assessments should, therefore, be designed to evaluate students' proficiency in these various practices.
- Well-designed science assessments should present students with real-world scenarios and challenges that demand thoughtful analysis and creative solutions. By emphasizing these skills, assessments promote deeper learning and better preparation for the complexities of the modern world (Achieve, 2013).

2.9 Principles for Designing Mathematics Assessment

- Assess fundamental mathematical processes and critical concepts coherently.
- Provide various entry points through visual aids, technology, or response formats.
- Encourage diverse problem-solving approaches and allow for multiple correct answers.
- Present stimuli and prompts concisely, clearly, and without unnecessary difficulty.
- Utilize technology enhancements to scaffold student reasoning or enhance item rigor.
- Situate items within meaningful and relatable contexts for diverse student backgrounds (Yuan, 2014).

2.10 The Assessment System In Pakistan

An examination is a tool that is being used to evaluate the performance, achievement ability in a subject. Seven different education policies have been introduced since the creation of Pakistan. The focus of all these policies was to promote quality education but the result remained unsatisfactory mainly because of the incompetent assessment procedure. Our assessment methods only encourage children to reproduce exactly what is written in books instead of making then analyze things, explore new concepts, this trend in assessment hampered the way toward quality education. UNESCO (2007), expressed concerns about Pakistan's educational assessment system, highlighting its shortcomings in effectively educating children basic skills. The main reason is poor teaching methods, lack of proper supervision, and training modules for a teacher, and incompetent assessment criteria.

Assessment needs to guide the students, what are their weak areas, where improvement is needed, and what skill should be polished more. Unfortunately in Pakistan, it is subjective with little or no feedback at all, student only gets their result without knowing where they are lagging.

The Ministry of Education (2009), declared that assessment criteria will onward focus on the competence of a student in a specialized area. Periodic reviews of assessment systems along with the introduction of advance assessment strategies should be done so that a balance is created between the use of formative assessment approaches and the summative approach of high-stake examination.

2.10.1 History of National Assessment System in Pakistan

2.10.1.1 Emergence of LSAs in Pakistan: LSAs are actually regarded as an important parameter for evaluating the education system, monitoring its quality, and making decisions on strategies and procedures for improving the quality of teaching and learning. In the Past, Government and non-governmental organizations have made various efforts to collect student performance data through a variety of tools including achievement tests, school-based surveys and in-house surveys. The tradition of scrutinizing foreign higher dividends has its roots in colonial times during British governance, and this practice persisted after gaining independence in 1947. Notably, Pakistan lacked an independent inspection board at that time. The inaugural Board for Secondary and Intermediate Examination (BISE), was founded in Karachi in 1951, followed by the establishment of several more boards in subsequent years.

The Sharif Commission (1959), introduced tests based on internal schools in the initial grades, efforts have been made to enhance students' performance in external examinations. Nevertheless, it has been reported that the situation has not improved in recent years due to challenges like teachers with limited skills, inadequate infrastructure, and a shortage of assessment tools. The NEP report of 2017 pointed out this issue, leading policymakers to establish the National Education Assessment System (NEAS), at the government level and Provincial / Local Education Assessment Institutions (PEACs / AEACs). NEAS was established as an assessment organization in 2003 as part of a World Bank funded project – Education Sector Reform (ESR) program – with the Ministry of Education. NEAS received technical and financial support through this five-year project. Following completion of the project period, the Ministry of Education included NEAS and the accompanying "provincial and areas education assessment centers" (PEACs/AEACs) as entities within federal and provincial education departments.

NEAS is now the National Assessment Wing (NAW) of the Pakistan Institute of Education (PIE).

NAW's mandate is to conduct periodic national, sample-based assessment studies at the elementary level in core subject areas. Findings from these assessment studies are key sources of evidence for policy interventions and capacity building.

Since its inception, NEAS, with support from PEACs/AEACs, has conducted six large-scale, sample-based assessment studies across the country. For example, NEAS has led Periodic assessments in Grades IV and VIII in the core subjects of mathematics, science, and language (Urdu & English). NEAS targets Grades IV and VIII to ensure alignment with international assessments. In addition to six national assessment studies, NEAS also spearheaded Pakistan's participation in the 2019 administration of TIMSS. Pakistan became the first country in South Asia to participate in this international study, with the expectation that analyses and insights from this initiative will enable sustained improvements in mathematics and science learning.

The role of NAW in developing and executing robust assessment studies is especially critical given the severe learning challenges in Pakistan. Pakistan has a high incidence of learning poverty owing due to a range of structural factors. The out-of-school population ranks second highest globally, with approximately 22.8 million children between the ages of 5 and 16 not receiving formal education. This figure accounts for 44% of the entire population within this age bracket (PSLM, 2018-19). According to the Pakistan Economic Survey for the fiscal year 2021-22, it is estimated that in the year following 2020-21, around 14.4 million students were enrolled in pre-primary education, while primary education, covering grades 1 through 5, had an enrollment of roughly 25.7 million students. Additionally, middle education, spanning grades 6 through 8, saw an enrollment of approximately 8.3 million students. In this context, it is important to ensure continued improvement in the assessment system so that it informs remedial policies and actions that are beneficial for the teaching and learning landscape in the country.

2.10.2 National Achievement Test (NAT)

Different type of assessments of student achievement in Pakistan was held in past for various Grades like for Grade 3,4,5,6 and 8 aimed at evaluating the educational outcomes of students. Student achievement is the main purpose of such assessment, also to set the level of learning at respective grades and to help feed into a policy for improving the education system. The poor performance of the school and students with low-level competencies is the major concern of different countries in the world. Since the efforts were being made to develop a mechanism for monitoring and evaluating student learning outcomes at national level. This type of assessment is known as the National Assessment System (NAS) (UNESCO, 2007).

The NAS selects key areas in the curriculum at selected grade and level on a regular interval which helps to measure the achievement level of the student. The major objective of this all exercise is to assess the achievement of the student which automatically reflects the education system of that particular area/country.

To identify the strength and weaknesses of the curriculum can be a judge by the data received on the performance of students at various subdomains of the curriculum. This data on student achievement correlates with other factors like student-related, school-related or teacher-related factors and information of their relationship can obtain which can help policymakers to focus on those areas that are likely to contribute to learner achievement levels through examination assessment.

In Pakistan, NAT is one major Large-scale assessment carried out across the country. Its results give valuable data to the policymakers and implementers on the performance of the education system and other factors that affects the performance of the student. This assessment plays an important role in determining the effectiveness of education system and the effect of investments in education and how successfully it contribute to student learning by delivering resources.

2.10.2.1 National Assessment Wing: NAW is empowered to conduct extensive national assessments throughout Pakistan, offering insights into the performance of the existing education system by evaluating students' learning achievements. During the project's duration, NAW conducted several assessment cycles in 2005, 2006, 2007, and 2008, followed by assessments in 2014 and 2019 after the project concluded.

The objectives of NAW are:

- To provide data related to inequality in student performance due to geography and gender
- To evaluate how well the curriculum is translated into knowledge and skills.
- Identifying correlates of student achievement with other factors
- Guiding Teachers' Efforts and Boosting Students' Achievements

2.10.2.2 Different Assessment Frameworks

For NAT, NAW used the following assessment framework for different subjects which are based on two strands, content strand, and ability strand. The Achievement tests used by the NEAS were based on the National Curriculum, 2006.

The procedure for developing these achievements are given below:

• Knowledge, Understanding, and Application the three important levels of the cognitive domain of bloom's taxonomy were considered as a baseline for all the tests of different subjects, keeping in view national curriculum content areas and objectives.

• For test especially the table of specifications which contains content, learning outcomes, number of items and type of item was designed.

• The test was based on multiple choice and constructed response, completion, and matching items. However, to provide better standardization of test setting and marking. Multiple choice and constructed response items were considered.

2.10.2.3 Mathematics Assessment Framework

The framework for Mathematics for Grade IV and Grade VIII is based on learning outcomes outlined in the 2006 national curriculum that has been implemented across all of Grades I – VIII. This framework delineates the purpose, format, content, and cognitive domains addressed in the mathematics tests. It is not meant to serve as a guide for instruction or educational direction.

NAW designed the framework to facilitate the development of test items and guide subsequent test construction. It includes learning outcomes from the National Curricula 2006 for Grade IV and Grade VIII. These will be covered in the test as content domains. The cognitive processes associated with the measurement of mathematics are also detailed under cognitive domains. The framework includes tables of specifications (test blueprints) from which valid, reliable and comparable tests can be constructed each year for pilot testing and large-scale assessment across the country.

a) Cognitive Domains for Grade IV and Grade VIII

For content knowledge to be internalized effectively, and applied to diverse situations, it is important to ensure simultaneous focus on students' cognitive domains and skills. Under the NC, the cognitive domains have been made consistent with the TIMSS assessment framework to ensure greater alignment of Pakistan's learning assessments with international standards. The cognitive domains outlined in the NC. Within the domain of knowledge, students are expected to acquire a solid understanding of words, symbols, and essential concepts, with a focus on their precise application and comprehension. Moving to the domain of application, students should demonstrate their ability to effectively employ mathematical principles and procedures when tackling real-life situations, encompassing numeric, algebraic, geometric, and statistical scenarios. Lastly, the domain of reasoning calls upon students to apply their pre-existing mathematical knowledge in novel contexts, demanding the analysis, synthesis, and evaluation of information to solve complex real-world problems while critically assessing data sufficiency and consistency (NAT Report,2016). The NAT reflect the same cognitive domain proportions as the TIMSS papers for both Grade IV and Grade VIII.

Cognitive Domain	Percentages	
	Grade IV	Grade VIII
Knowing	40%	35%
Applying	40%	40%
Reasoning	20%	25%

b) Curriculum Aims and Content for Grade IV

The National Curriculum outlines a list of aims meant for students at the primary level. At the primary level, the curriculum aims to instill the ability among students to employ basic mathematical skills in everyday use. This ties into the aim of ensuring that students develop the ability to think logically in applying basic mathematical knowledge to diverse situations. Another aim of the curriculum is to ensure that students have complete conceptual clarity of foundational topics. A deep understanding of these concepts is hence essential for children's educational journey beyond the primary level.

In pursuit of these aims, the Ministry of Education has structured the curriculum to include content strands that at one level provide students with necessary skills to face diverse challenges, and at another level provide students with clarity on important concepts that serve as building blocks for subsequent learning at higher levels of education.

These content strands are:

- Number and Algebra
- Measurement and Geometry
- Statistics and Probability.

Table 2.1 Mathematics Assessment Framework for Grade 4 and Grade 8

Grade 8			Grade 4		
Content strands	**Proportional weightage in percent**	**Content percentage wise distribution**	**Content strands**	**Proportional weightage in percent**	**Content percentage wise distribution**
Numbers & operation	**35 percent**	Operation on sets. Real numbers Number system	**Numbers & operation**	**45 percent**	Numbers & arithmetic operations Factors and multiples

		Financial mathematics			Fractions Decimals and fractions
Geometry	**35 percent**	Fundamental of geometry Practical geometry Area and volumes Demonstrative Geometry	**Measurement & Geometry**	**35 percent**	Measurement Geometry
Introduction to trigonometry	**5 percent**		**Information handling**	**10 percent**	
Algebra	**20 percent**	Polynomials Factorization, simultaneously equation			
Information handling	**5 percent**				

Table 2.2: Distribution of Items by Mathematical Abilities for Grade-8 and Grade 4

Level of learning (mathematical ability)	Weightage according to Curriculum Objectives Grade-8	Level of learning (mathematical ability)	Weightage according to Curriculum Objectives Grade-4
Connectional Understanding	32%	Conceptual understanding	40 percent
Scientific Investigation	42%	Procedural knowledge	40 percent
Practical Reasoning	26%	Problem solving	20 percent
Total	100%	Total	100%

Table 2.3: Table of Specification showing Distribution of Test Items by Content and Level of Learning (Mathematical Abilities) for Grade 8

Ability Level	Number & Operations (35%)	Geometry (35%)	Algebra (20%)	Trigonometry (5%)	Information Handing (5%)	Total
Conceptual Understanding	6	6	5	1	1	19
Procedural Knowledge	9	9	5	1	1	25
Problem Solving	6	6	2	1	1	16
Total No of Item	21	21	12	3	3	60

Table 2.4: Table of Specification showing Distribution of Test Items by Content and Level of Learning (mathematical Abilities) for Grade 8

Ability Level	Number & Operations (35%)	Geometry (35%)	Information Handing (5%)	Total
Conceptual Understanding	6	6	1	19
Procedural Knowledge	9	9	1	25
Problem Solving	6	6	1	16
Total No of Item	21	21	3	60

2.10.2.4 Science Assessment Framework

The Science framework utilized in NAT is founded on a consensus regarding the essential components of science education and learning against which student achievement is assessed. This is a two-dimensional framework: At content level the subject of science is divided in the fields of Life Science, Physical Science, and Earth Science. This is in accordance with distribution used in curriculum In terms of level of learning three levels of learning are used in the framework which includes conceptual understanding, scientific investigation, and practical reasoning.

Table 2.5 Science Assessment Framework

The Science framework places emphasis on the knowledge, skills, and competencies necessary to foster a deep understanding of science among students. It is constructed in the form of practical problem-solving tasks that involve design and use of materials while considering the students' developmental level. It is based on the understanding of desirable elements of science education against which student attainment ought to be measured. It covers content domains and learning outcomes for science as outlined in the National Curriculum 2020 for Grade IV and the National Curriculum 2006 for Grade VIII. The cognitive dimension specifying the thinking processes to be assessed are also detailed under cognitive domains.

a) **Cognitive Domains for Grade IV and Grade VIII**

Cognitive Content	**Knowing (40%)**	**Applying (40%)**	**Reasoning (20%)**	**Total**
Life Science (45%)	18%	18%	9%	45%
Physical Science (35%)	14%	14%	7%	35%
Earth and Space Science (20%)	8%	8%	4%	20%
Total	40%	40%	20%	100%

For content knowledge to be internalized effectively, and applied to diverse situations, it is important to ensure simultaneous focus on students' cognitive domains and skills. Under the National Curriculum, the cognitive domains have been made consistent with the TIMSS assessment framework to ensure greater alignment of Pakistan's learning assessments with international standards. The cognitive domains outlined in the National Curriculum are: In the "Knowing" domain, students are tasked with acquiring a solid grasp of fundamental concepts, facts, and their underlying principles, emphasizing the precise application of these ideas, encompassing definitions, relationships, and representations. Within the "Applying" domain, students are expected to adeptly choose and employ pertinent scientific concepts and procedures when addressing real-world challenges. The "Reasoning" domain necessitates students to utilize their prior scientific knowledge in novel contexts, involving the identification, formulation, and resolution of real-life problems through critical analysis, synthesis, and evaluation.

b) Curriculum Aims and Cognitive Domains for Grade IV

The National Curriculum outlines salient goals associated with Science education at the primary level. These include: enabling students to develop critical thinking and curiosity about scientific and technological processes, encouraging learning through inquiry; providing a sound conceptual basis for children to acquire and benefit from scientific knowledge throughout their academic careers; preparing students to situate prevailing socioeconomic and environmental issues within the realm of scientific inquiry; and developing a lasting interest among students to explore further academic studies leading to careers in the fields of science, technology, and environmental science.

The curriculum at Grade IV level is structured across three content strands. The content strands containing specific topics/domains are meant to build a robust foundational understanding of basic scientific concepts as well as expose children to how such scientific processes are linked with and at the source of what they see around them. The three content strands are:

- **Life Science**
- **Physical Sciences**
- **Earth and Space Sciences.**

c) **Curriculum Aims, Content and Cognitive Domains for Grade VIII:** The National Curriculum 2006 emphasized the need to enhance scientific literacy in the country. Building on students' learning at the primary level, the objective of the secondary-level curriculum is to enhance students' understanding of scientific processes, thereby increasing their interest in science as not just an in-class subject, but as a wide-ranging field of inquiry and exploration. Learning at the secondary level is critical to the objective of shaping students' minds towards acquiring scientific knowledge, using it as a lens to view social and environmental challenges around them, and hopefully pursuing further studies to adopt fulfilling careers in science that not only benefit them but also their communities and society. The National Curriculum identifies the following content strands for Grade VIII:

- Life Sciences
- Physical Sciences
- Earth and Space Sciences

Table: 2.6: Proportional Weightage of Various Content Strands for Grade-4 and Grade 8

Content Strands	Weightage	Percent Weightage	Content strands	Weightage	Percent weightage
	Grade 4			**Grade 8**	
Life Science	32	45	Life Science	30	60
Physical Science	17	28	Physical Science	10	20
Earth Science	16	27	Earth Science	10	20
Total	60	100	Total	50	100

The content areas in science are as given in table 2.6. The distribution of relative weightage of each content area in accordance with the weightage given in the curriculum (calculated based on SLOs for each topic in the curriculum).

Table 2.7 Proportional Distribution of science assessment by level learning.

Level Learning	Proportional Weightage by level	Percent allocation by each level
Conceptual Understanding	32	53
Scientific Investigation	22	37
Practical Reasoning	6	10
Total	60	100

Table 2.8 Table of Specification of Science Assessment for Grade-4

Learning Area	Life Science	Physical Science	Earth Science	Total
Connectional Understanding	15	9	8	32
Scientific Investigation	10	6	5	22
Practical Reasoning	2	2	2	6
Total	27	17	16	60

Table 2.9 Table of Specification of Science Assessment for Grade-8

Learning Area	Life Science	Physical Science	Earth Science	Total
Conceptual Understanding	15	7	8	26
Scientific Investigation	10	3	4	17
Practical Reasoning	5	1	1	7
Total	30	10	10	50

2.11 Different other Large Scale Assessment in Pakistan

In Pakistan, educational assessments encompass the Standardized Achievement Test (SAT), targeting all students in Grades V and VIII, with a participant count of approximately 300,000 in public and select private schools within Sindh (SAT Report, 2016). The SAT evaluates students in Mathematics, Science, and languages, aligning test items with the national curriculum and textbooks. Language proficiency is tested in Sindhi, Urdu, and English, catering to provincial, national, and official language needs. Data on factors impacting student achievement, such as socio-economic status, parental education, and teacher qualifications, is collected through district surveys. Item writers, who are subject specialists, create test items, subsequently reviewed by the Institute of Business Administration (IBA). Test items are categorized into knowledge, understanding, and application levels, with weightages of 40%, 35%, and 25%, respectively. The test administration is coordinated by Talaqua Officers (TOs) at the sub-district level, with TOs overseeing logistical aspects, including invigilation and security. Results, available at various levels, highlight consistent low performance in core subjects over the past years.

In contrast, the Annual Status of Education Report (ASER) in Pakistan, collaboratively managed by Idara-e-Taleem-o-Aagahi (ITA) and international partners, conducts nationwide household surveys annually, assessing basic literacy and numeracy skills among primary-grade students aged 5-16 (Annual Status of Education Report, ASER Pakistan, 2017). ASER surveys employ face-to-face interviews and pencil-and-paper tests in a household setting, with data collected by volunteers, often recent graduates or university students, who undergo a three-day training. Beyond test results, ASER captures valuable information on student learning gaps, gender-based and province-based disparities in language and numeracy skills, enrollment rates, gender equity, and disability statistics. ASER consistently underscores low student performance in reading and numeracy skills over multiple years, offering crucial insights into Pakistan's educational landscape.

The educational assessment system in Pakistan has been deeply influenced by its historical ties to British colonization (Khattak, 2012). Its main purpose has traditionally been to measure students' readiness for progression within the educational hierarchy. Important consequences for the future career path of students are especially critical high school exams at the matriculation and secondary level. These exams are overseen by various provincial examination boards and are conducted in languages such as Urdu, English, Pashto and Sindhi, although standardization remains a challenge. Unfortunately, many of these assessments fall short of effectively assessing the quality of learning, often emphasizing rote memorization and selective study rather than encouraging analytical and critical thinking (Ministry of Education, 2007).

In the context of upper elementary grades (5-8), assessment practices at the provincial level are still evolving, with some provinces not yet fully in place. Given the deficiencies in the current assessment system and the urgent need to improve assessment of student learning outcomes, education reforms in the early 2000s, followed by the National Education Policy of 2009, recognized the need to strengthen assessment capabilities at the school, provincial, and federal levels. According to Greany and Kellaghan (2008), a national assessment system in Pakistan should primarily focus on:

- Provide valuable feedback to policymakers, politicians and the wider education community on concrete, impactful outcomes.
- Report on student learning levels and align them with common curriculum expectations in subjects such as science, math, technology and life skills.
- Identify strengths and weaknesses in student learning while evaluating the performance of different subgroups of students.
- Facilitate comparison of student performance in Pakistan and internationally and recognize the link between the quality of education and the development of a knowledge-based economy.
- Track shifts in student achievement over time and evaluate the impact of policy decisions, particularly in terms of resource allocation.
- Explore the factors that contribute to student achievement, both in and out of the school environment.
- Assess the adequacy of teacher professional development, services and resources and bring them into line with government standards.
- Create an accountability framework that holds teachers, administrators, students, schools, institutions, and relevant agencies accountable for educational policy and decisions.
-

2.12 Different Stakeholders Perception Regarding Large Scale Assessment

2.12.1 Perception

The term “perception” comes from the Latin verb perception (Stevenson & Waite, 2011), referring to a dual process of acquiring knowledge by stimulating the senses and knowledge (Maund, 2003).

2.12.2 Teachers’ Views on LSA

Different educational principles and theories support that the teacher's views and beliefs are essential to any assessment process as they influence different teaching and learning processes. Teachers have an important role to play in identifying what is needed or what is best for their students as they focus on their day-to-day teaching. It not only affects teaching habits and classroom behavior but also, is related to student achievement. (Ringstaff, & Sykes, 1985). Recent investigations have indicated that educators in North America express unfavorable sentiments towards Large-Scale Assessments (LSAs) (Rogers, 2014). In a study by Klinger and Rogers (2011) examining teachers' perceptions of LSAs in Ontario and Alberta, it was discovered that teachers in both regions recognized LSA results as valuable information for enhancing their students' learning and academic progress. Nevertheless, teachers raised concerns regarding LSAs, encompassing issues such as parents' challenges in interpreting provincial assessment outcomes, the publication of school standards based on LSA scores, the use of LSA scores for evaluating educators, and the manner in which LSA results were conveyed (Klinger & Rogers, 2011). When considering teachers' perspectives on large-scale assessments, the initial assumptions might not capture the full complexity of their views. While examining teacher opinions on government-mandated large-scale assessments, it becomes apparent that their attitudes toward these assessments are multifaceted, encompassing both positive and negative aspects. For instance, Ballard and Bates (2008) conducted research study that indicated that teachers appreciated large-scale assessments for the valuable data, trends, patterns, and comparisons they provided regarding students, classes, and schools. However, these teachers also expressed concerns about biases in some assessment questions, which they believed could create undue pressure and unrealistic expectations for certain students. Furthermore, they felt that large-scale assessments often failed to accurately reflect students' actual abilities, regardless of the time invested in test preparation. A similar pattern emerges in Skwarchuk's (2005) study in Manitoba, where the majority of surveyed teachers held predominantly "very negative" views regarding mandated testing. However, a small contingent of teachers supported LSAs as a source to potentially enhance standards and promote uniformity in teaching practices across the province. The impact of these assessments on teaching practices seems to vary based on the stakes involved. In research conducted by Abrams et al. (2003) in various U.S. states, teachers in high-stakes assessment situations tended to allocate less time to teaching subjects and materials that were not tested compared to teachers in low-stakes assessment situations. This shift in teaching practices, driven by the implementation of state-testing programs, was generally viewed negatively by many teachers and was seen as potentially affecting the quality of instruction. Similarly, Mertler (2010) conducted interviews with 1,543 teachers in Ohio and found that state-mandated low-stakes testing also had negative impacts on instructional and curricular practices. Teachers reported changes in teaching methods and a reduced emphasis on non-tested content. Teachers in Ontario reported similar concerns about the impact of large-scale low-stakes assessments on classroom practices, including a tendency to focus predominantly on assessed areas, potentially at the expense of other curriculum content. This pressure often led to a rushed approach in teaching, limiting opportunities for student exploration, investigation, and in-depth understanding. Furthermore, Klinger and Rogers (2011) conducted a study involving Grade 3, Grade 6, and Grade 9 teachers in Ontario. Their findings suggested that teachers had, at best, neutral perceptions of low-stakes large-scale assessments in terms of their impact on teaching practices. Some teachers expressed uncertainty, while others believed these assessments did not necessarily improve teaching. It is noteworthy that the Elementary Teachers' Federation of Ontario (ETFO, 2000) found that a significant percentage of teachers felt that EQAO tests had a negative impact on their role as educators.

2.12.3 Teacher Association and Organization Views

Several teacher organizations have raised concerns regarding LSAs as well. For instance, the Manitoba Teachers' Society (2010) asserted that large-scale testing lacked educational value due to its complex nature and its inability to adapt to the curriculum and ongoing instruction. The Alberta Teachers' Association (2009), emphasized that accountability should not revolve around static assessment points. Instead, they favored early classroom assessment practices that allowed teachers to tailor their instruction to the specific needs and learning styles of their students. This organization firmly believed that it was the responsibility of teachers to develop assessment approaches that aligned with the curriculum and addressed the diverse learning needs of their students.

These teacher perspectives on LSAs reflect doubts about the credibility of these assessments, which could potentially impact how teachers utilize the findings from LSAs. If teachers do not perceive LSAs as valuable, they may be less inclined to integrate the results into their teaching methodologies, thereby raising questions about the overall utility of LSAs in enhancing teaching and learning.

In 2013, the Physical Education Teachers Federation (PEITF), publicly opposed the budget allocation to LSAs, expressing dissatisfaction with provincial assessments (Wright, 2013). Gilles Arsenault, the leader of the PEITF union, even called for the complete abolition of LSAs to facilitate the hiring of more teachers (Zwaagstra, 2014). Furthermore, in alignment with the opposition voiced by teachers, three out of the four candidates vying for provincial leadership in Prince Edward Island (PEI) openly called for the elimination of LSAs (PEI Home & School Federation, 2015). Similar to the sentiments expressed by teachers in other regions of Canada, the stance of PE teachers on LSAs has had a discernible impact on the trajectory of LSAs. Their vocal opposition to LSAs may have influenced public opinion regarding the value and role of LSAs in education.

2.12.4 Principals' Views on the LSA

LSA data provides school leaders with the opportunity to showcase their leadership abilities, enhance the quality of education, and boost overall school performance. In a study conducted by Volante and Cherubini (2010), all six principals surveyed indicated that LSA data serves as a foundational element for incorporating annual development plans. On a grassroots level, 11 principals expressed their dedication to analyzing data to establish connections between LSA data and the curriculum. Despite acknowledging the impact of political factors on participant responses, these researchers believe that principals recognize the importance of utilizing data to enhance teaching and learning within their schools.

Another investigation into principals' perspectives on LSA in Saskatchewan revealed that the majority (78%) of 65 principals reported a positive influence of provincial assessments on teaching and learning in their schools (Prytula, Noonan, & Hellsten, 2013). These principals play a crucial role in guiding decision-making, participating in initial planning, scheduling, and classroom instruction. This finding was supported by another study in Saskatchewan, which emphasized that many principals viewed provincial LSAs as a vital component of school-level planning (Newton & Viczko, 2010). The principals' opinions on the use of LSAs to steer teaching and learning align with the third function of LSAs, which involves utilizing LSA outcomes to inform education and learning. Unfortunately, there is currently no available study examining principals' perspectives on LSAs in physical education.

2.12.5 Parental Views on the LSA

Only a limited number of studies have explored Large-Scale Assessments (LSAs) from the perspective of parents. In a study conducted in 2004 by Osburn, Ritter, Stegman, and Suitt, they found that in a well-functioning school district in Northwest Arkansas, over half of the parents (104 out of 190) with children in fifth-grade classes considered LSAs to be highly important. Furthermore, more than eighty percent (167 parents) expressed a keen interest in their children's performance on these tests. However, these parents expressed dissatisfaction with the lack of communication and clarity when it came to interpreting the test results.

Another study conducted by Freeman and Williams (2011) examined the viewpoints of a select group of parents on LSAs. They discovered that parents from culturally diverse backgrounds generally had a positive perception of LSAs. They believed that the LSA placement process promoted a fair assessment of students' skills. In Canada, specifically in Ontario, Davie, Hart, and Livingstone (2001) found that 46% of parents (271 individuals) believed that provincial assessments effectively measured student performance and could serve as valuable tools for enhancing student learning at the primary level. At the secondary level, a substantial 76% of parents supported provincial exams and agreed that these results should be used as indicators for high school graduation.

However, according to a 2012 survey conducted by the Ontario Institute for Studies of Education (OISE), only 42% of parents (42 out of 101) were able to identify a significant role for the Education Quality and Accountability Office (EQAO). The EQAO is responsible for administering provincial LSAs in Ontario to assess student achievement in literacy, numeracy, and mathematics at key educational stages. Nevertheless, 63% of these parents reported having read or heard about test results from across the province in the past year.

In contrast, Mu and Childs (2005) conducted a study in which 70% of parents in Ontario believed that LSAs provided accurate assessments for students, schools, and school boards. Moreover, 72.8% of parents felt that LSAs were developed and scored with consideration of students' cognitive levels. Most parents also agreed that LSAs contributed to greater accountability in school programs. Mu and Childs (2005) found that parents sought more information about additional social resources related to LSAs beyond what the school provided. They also expressed the desire for LSA results to be easily accessible, along with sample papers and other resources to support their children's learning at home.

It's worth noting that some of the studies that followed Mu and Childs revealed varying levels of awareness regarding the purpose and value of LSAs. This variation might be attributed to the novelty of LSAs in the Ontario education system, potentially affecting their role in holding schools accountable. Test scores may not have been sufficient to mobilize the parent community to advocate for improvements in the education system.

2.12.6 Student's Views on LSA

Among the limited studies on students' perceptions of LSA, previous studies have shown that the most successful students enjoy LSA due to factors related to intrinsic motivation and a strong belief in success. In contrast, high-achieving students did not endorse LSA because they had little knowledge about them and their method of eliminating LSA was less common than most high-achieving students (Klinger & Luce-Kapler, 2007).

2.12.7 General Public Views on LSA

Recent research indicates that there is a consistent and enduring level of public support for Large-Scale Assessments (LSAs) in education (Phelps, 2005). The Ontario Institute for Studies in Education (OISE) has been conducting a comprehensive public survey called "Public Attitudes to Education in Ontario" every three years since 1978. This survey compiles public sentiment regarding education in the province. In the most recent OISE survey conducted in 2012, which gathered responses from 1,016 adults, it was found that provincial LSAs continue to enjoy substantial backing, with 70% of respondents expressing strong support for provincial Level 2 LSAs. Interestingly, 67% of participants advocated for teacher exams as a means for students to earn grades, rather than relying on LSA scores.

Regarding accountability and LSAs, 64% of respondents agreed that the Education Quality and Accountability Office (EQAO) survey helps maintain accountability within the education system, ensuring transparency to parents and taxpayers. Surprisingly, although only 29% of the public claimed to have some knowledge of EQAO's role in Ontario's education system, 45% reported having read or heard about national test results in the past year.

Moreover, nearly two-thirds of respondents expressed support for national primary school testing in some form. Of these, just over half (53%) favored the current system of measuring individual student results each year, while only 11% supported LSAs, which are utilized by a relatively small percentage of students. A significant majority, two-thirds, believed that the national survey could contribute to improving the quality of education in Ontario.

However, despite the broad public support for LSAs in Ontario, there is a lack of consensus regarding their substantial impact on student success. The proportion of communities that believed LSAs had a significant impact on student achievement decreased from 50% in 1998 to 40% in 2012 (Hart, 2012).

Regular assessments of public opinion on LSAs in Ontario have provided valuable insights into the role of LSAs in education. Although not all stakeholders share a similar view of the impact of LSAs on educational quality, there are enough responses to support ongoing accountability efforts in Ontario. Importantly, there is no equivalent study in the province of Prince Edward Island (PE), making this research essential for understanding public sentiment regarding LSAs in that region.

2.13 Achievement Test

Constructing and validating tests, especially for measuring academic achievement, involves a series of intricate steps and procedures that encompass various concepts and underlying variables. It is essential to follow specific protocols to ensure that the test aligns closely with the intended outcomes. According to Haladyna and Downing (2011), two critical steps in test development are: (i) item development, which includes defining content, preparing test specifications, building an item pool, validating content through expert judgment, conducting pilot tests, and analyzing and revising test items; and (ii) item validation through item analysis. These processes are interconnected and carried out meticulously to create valid and reliable instruments for estimating both item and person abilities. Validity serves as the cornerstone of any assessment system, whether the test is standardized or locally-designed, aiming to produce accurate estimates of examinee abilities that can support valid inferences.

2.13.1 The Quality of National Assessment

The term "quality" encompasses various aspects of students' educational experiences, including safe and well-resourced learning environments, responsive curricula, efficient teaching methods, capable educators utilizing interactive teaching approaches, and the actual learning process by students (Schubert 2005). Nonetheless, in national assessment studies, the primary emphasis is placed on evaluating the quality of education by focusing on the cognitive results of learning, which essentially means assessing what students have acquired, and formulating approaches to elevate these achievements. This priority is in accordance with target 6 of the Dakar Framework for Action, which underscores the need to enhance education quality to attain tangible and quantifiable learning accomplishments (UNESCO 2000).

Acknowledging the paramount importance of student learning in national assessments. Four prerequisites must be satisfied to ensure that (a) the test employed effectively mirrors the accomplishments that schools intend to cultivate, and (b) the acquired data aligns with the requirements of the users (Beaton & Johnson 1992).

First and foremost, it's imperative to guarantee that a test accurately assesses the knowledge and abilities specified in a curriculum or framework, such as reading. This involves taking into account factors such as the significance of the curriculum, the cognitive complexity involved, the appropriateness of the language used, and how meaningful the test items are for students (Haertel & Herman 2005). Tests should encompass more than just evaluating individual skill elements or memorization, as this approach can hinder the cultivation of advanced cognitive capacities such as critical thinking, problem-solving, and the handling of unconventional challenges. Test creators should strive to develop tools that provide insights for policymaking and decision-making, ultimately contributing to enhancements in curriculum and instruction that facilitate the acquisition of valuable knowledge and skills (Mannion et al., 2018).

Secondly, a test should evaluate the knowledge and skills of students at a level that aligns with their abilities. It becomes problematic when a test relies solely on curriculum guidelines that set unattainable standards for student performance. In such instances, the test might prove overly challenging for students with lower levels of achievement, potentially overlooking their achievements. To tackle this issue, the process of test development should encompass not just the standards outlined in the intended curriculum but also the real academic accomplishments of students within schools. Involving practicing teachers in item development and selection, as well as conducting careful field trials with a diverse sample of students, can help ensure that very few students answer all items correctly or incorrectly.

The third essential criterion for obtaining reliable insights into students' expertise and skills within a specific curriculum domain is that their performance should not be affected by their competence in unrelated areas that are not the primary focus of the intended assessment. (Messick 1989). For instance, a science or mathematics test should not heavily rely on language skills to the extent that performance becomes dependent on students' reading ability rather than their knowledge in science or mathematics. This issue arises when it cannot be assumed that all students possess the same level of reading skill, especially when the test's language differs from what they typically use.

Lastly, if assessment results are meant to monitor changes over time, the assessment instruments must be comparable. This can be achieved by using the same test securely between administrations or through scaling with Item Response Theory if different tests are employed. Best practices involve carrying over a subset of items from one test to another to establish a strong link between assessments. Additionally, it is essential to ensure that student samples and administration procedures remain equivalent. Factors such as varying exclusion criteria for students with learning difficulties or differing conditions (e.g., response rates) should be taken into account when comparing students' achievements at different points in time.

Achievement tests gauge an individual's grasp of subject matter knowledge. From a technical standpoint, they assess the practical impact and behavioral alterations brought about by a school subject within a child's overall personality. According to Best (2007), to measures the learner present level of performance or what he/she learned achievement test is conducted. D'Sa and Visbal Dionaldo (2017) stated that more emphasis is seen towards the quality of test used for assessment as limited no of teachers are trained for developing a quality objective test and skills performing item analysis (Mannion et al., 2018). Rehman, Aslam and Hassan (2018) believe that lack of item analysis can have a negative effect on entire test. Through evaluation will get compromised, it will not only cast an adverse impact on the grade of pupil but in the long run on their career and life ahead as well (Reichert, 2011). Studies carried out by Nedeau-Cato, Laughlin, and Rus (2013) and Hijji (2017) concluded that more than 85% of items have at least one or more flaws which strengthen the theory that conducting item analysis is essential for every exam to make it more purposeful and error free. Item analysis involves the examination of four key components: Difficulty Index (DI), Discrimination Index (DI), Distractor Efficiency (DE), and the assessment of reliability through the application of the Kuder-Richardson Formula 20 (KR20). Different techniques were used to identify and maintain the quality of the achievement test.

2.13.1.1 Validity

Validity stands as a cornerstone in the field of psychometrics, as it directly assesses the extent to which an assertion concerning a student, based on assessment data, is solidly supported (Cronbach, 1989). The process of establishing validity entails articulating this justification, scrutinizing the foundational beliefs and theories, and evaluating its robustness and trustworthiness through a range of corroborating evidence. It necessitates the identification of factors that could potentially weaken this justification, exploring alternative explanations for performance outcomes, and incorporating these considerations into the assessment system to minimize inaccuracies in inference.

Cronbach and Meehl (1955) emphasized the importance of construct validation when interpreting a test as a measure of an attribute or quality that is not explicitly defined by operations. Construct validation involves making claims about individuals based on observations, rather than solely describing the observations themselves. Earlier approaches to validity categorized it into different types, such as content validity, predictive validity, convergent, and divergent validity. Various forms of validity can be viewed as distinct lines of rationale and sources of evidence that bolster a singular construct validity.

Embretson (1983) differentiates between two categories of validity arguments: one that elucidates the rationale for why data gathered in a particular manner should provide understanding of the targeted skill or knowledge (known as construct representation), and another that explores the connections between the resultant scores and other factors to support the argument (referred to as nomothetic span). Historically, validation studies predominantly leaned towards nomothetic arguments, utilizing scores from assessments in their ultimate or near-final formats. Nevertheless, the design of these tests was primarily shaped by item format and content criteria, with less emphasis on considerations linked to construct representation, be it theoretical or empirical.

The "cognitive revolution" that occurred in the latter part of the 20th century played a pivotal role in bringing scientific credibility and effective techniques for incorporating the essence of concepts into the initial phases of test development (Embretson, 1983). Presently, both forms of argumentation are deemed valuable, with validation processes rooted in nomothetic breadth having a more established footing, whereas those founded on construct representation are still undergoing development.

2.13.1.2 Reliability

Reliability pertains to the extent to which data adequately support a claim, assuming the warrant is appropriate and alternative hypotheses are properly eliminated. Even if the reasoning is sound, the available data may not provide sufficient information to substantiate the claim. In probabilistic measurement models, reliability can be quantitatively expressed. It is important to note that data collection procedures can involve multiple steps or features, each impacting the evidential value of the data.

Evaluating the sufficiency of evidence frequently entails the concept of replicating the measurement procedure. This concept has played a central role in characterizing the reliability of assessments since the work of Spearman (1904), similar to its role in the physical sciences. For example, if a stone is weighed ten times, and each measurement yields a slightly different result, the variability among the measurements act as a signal of the level of uncertainty linked with the measurement process. However grasping the essence of replicating the measurement process can become more intricate when the procedure encompasses multiple phases that permit variations (such as on distinct occasions, with different tasks, or involving different assessors). It can also be challenging when certain stages cannot be duplicated (for instance, if an individual gains knowledge from undertaking a task, a subsequent attempt does not gauge the same depth of understanding) (Brennan, 2000).

2.13.1.3 Psychometric Instrument

Originally, psychometric instruments were primarily focused on measuring intelligence. The main objective of psychometrics, according to Samejima (1997), is the mathematical modeling of humanbehaviour.

Psychometric tools are used to measure a variety of personality traits, such as early childhood dev elopment, ability, achievement, and intelligence. These tests play an important role in society and must be carefully designed and evaluated. A strong test model should establish a strong relationsh ip between test items and potential scores, allowing good design to achieve the desired score distri bution and control parameters. Psychometricians specialize in developing and refining tests, testin g procedures, data evaluation tools, and methods for understanding and evaluating results. Sameji ma (1997) defines psychometrics as "mathematical modeling of human behavior".

The primary goal of educational and psychological measurement is to assess the cognitive abilities of test takers and represent these abilities as dependable numerical scores. This process entails creating exam questions, deriving observed scores from the test performance, and utilizing these observed scores to estimate the true scores.

When a test demonstrates a sufficiently high level of reliability, the observed score can be considered a trustworthy approximation of the true score. The consistency of measurement hinges on the reliability of the test itself, ensuring that it consistently produces comparable results under various conditions and settings, including different evaluators and testing environments.Validity is another critical aspect of measurement, denoting the extent to which the results of a test can be legitimately relied upon for a specific purpose. Validation is not an optional component but rather a fundamental consideration in the design and evaluation of tests. Validity is established through evidence and theory that support the interpretations of test scores in the context of their intended use (Scott & Mead, 2011). It's important to note that validity applies to the inferences drawn from test scores, not just the instruments themselves. Consequently, validity must be demonstrated for each intended interpretation. Test scores reflect the underlying construct to varying degrees of precision but never with absolute accuracy. Samuel J. Messick introduced a novel approach to understanding validity in a series of publications (1975, 1980, 1989, 1990, 1994a, 1994b, 1995, 1998). He advocated for a unified perspective on validity, with construct validity serving as the overarching concept encompassing various aspects of test validity. Messick argued that the traditional concept of validity, including content validity (verified by experts), criterion validity (e.g., correlation of test scores with another test), and construct validity, has limitations because it does not take into account the broader societal impacts of tests. In light of this, Messick (1995) presented a framework including six distinct aspects of construct validity that should be considered: (1) content validity, (2) substantive validity, (3) structural or internal validity, (4) external validity (5) generalizability, and (6) subsequent validity.

Content validity, which underlies Messick's (1995) construct validity framework, refers to how well different dimensions of a subject are represented in a test. This includes assessing the relevance and technical quality of test content, including aspects such as scoring procedures, rubrics, instructions, wording, and the items themselves. To determine content validity, experts or individuals can evaluate several aspects: (1) the alignment of the test specifications with the construct domain, (2) whether the test items adequately represent the construct domain, (3) whether the content across all related test documents is consistent with the given design, and (4) whether assessment procedures, including rubrics, are relevant to what is being measured. Essentially, content validity relies on expert judgment to determine how effectively the test content mirrors the construct domain.The substantive element involves delving into the cognitive processes that individuals engage in when responding to test items (Lane, 1999). This analysis can be conducted through methods like think-aloud protocols and interviews with test takers, internal pilot tests conducted during test development, and external reviews by subject matter experts (Hamilton, Nussbaum, & Snow, 1997; Magone, Cai, Silver, & Wang, 1994). Detailed descriptions of the cognitive processes expected from test completions, especially in secondary data analyses, provide valuable evidence for the substantive aspect in Messick's framework.

Addressing the internal or structural element, researchers often employ factor structure analyses, including inter-correlations among test items, principal components analysis, factor analysis, and differential item functioning (DIF) (Nussbaum et al., 1997). However, caution is necessary when examining the dimensionality or internal structure of Large-Scale Assessments (LSAs) tend to be better suited for continuous data compared to ordinal data when utilizing traditional factor analysis methods or dichotomous data. For such data, alternative methods like correlations or NOHARM are more appropriate.

The external element focuses on exploring the connection between test scores and other measurements, including concurrent and predictive validity (Messick, 1994). Typically, this involves correlations between two assessments measuring the same or different constructs. Generalizability investigates the uniformity of student achievement across various populations and environments. Researchers can vary the outcomes of test utilization, such as high-stakes versus low-stakes scenarios, and evaluate how these variations affect the interpretation of scores. Finally, the consequential element considers the intended and unintended outcomes of the assessment, encompassing the societal and value implications of score interpretations (Messick,1998). Consequences may include impacts on teaching and learning, student perceptions, and future educational decisions. These consequences can be both positive and negative, depending on various contextual factors and interpretations.

Michael T. Kane introduced an argument-based approach to validation (1990, 1992, 2001, 2002, 2004, 2006, & 2009). Kane's approach involves constructing a validity argument is a rationale that validates the usage of a test and employs various methods to substantiate the utilization and understanding of test scores. The argument comprises two parts: validity and interpretive. Validity entails an overall assessment of evidence supporting the intended uses and interpretations, while the interpretive aspect specifies the questions or assumptions to be addressed during validation. Multiple types of evidence may be required to evaluate specific assumptions, making the interpretive argument an evolving, hypothesis-driven process. Kane's approach places considerable emphasis on the role of validators in proposing interpretations and uses of test scores, giving them flexibility in adopting a narrow or broad view of validity based on their goals.

In contrast, Lissitz and Samuelsen (2007), proposed a more specific "content validity" approach, focusing on the internal structure, practical content, and reliability of the test. They de-emphasized the idea of a unitary construct validity and argued for a dual-part evaluation of the test itself and its relationship with other criteria or measures. Additionally, Borseboom et al. (2009), questioned the concept of a unitary construct validity, suggesting that researchers' interpretations of validity could vary widely, leading to differing score interpretations.

Measurement's purpose is to portray an individual's attributes using valid and well-founded theoretical models while also ensuring reliability. The interpretation of obtained test scores is conducted in a rigorous and scientific manner. Nonetheless, measurement in the social sciences often grapples with the challenge of unreliability, as measuring the same attribute on different occasions may yield disparate results (Revelle, 2013).

The accuracy of collected information is intertwined with the capabilities and modernity of the techniques applied. The nexus between the theory employed in test construction and implementation, as well as the measurement and evaluation of the test, is pivotal for generating realistic and pertinent information (Cook & Beckman, 2006).

Different theories can be applied in the psychometric process, allowing the test to be studied from various angles and the items to be evaluated based on different theoretical frameworks.

2.14 Item Analysis

According to Thompson and Levitov (1985), item analysis assesses individual items concerning external criteria or other items on the test. Norm-referenced tests (NRTs) developed for instructional purposes or educational research, as well as criterion-referenced tests (CRTs) comparing students' performance to pre-established criteria, benefit from item and test analyses. Researchers like Galton, Pearson, Spearman, and Thorndike have significantly contributed to the theory of test item analysis. Items may fail to meet quality standards due to flaws in the question or content instruction (Krishnan, 2013). It is a procedure that involves evaluating both students' responses and test questions to assess the quality and quantity of individual items as well as the overall test. Many test creators aim to verify that their tests' content validity by having external experts assess the items for relevance, difficulty, clarity, and other factors. Although this technique is commonly accepted, a closer examination exposes a number of flaws that must be avoided if professional counsel is to be truly beneficial. Through this, the quality of every item is evaluated by examining the student’s response to a specific item, highlighting the overall quality of the test. Through this method, instructors get a chance to not only increase their construction skills but also help them to identify areas or contents that need greater emphasis/clarity. Item analysis also highlights specific methods that are used to evaluate an item both quantitatively and qualitatively to assist the test developer in enhancing the test by including or removing items that do not meet the minimum acceptable criteria.

Both the qualitative and quantitative review of the item is necessary where on one hand an expert of the material or content help in assessing whether the quality of the item is up to the mark and will be successful in achieving the required target at the same time the quantitative review complete the remaining points that are normally skipped during qualitative reviews such as the multidisciplinary nature of the content or the demographic characteristics of the test taker. These issues are highlighted by statistical data and that is why quantitative analysis is done after the administration of the test (Beck, 2020).

2.14.1 Test Theories and Model

The terms "test theory" and "test models" are frequently used interchangeably, even though they have a specific technical difference. Test theories furnish a comprehensive structure that establishes a link between readily observable variables (Similar to test scores and item scores, there are also latent variables, including true scores and ability scores (Bejar, 1983). However, to be useful, a test theory needs to be fully specified as a concrete model. Test models are specific formulations within the test theory, offering detailed specifications of the relationships between test-theoretic concepts and formulating assumptions about these concepts and their interrelationships. The suitability of test models can be evaluated by examining particular test data conducting empirical investigations, or assessing model fit. Models are always incomplete representations of the test data they are fitted to. Given an ample amount of test data, models can be either disproved or determined to be inaccurate. The focus should not be on whether the model is right or wrong, but rather on whether it fits the data well enough to effectively guide the measurement process. Statistical evidence and judgment play important roles in addressing this question (Bejar, 1983).

Classical test models are often called "weak models" because they have relatively lenient assumptions that can be easily satisfied by test data. However, not all classical test models are weak; some, like the binomial test model, are considered "strong models" due to their strict assumptions about error score distribution. Item response models are also classified as robust models since they impose stricter and less easily satisfied assumptions on the test data (Bejar, 1983). For example, one, two, and three-parameter logistic models presume that test items gauge a single shared trait or ability. On the other hand, classical test models do not make as stringent an assumption; instead, they posit a common factor structure across parallel test forms (Bejar, 1983).

Bejar's (1983) paper delves into two prominent psychometric theories, namely, random selection theory and item response theory. Random sampling theory comprises two methodologies: the classical theoretical approach and the generalization theory approach. Classical Test Theory (CTT) has been a widely employed method over the years to evaluate the reliability and various attributes of measurement instruments. CTT introduces three key concepts: the test score (also known as the observed score), the true score, and the error score (or random measurement error).Hambleton and Jones (1993) formulated a mathematical model referred to as the "classical test model," expressed by Equation (1): $X = T + E$. In this linear model, the observable test score (X) is connected to the combination of two latent variables: the true score (T) and the error score (E).Since the true score cannot be directly observed, it must be estimated from an individual's responses to a set of test items. A student's ability is determined by the number of correct answers (Bichi, 2016). However, solving the CTT equation requires certain simplifying assumptions. Central assumptions within CTT involve the independence of true scores and error scores, the average error score among examinees being zero, and the absence of correlations among error scores on parallel tests (Hambleton and Jones, 1993).

Classical test analysis (CTA) employs traditional item-specific and sample-related statistics, which encompass assessments of item difficulty, item discrimination estimates, analyses of distractors, and related measures (Bichi, 2016). CTT analyses predominantly focus on examining evaluations at the test score level rather than at the item score level, which includes measures of test-level reliability (test-level statistics).

Classical test theory (CTT), alternatively known as classical true score theory, is a simple model that elucidates how measurement errors can impact observed scores (Marcoulides, 1999). It primarily concentrates on estimating the reliability of test results. The true score represents the average of all observed scores gathered under different conditions and with different items, providing the most unbiased estimate of an individual's ability. In each instance of test administration, the observed score is likely to deviate from the true score due to random error (Suen, 1990). In accordance with CTT, every test score is regarded as a realization of a random variable X, consisting of a true score (T) and an error score (E), as denoted by the equation X=T+E (Steyer, 1999).

In classical test theory (CTT), the true score is not directly observable and must be estimated from an individual's responses to a set of test items. CTT posits that observed scores come with error, and the objective is to minimize this error in measurement development (McBride, 2001). Hence, the test's reliability and the calculation of the reliability coefficient are essential. The knowledge of the reliability coefficient enables the estimation of error variance. The square root of this error variance is known as the standard error of measurement (SEM), which assists in establishing a confidence interval for a more precise estimate of the true score.

In classical test theory (CTT), reliability is viewed as an attribute of the test data itself rather than the assessment. As per APA (1999) standards, when reporting reliability, it's imperative to provide details about the methods used for coefficient calculation, sample characteristics, and data collection conditions. Nevertheless, the reliability estimates derived from these procedures are sample-dependent and have practical limitations when designing or evaluating technology-enhanced assessments (Scott, 2010).

2.14.1.1 Classical Test Theory

Item response theory and random sample theory are two important psychometric theories for the analysis of measurement techniques, according to Bejar (1983). There are two distinct approaches to random sampling theory: the classical theory approach and the generalized ability theory approach. Classical True Score Theory (CTT) is a simple model that provides an explanation of how errors in measurement can influence the scores we observe (Marcoulides, 1999).The first theory of measurement is called classical test theory (Gulliksen, 1950), and it has been used for many years to analyze the psychometric properties of personality assessments. Estimating the reliability of test results that have been observed is the primary objective of this theory. When a test is administered to a specific set of items, at a particular moment, under reliable conditions, it yields an observed score for the examinee. Under all conceivable conditions, across various instances, using all similar possible items, the average of these observed scores serves as the most impartial estimate of the individual's true ability, and this average is referred to as the true score. However, in any single instance of test administration, the observed score is likely to deviate from the true score, a variance known as random error score (Suen, 1990). Within Classical Test Theory (CTT), each measurement (test score) is considered as a value of a random variable X comprising two components: a true score and an error score (Steyer, 1999).

2.14.1.2 Item Response theory:

The basic goal of educational and psychological assessments is to identify a test-takers ability to think and portray them as an accurate numerical result. However, the benefits and outcomes of education cannot be experienced in our educational system unless they are measured. This demonstrates how measurements have distinct functions in education, aiding in decision-making and providing feedback. Except in cases where there is no concern for feedback, measuring in education cannot be completely disregarded. The primary goal of the measurement is to identify, assess, or rate people in order to gather crucial data that may be used to various sorts of actions or choices. Measurement must provide extremely trustworthy and secure information in order for these actions or judgements to be appropriate for their intended use. Item Response Theory (IRT) is a statistical framework used to analyze and model the response data obtained from a questionnaire or test. It is widely used in educational and psychological research to measure the latent traits (abilities, attitudes, and knowledge) of respondents based on their responses to a set of items. Item Response Theory (IRT), as a tool for the analysis of educational assessments, has now found application in diverse fields, spanning sociology, political science, psychology, human development, business, and communications. (Templin, 2012). The CTT framework was initially the most popular one for generating and analyzing standardized examinations. Since the early 1970s, IRT has predominantly supplanted CTT as the primary theoretical framework utilized in this scientific domain (Wiberg,2004). IRT permits users to define a mathematical function for describing the connection between an underlying trait, denoted as θ, and the likelihood that a test taker with a specific θ will answer a test item correctly. In the period preceding the 1980s, IRT research primarily focused on tasks such as estimating model parameters, assessing the fit of the model to data, and applying these models to various testing scenarios involving multiple-choice items with binary scores (yes/no, 1 or 0). Performance evaluations, polytomous response formats, and multidimensional qualities were all studied in depth, as was work on computerized adaptive testing. As a result of this broader focus, a wide range of new IRT models emerged, allowing researchers to handle complicated challenges not just in achievement testing but also in fields such as attitude, personality, cognitive, and developmental evaluation (Gierl, & Bisanz 2001).

2.14.1.2.1 The One-Parameter Logistic (1PL) Model

The most frequently utilized and easiest of the three IRT models, according to Ogunsakin (2015), is the one-parameter model. The sole parameter in the IRT-1PL model that the model estimates is the difficulty parameter bi.

2.14.1.2.2 Two-Parameter Logistic Model According to David (2011), the two-parameter model serves a similar purpose as the one-parameter model, but it introduces variability in both item discrimination and item difficulty parameters among different items. The key distinction of the IRT-2PL model from the 3PL model is that it assumes pseudo-guessing does not significantly affect item fit, or more commonly, it does not apply to the data, such as in cases involving rater-scored constructed-response items.

2.14.1.2.3 Three-Parameter Logistic (3PL) Model In these models, three parameters - difficulty, discrimination, and guessing - are employed. The inclusion of a guessing parameter is particularly advantageous for assessments like true-false and multiple-choice tests in the three-parameter model.

2.14.1.2.4 The Four-Parameter Logistic (4PL) Model Expanding on the standard three-parameter logistic (3PL) model, the four-parameter logistic (4PL) model introduces an upper asymptote that may vary from 1, as highlighted by Kalolina (2013). The four-parameter logistic model (4PLM) is built on the assumption that even highly proficient test takers can make errors, sometimes due to factors like carelessness.

2.15 Comparison of CTT and IRT

IRT and CTT are two approaches used in educational measurement, but they differ in their fundamental estimation of person ability. CTT is test-centered, meaning that the focus is on the overall test score and its relationship to the examinee's ability. On the other hand, IRT is item-centered, which means that it focuses on the connection between how a test taker responds to individual items and their level of ability.

One key difference is that two examinees who have the same total number of correct items on a CTT-based test may not be assigned the same ability estimate with IRT. IRT provides more flexibility in estimating person and item parameters independently, whereas CTT does not offer the same level of precision.

IRT models may be more mathematically complex and harder to interpret compared to CTT. They also require larger sample sizes to obtain accurate and stable parameter estimates, although some IRT models like the Rasch one-parameter model can work with moderate sample sizes.

However, IRT offers several advantages in educational measurement. It allows for the construction and assessment of measurement instruments, estimation of valid and invariant person and item parameters, and flexibility in solving measurement problems. One significant advantage of IRT is its ability to estimate item parameters independently of examinee characteristics and estimate individual abilities independent of item characteristics. This allows for person-free item calibration and item-free object measurement, which are desirable in scientific measurement.

IRT also enables the development of fair and equitable measurement scales and provides targeted measurement of individual abilities. It allows for the selection of suitable items to extract maximum information about an individual's ability level. Additionally, IRT can be applied in various measurement contexts such as test equating, tailored testing, adaptive testing, and differential item functioning analysis. While CTT is test-centered and simpler to operate, IRT offers advantages in constructing measurement instruments, estimating person and item parameters, and providing flexible and targeted measurement. It allows for person-free item calibration, item-free object measurement, and the development of fair and equitable measurement scales. Students with varying levels of knowledge and application skills in the subject matter.

Classical Test Theory (CTT) offers a straightforward approach to item analysis, making it suitable for small-scale exams or groups without extensive psychometric knowledge. Item difficulty is quantified using proportions, averages, counts, and correlations.

For dichotomous items, item difficulty is determined by the proportion (P value) of examinees who answer it correctly. A higher P value (e.g., 0.95) indicates an easier item, while a lower P value (e.g., 0.35) suggests a more difficult item. Sometimes, item difficulty is referred to as item facility due to the inverted scale (lower value = higher difficulty). Polytomous items are assessed based on the mean score. For example, if an essay item is scored on a scale of 0 to 5 points, an average score of 1.9 indicates difficulty, while a score of 4.1 suggests ease.

Item discrimination is a desirable trait in psychometrics, as it allows for differentiation among examinees. The point-biserial item-total correlation (Rpbis) is the primary statistic used in CTT for evaluating item discrimination. It measures the correlation between scores on the item and the total test score. A strong item that effectively measures the topic will show higher-scoring examinees also answering the item correctly, resulting in a correlation of 0.20 or higher. If the correlation is around 0.0, the item lacks discrimination and is considered ineffective for the exam.

Item Response Theory (IRT) is a sophisticated method for item analysis that deals with a range of psychometric functions, including item analysis, equating, and adaptive testing. Implementing IRT necessitates substantial sample sizes (typically ranging from 100 to 1000 responses per item) and demands expertise, often provided by a Phd level psychometrician. This approach is not well-suited for small-scale assessments like classroom quizzes but finds extensive application in larger-scale exams such as K-12 benchmark assessments, university admissions tests, and professional certification examinations.

2.15.1 Paradigm Shift from CTT to IRT

Despite its simplicity in terms of mathematical operation, CTT has served the measurement community for more than 50 years and accomplished more than anticipated. Due to CTT's failure to address several measuring issues in the research community, IRT was established as its substitute. The limitations of CTT models are confirmed by Minh (2004), who urges on measurement experts to create new frameworks to get over these restrictions. Among other things, these restrictions include the following:

The test score and the real scores are test dependent, while item difficulty and item discrimination are group dependent. This implies that the test taker's result is dependent on the particular collection of test items used. Minh inferred from this limitation, that the person being tested typically scores lower on tough examinations and higher on simpler ones, despite the fact that his/her ability is consistent across those tests. The fact that CTT models are predicated on the idea of uniform measurement errors among examinees is another drawback of CTT. The concept of parallel testing is yet another significant CTT drawback that IRT resolves. According to the literature review, CTT is test-oriented rather than item-oriented. This demonstrates how challenging it is for researchers to forecast and establish how each test question will be responded to by the examinee. These cause a paradigm change from CTT to IRT in order to get beyond its limitations. The framework for Item Response Theory (IRT) that supersedes the limitations of Classical Test Theory (CTT) includes characteristics such as item statistics that remain consistent regardless of the group, scores that depict the examinee's proficiency independently of the specific test, a mechanism capable of addressing issues related to parallel testing, and a model capable of establishing connections between items and the examinee's aptitude. Similar to CTT, IRT starts with the premise that a test taker's reply to a certain item is influenced by an imperceptible mental characteristic of that test taker. The terms "trait" and "ability" are used to describe such quality. Latent traits or latent abilities are characteristics that cannot be straight away detected. IRT tries to describe how a test taker's latent ability and the likelihood that they will properly answer a particular test item relate to one another. According to (Minh 2004), an item characteristic function or item response function is used to model this connection.

2.15.2 Application of CTT in Measurement in 21st Century

For many years, Classical Test Theory has been employed as a method to assess the reliability and other attributes of a measuring instrument. Classical test theorists utilize conventional items and statistics that depend on the sample, this includes various like item difficulty, item discrimination estimates, distractor analysis, item-test inter-correlation, and more. As per Freeman, item difficulty serves as an estimate of the level of skill required to successfully answer an item. On the other hand, discrimination power represents the comparison of the high score with a low score on the same test.

2.15.3 Item Analysis Procedure

The performance of the test takers in a summative assessment was examined to ascertain the difficulty index (DI) and item discrimination (DI) of multiple-choice items. The item difficulty index is computed by taking the percentage of correct responses among the total number of responses.

A higher difficulty index indicates easier items. An acceptable range for the difficulty index is between 20% and 90%, with items between 40% and 60% considered excellent. Items with a difficulty index below 20% (too difficult) or above 90% (too easy) require modification. The difficulty index represents the relative frequency of correct responses rather than an intrinsic characteristic of the item (Hotiu, 2006).

The item discrimination index is computed as the point biserial correlation, measuring the relationship between answering the item correctly and the overall score on all remaining items. This index is employed to evaluate how effectively an item distinguishes between students who achieve higher and lower test scores. Guidelines by Ebel (1979) categorize items based on their discriminating indices, with values above 0.40 considered satisfactory, 0.30 to 0.39 indicating little or no revision required, 0.20 to 0.29 suggesting marginality and need for revision, and values below 0.19 suggesting elimination or complete revision.

It's important to consider the context of the test when interpreting item discrimination indices. Items with low discrimination may be ambiguously worded and should be examined. Negative discrimination indices should be investigated to understand the reasons behind them. Factors influencing item discrimination include item difficulty and the item's purpose in relation to the overall test (Mehrens & Lehman, 1991).

2.15.3.1 Item Difficulty: This assesses if a test question is too easy or too hard. Items with difficulty below 30% or above 90% should be revised or replaced as they don't effectively differentiate between students who know the material and those who don't.

2.15.3.2 Item Discrimination: This measures if a test question effectively distinguishes between students who understand the material and those who don't. It's crucial for an item to separate Item discrimination measures the extent to which a test item differentiates between high-scoring and low-scoring students. It focuses on whether the right people are answering an item correctly, rather than just how many. Two indices are commonly used: the discrimination index (D) and discrimination coefficients. The discrimination index is calculated by comparing the proportion of correct responses in the top 27% of students to that in the bottom 27%. A positive D indicates higher-scoring students answered correctly, while a negative D suggests the opposite. Discrimination coefficients, such as the point biserial correlation coefficient (rpbi), consider every student's response. A high rpbi indicates that higher-scoring students choose the correct answer, while low or negative values identify problematic items. Ebel and Frisbie (1991) offer guidelines for interpreting discrimination indices, with values above 0.40 indicating satisfactory item performance. To assess test reliability, internal consistency measures like Cronbach's Alpha or KR-20 are used to evaluate how well items within a test measure the same underlying concept.

2.16 Validity and Reliability in CTT Model

Assessments (LSAs) pose a challenge in achieving comprehensive validity and reliability because attempting to create a single assessment that effectively measures all significant dimensions can be demanding as these aspects often compete or conflict with each other (Zhao, 2018). Finally, the substantial costs associated with developing and frequently administering numerous LSAs pose practical limitations. Consequently, LSAs have predominantly focused on assessing skills and knowledge in limited areas, typically emphasizing cognitive abilities like mathematics, language, and science (Emler et al.,n.d).

Validity in research refers to how well a measurement instrument truly evaluates the specific attribute or concept it is designed to measure. In the Classical Theory of Measurement, three approaches are utilized to assess the accuracy of a research instrument.

Face validity pertains to the subjective assessment made by researchers regarding the instrument's apparent suitability and relevance. It involves determining whether the instrument's items appear to be pertinent, reasonable, unambiguous, and clear. While some scholars question the true value of face validity as an indicator of validity, it can still be quantitatively assessed by involving experts in the field of study (Oluwatayo, 2012).

Construct validity involves gathering evidence from various studies employing a particular measuring tool. This process entails investigating the connection between the measure being assessed and variables that are either established to be related or are theoretically linked to the construct being measured. For example, when assessing the quality of life, one would anticipate obtaining lower scores from chronically ill individuals as opposed to healthy college students. Correlations that conform to the anticipated patterns enhance the body of evidence supporting construct validity.

Criterion-Related Validity: Criterion-related validity provides insights into how well scores on the new measure correlate with other measures of the same construct or very similar underlying constructs that should theoretically be related. It is crucial that these criterion measures themselves are valid. Predictive validity and concurrent validity are two types of criterion-related validity (Carole & Winterstein, 2008).

In essence, these methods serve as essential tools for ensuring that a measurement instrument accurately captures the intended construct or concept without introducing bias or errors.

Reliability, according to the Classical Theory of Measurement, acknowledges that any score comprises both the "True Score," which remains unknown, and an "Error Score" introduced during the measurement process. Various methods exist for estimating the reliability of a measure:

Reliability as Equivalence: This involves two forms of reliability—alternate or parallel form and inter-rater form. Alternate or parallel form reliability necessitates the development of two instrument forms with the same content domain, test specifications, number of items, item format, and similar difficulty and discrimination indices.

Test-retest reliability measures a test's consistency over time by comparing results from multiple test administrations to the same group of individuals.

Internal consistency assesses the agreement among items within a single test, assuming that items measuring the same concept should correlate. Common methods include Cronbach's Alpha, Split Half, and KR-20 & 21 (Carole & Winterstein, 2008).

2.17 Reporting and Dissemination of Data

According to Best (2007), purpose of conducting test is to calculate the learner present level of performance. One should considered the following points during reporting,

• Use simple language, best for attractive media products such as videos.

• Clearly identify participants, tailor events and products according to their needs.

• Consider hiring community leaders and trustees as lawyers

• Dissemination of information to the media. Events must be well advertised.

• Use simple, easy-to-understand slogans and messages.

• Make a backup copy of all broadcast or distribution material to a wide audience (including PowerPoint presentations) with technical support information.

Reporting involves the analysis and interpretation of learning assessment data, with the aim of preparing findings for dissemination. This process serves the purpose of understanding how these findings align with predefined policy objectives and program goals.

The reporting process encompasses:

- Why? Understanding findings in relation to program policy goals.
- What? Describing the outcomes and investigating the connections between outcomes and the surrounding contexts.

- How? Presenting achievement levels in various formats, such as averages, performance levels, or subgroup breakdowns, both as snapshots at a specific point in time and as trends across assessment cycles.
- Who/What? Reporting occurs at multiple levels, including individual student, school, subnational, and national or cross-national levels.

Dissemination involves:

- Preparation of Dissemination Products: Creating one or more materials that convey the assessment findings.
- Informing Targeted Audiences: Ensuring that the intended audience(s) receive and understand these findings. Dissemination products and platforms can include analytical reports, policy briefs, media reports, press releases, an assessment database, blogs, and social media. Dissemination should be guided by an effective strategy, incorporating various approaches to cater to diverse audiences.
- Use of Data for Policy and Practice: Policymakers use the reported and disseminated results for various purposes:
  - Monitoring and evaluating education policies and developing monitoring mechanisms.
  - Implementing policies or enhancing their effectiveness.
  - Raising awareness about critical issues and prioritizing them for reform.
  - Formulating or designing policy options.
  - Holding stakeholders accountable for their roles in the education system.

2.17.1 Principles for designing and utilizing data

Principles for designing and utilizing national data platforms encompass several key elements. Firstly, a clear institutional framework must be established, delineating responsibilities for the development, implementation, and upkeep of the National Reporting and Dissemination Platform. This should be guided by relevant laws, guidelines, and standard operating procedures, often with the National Bureau of Statistics at the helm. Secondly, the platform's objectives, scope, and functionalities should be collaboratively determined with input from all stakeholders and users, ensuring alignment with both local and international statistical standards. Thirdly, sustainability considerations are paramount, with decisions on purpose, scope, and resources factoring in long-term capacity and support, both internally and externally. Lastly, adherence to international and national statistical standards is essential across all components of the platform to facilitate integration, customization, and enduring growth.

2.17.2 Challenges in LSA Reporting and Assessment

The primary challenge is how to effectively utilize assessment results to inform stakeholders and drive improvements. The study reveals that the lack of human resources is the most frequently reported issue (20 out of 24), followed by insufficient technical capacity and infrastructure (17 out of 24), and limited financial resources. It is crucial to leverage results from international, regional, and national assessments to inform policy decisions and empower stakeholders, including parents, teachers, and school administrators, to enhance their practices. Moreover, comprehensive dissemination of learning assessment results to the general public enables valuable feedback for policymakers.

Overall, the findings underscore the significant challenges in terms of inadequate human resources and technical capacity for data utilization and dissemination, mirroring the issues faced in data analysis. Previous studies have emphasized the necessity for countries to train national experts and professionals not only in operating assessment processes but also in reviewing and critiquing assessment designs, improving test instruments, and employing advanced statistical analysis. By doing so, learning assessment results can be effectively employed to enhance education quality, improve system efficiency, monitor student performance and school effectiveness, and identify key intervention areas (Ho, 2012).

Despite the existence of national and international assessments (LSAs) in the region, the results are not effectively guiding policy changes to enhance education systems. Most participants prioritize creating summary reports of LSAs and struggle to connect these assessments to policy decisions. Instead, they tend to focus on isolated issues like inadequate teacher training or multigrade classrooms. Consequently, countries are not fully leveraging their investments in LSAs to improve education.

Establishing LSAs as a Consistent Monitoring Tool: LSAs lack a stable and well-defined position in the participating countries. Their presence often depends on opportunistic factors, such as donor funding availability, or is ad hoc, with primary grades receiving more attention than others. This lack of stability hinders countries' ability to compare data across assessment periods and monitor the impact of policies effectively.

Insufficient Structural Support for LSA Utilization: The strategic utilization of LSAs appears unclear to both education planning and assessment department representatives within the Ministry of Education. When asked to identify key policy stakeholders for whom evidence should be prepared, many individuals showed little involvement in such activities. Additionally, some participants, even within assessment units, face challenges in conducting basic statistical analyses or ensuring data quality and compliance with standards. Delays in publishing reports are common, highlighting the need for experts to oversee the process, including vendor selection, contract management, evaluation, and quality assurance within the Ministry.

2.17.3 Approaches Used During Reporting

Analytical techniques employed for benchmark reporting (Mullis, 2012) are designed to illustrate performance standards. Benchmarks themselves are not direct indicators of student performance relative to each other. In extreme situations, it's conceivable for all students to meet or surpass a particular benchmark. These benchmarks are intricately tied to performance standards. For instance, a high benchmark on an international scale corresponds to each skill level, with these levels describing what students generally need to achieve and demonstrate to be classified as top performers. In each country, the analysis was performed to report the percentage of students accessing each international benchmark and the trend in the percentage of children accessing them. Regarding the connection between performance and contextual elements, IEA studies also highlighted that contextual factors across various themes, collecting data from diverse questionnaires. Furthermore, in 2011, TIMSS and PIRLS introduced policy-aligned scales encompassing aspects such as home-based learning, available learning resources, and conditions for teachers, the school environment, and students' attitudes toward learning for the first time. The surveys compared areas on these scales that are consistent with policy and practice (Mullis et al., 2012). It primarily focuses on assessing the measure's performance and examining variations in measured performance over time and across groups categorized by different areas of interest. It does not involve conducting correlation, variance, regression, or other complex analyses. At SACMEQ, student and teacher performance is assessed using the Rasch model object response theory. A single worksheet was created to display the international outcomes of SACMEQ III. It contains average scores that were computed and compared among various countries, as well as across gender and socioeconomic status, as determined by the moderate indicator within the SES category. Additionally, the worksheet includes information about schools, categorizing them according to the eight skill levels identified within each learning and mathematics domain (Hungi et al., 2010).

In a separate worksheet dedicated to international outcomes, a comprehensive analysis has been carried out to pinpoint significant student and school-related factors that impact academic performance.This analysis also scrutinizes variations between school and inter-school accomplishments, as well as evaluates disparities concerning social and gender factors. On success compare after controlling other factors that contribute to success (Hungi, 2011). Some worksheets from SACMEQ III list frequencies of varied content of interest, but do not link these contextual and functional aspects (Hung, et al., 2011). SACMEQ national level reports categorize achievement scores based on specific factors of interest, including region, gender, school location (rural/urban), and socioeconomic status (low/high) (Monyaku & Mmereki, 2011). These reports also analyze how the assessed scores have evolved across various SACMEQ assessments. Additionally, they identify the percentage of schools that fall into one of the predefined eight skill levels for each of the two domains. In LLECE reports, test results are presented using a unified, uninterrupted scale for each subject area. This scale is constructed by applying the Rasch model and item response theory. LLECE uses hierarchical line modeling to analyze aspects related to student achievement, to summarize results. Hierarchical line modeling is a complex type of regression of small standard squares (OLS) used to analyze variability in variable results when predictive variables are at different hierarchical levels; for example, students in a class share diversity depending on their general teacher and the general class.

2.18 Reporting and Dissemination of LSA Worldwide

2.18.1 At International Level:

Regarding the dissemination of International Assessment (IA) results, most countries share the assessment findings by publishing reports in hard copy or online formats and providing feedback to students, parents, and teachers. Additionally, many countries publish the results on the internet or through press releases and organize seminars, workshops, or conferences for various stakeholders. Only a few countries utilize social media to disseminate the results.

Regarding the intended audience for disseminating IA results, all nations share their findings with policymakers, school principals, and educators. The majority also report the results to academicians/researchers and local government officials, while some countries report to donors/partners, students, and parents. From a wider standpoint, the primary objective of IAs is to assess students' cognitive abilities and provide valuable insights for the improvement of educational policies. These assessment results serve as vital empirical information that nations depend on for the examination and enhancement of their educational curricula.

OECD in PISA is entrusted for compiling of overall results in a shape of an international report. These reports are used for providing analytical results, feedback ad recommendations on these results. PIRLS and TIMSS also adopt the same method where international reports are made by their respective study centers. These research centers also prepare encyclopedias, whose role is to present data from curriculum questionnaire answered by representative of participating countries. This data is not for analysis purpose but to be there for comparison as and when needed. This encyclopedia also include separate chapters for each participating countries that provide details of their curriculum, language, policies and teaching methods.

IEA in PIRLS and TIMSS also cover the technical aspect of its evaluation in their technical reports, which are published online. All these reports, results, encyclopedia etc. can be downloaded from PIRLS and TIMSS international study center website. IEA data center also created IEA international data Countries participating in these reports also publish their own reports for their policy makers, govt and educators etc. where not only national results are shared but also highlights issues of common interest in particular education system. A base analyzer and IEA data visualizer software application for facilitation of analysts.

2.18.2 At Regional Level: In terms of reporting and dissemination of Regional Assessment (RA) results, the assessment findings from all ten participating countries are disseminated through various means, including the coordination of seminars, workshops, and conferences, as well as the distribution of printed reports to relevant stakeholders. Additionally, more than half of these countries choose to release the results either on the internet or through press releases, ensuring that feedback reaches students, parents, and educators. A smaller subset of countries goes a step further by making the reports accessible online or by sharing the results through social media channels.

Regarding the target audience for reporting RA results, all ten countries report the assessment findings to policymakers, donors/partners, school principals, and teachers. Additionally, nine countries report the results to local government officials. Academicians and researchers are included in the reporting target audience for seven countries, while students and parents are included for six countries.

Much like international assessments, the analysis of RA (Regional Assessments) data serves the purpose of pinpointing disparities in students' cognitive capabilities and understanding the factors that impact these abilities. The primary focus for the majority of countries is leveraging RA data to enhance the professional development of educators. Nevertheless, this data also finds utility in a range of educational interventions, spanning from policy enhancements to the implementation of programs targeting students or specific areas of learning.

At regional level SACMEQ is one such organization that is working on making reports and working papers highlighting scoring trend, common errors and other issues in different subjects and divide the same on the basis of districts and other subgroups e.g. gender school location etc. all this data and reports are readily available on SACMEQ website. SACMEQ coordination center also work in publishing international working papers with regional level comparison. All those countries participating in SACMEQ prepare a national report, wherein all the aspects of education system e.g. background of education system, research management, and contextual information, student and teacher performance and policy recommendation all are allocated separate sections. This way SACMEQ is contributing a lot in providing assistance for compiling national reports. In many instance it has been countries used a source versions, what written in SACMEQ. Not only national report, SACMEQ also conducts workshop which focus is on specific parts of national reports. This type of assistance was required by the participating countries who wanted to have and in-depth knowledge of how to disseminate research results to stakeholders at regional and local level.

As far as, LLECE is concerned, the reporting method is divided into phases. In the first phase a general report is provided then in the second phase a more detailed report highlighting the contextual aspect. The main aim of this detailed report is to identify both contributing factors and potential adjustments to education policy changes.

2.18.3 At National level:

Regarding the utilization of National Assessment (NA) results, most of the surveyed countries (19 out of 21) reported using the assessment findings to evaluate and update education policies and curricula. Furthermore, a significant emphasis is placed on using NA results for the professional development of teachers, with 17 out of 21 countries highlighting this as a key priority. Additionally, 14 out of 21 countries indicated that these results are employed to implement intervention programs aimed at students or schools for specific themes or learning areas.

Similar to Regional Assessments (RA), the majority of these nations disseminate NA results through the organization of seminars or workshops (17 out of 21) and the publication of hard copies of reports (16 out of 21) for various stakeholders. Approximately half of the countries make the results (11 out of 21) or the reports (10 out of 21) available online. However, NA results are not commonly shared through social media channels (6 out of 21).

The primary recipients of NA results reporting include policymakers (20 out of 21) and school principals and teachers (19 out of 21) which aligns with the patterns observed in International Assessments (IA) and Regional Assessments (RA). Moreover, more than 80 percent of these countries also disseminate the results to academicians and researchers (17 out of 21).

In summary, the characteristics of NA data, including its objectives, utilization, methods of dissemination, and intended audiences, closely mirror those of IA and RA. The majority of countries leverage NA data to inform the development of national education policies, particularly by assessing students' cognitive abilities and monitoring changes in their performance over time. Consequently, NA results play a pivotal role in the review and enhancement of education curricula and in strengthening teacher capacity, with a secondary focus on organizing intervention programs for specific themes or learning areas.

2.19 Conceptual Framework for Reporting and Disseminating of Large-Scale Assessment

2.19.1 Conceptual Model for Reporting and Dissemination Assessment Results

A theoretical framework for using data to support decision-making in the management of educational districts was created by Breiter and Light in 2006. They define decision-making as a "very complicated, individual cognitive process that can be impacted by many external influences," which is at the heart of their concept (Breiter & Light, 2006). They disapprove of the idea that decision-making requires a myriad of different types of data, and instead advocate for logically reducing (collecting and organizing) large amounts of data, summarizing and analyzing the data to create information, and then turning the information into context-related knowledge to guide action (prioritizing and synthesizing).

Their model comprises four fundamental elements: data, information, knowledge, and decision-making. This approach acknowledges the diverse range of data and data sources that educational decision-makers encounter, without exclusive emphasis on assessment data. To enhance the quality of teaching and learning within schools, mostly studies adapted and developed a conceptual model based on Breiter and Light's methodology to effectively convey assessment outcome. Data must be reported in a way that enables important users to understand the data, such as district officials, school administrators, and teachers (Coburn, Honig & Stein, 2009).

The model defines information as data that has been properly analyzed and summarized to provide insight into the nature and scope of the recognized problem. Therefore, any report must convey pertinent information that either expands on what is already known or enlightens a new area of interest or requires more research.

By combining what is new with what is already known or available to improve the unpleasant condition and considering what the priorities are, knowledge expands on the information that is already available. The model's term for information is data that has Making decisions involves applying learned information to a situation in order to influence it in a desirable way and, in the case of knowledge derived from assessment, to enhance learning outcomes.

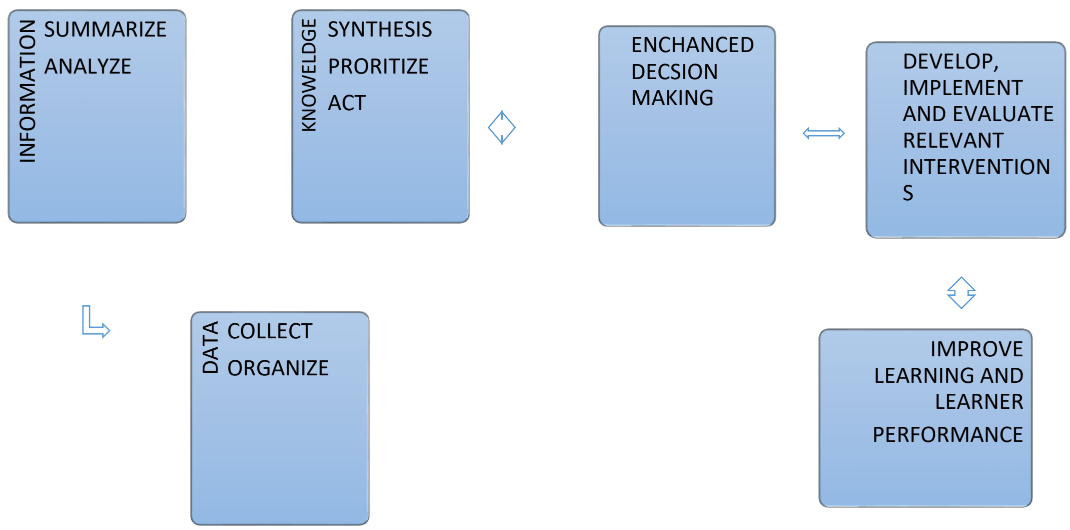


Figure 2.1 Framework for reporting and dissemination of LSA

A nation's participation in substantial educational evaluations and their potential impact on educational policy becomes more meaningful when there's a solid comprehension of assessment among the population. This understanding largely relies on the quality and value of available information. To ensure this level of understanding, sponsors and test planners should make an effort to disseminate information consistently throughout the entire duration of the project, not just waiting until the final results are available. This proactive approach ensures that the data generated from the assessment is better utilized for understanding the test's purpose and implications.

Although reporting typically takes place toward the end of a project, the assessment process can be more effective in terms of both public acceptance and its potential impact on policy if information dissemination takes place before, during, and after the evaluation, as emphasized by Best et al. (2013). The World Health Organization also employs effective dissemination strategies to ensure the proper circulation of data

Examine previous reporting and dissemination initiatives.

↓

Establish goals and aims.

↓

Define target audiences.

↓

Recognize the requirements of the audience.

↓

Select methods and deliverables for dissemination

↓

Specify dissemination platforms.

↓

Assess existing resources.

↓

Take into account timing and optimal periods for action.

↓

Assess the effectiveness of the endeavor.

(**World Health Organization, 2014**)

Figure 2.2 Dissemination strategy by WHO

2.20 Important Document that Disseminated at International Level

2.20.1. Assessment Frameworks

The publication of assessment frameworks serves as a pivotal step in securing the endorsement of assessments among a diverse range of stakeholders, including academic staff, policymakers, researchers, funders, and educators. Typically, these assessment frameworks for large-scale international evaluations are made public either before or concurrently with the assessment's implementation (Mullis & Martin, 2013; OECD, 2013). These documents are typically characterized by an exploration of contemporary viewpoints on the subject, a delineation of the test's focal areas, an outline of the content to be encompassed within the examination, a specification of the expected formats for responses, and a preview of how the results will be presented.

2.20.2 Questionnaire Frameworks

The questionnaire framework typically serves as the foundation for developing contextual questions. Having a list of contextual questions is crucial as it aids in identifying significant background factors that could be linked to student success. For instance, in PISA 2012, a questionnaire was released alongside the learning, mathematics, and science framework (OECD, 2013).The questionnaire outlines the overall purpose and objectives of the PISA policy and what kind of contextual information is required to achieve these objectives. PISA must address various aspects of student / family, school and system levels. Equality concerns demand specific consideration. The examination framework serves as the foundation for all testing procedures. It offers a transparent overview of the elements and topics encompassed by the evaluation, the intended audience, the means of assessment presentation, and the estimated assessment duration. Test frameworks serve as the cornerstone for every evaluation. Each framework delineates concerns that need attention and constructs a corresponding response.

2.20.3 Sample items

An effective way to provide more insight into the test, your goals and your target audience to have easily accessible sample material. This is usually done on a website or in a publication dedicated to the purpose. Generally, most of the test material is kept safe until the next test. This allows the same items to be used for two or more experimental controls to determine changes in outcomes at the time of intervention. However, other items may be released as test samples. Another potential source of sample materials can be items that were initially considered suitable but were ultimately not included in the testing phase. This can occur when creating several items that share a similar type, format, or level of difficulty. Typically, a sample of these items is provided along with comprehensive descriptions indicating the specific area of the curriculum to which an item relates, the skills it assesses, and occasionally, the international scores attributed to the item.

2.20.4 Questionnaires

Numerous major international and national assessments incorporate sets of questions that can be administered to students, teachers, school principals, or parents. These questionnaires serve the purpose of gathering data concerning the environment and circumstances in which the students taking the assessment are situated. The information collected can offer insights into various aspects, including interpersonal skills and the student's background, such as socio-economic status and language proficiency. Additionally, it can shed light on the teaching methods employed by educators and provide details about the school, such as its location and size. These questionnaires are typically made available to the public because they do not assess performance but rather provide valuable contextual information.

2.20.5 Visual Resources Audiovisual technology serves as an effective means to enhance comprehension of the testing program. It offers a vivid visual representation of the test's proceedings and can provide viewers with a practical demonstration of how the test is conducted. An illustrative example is provided by the OECD in the form of a collection of videos titled "Powerful Players and Successful Translators for Education" (OECD, 2011). These videos explore the educational systems of different countries.

2.20.6 Encyclopedia to explain educational programs as used in Trends in International Mathematics and Science Study (TIMSS) and PIRLS

2.20.7 National reports

These reports give countries the opportunity to focus on issues that affect their situation.

2.20.8 Technical reports

Following the release of international reports, it is customary to conduct further in-depth analyses to spotlight specific areas of policy interest. For instance, there might be a significant interest in examining disparities in academic performance between male and female students. Indeed, there is a notable report titled "Gender Differences in Achievement" (Mullis et al., 2000) based on TIMSS data, and another report named "Equally Prepared for Life: How 15-Year-Old Boys and Girls Perform in School" (OECD, 2009) that relies on PISA data. Additionally, a thematic report that focuses on the relationship between academic achievement, mathematics, and science at the primary school level (Grade 4) was produced using data from both TIMSS and PIRLS (Martin & Mullis, 2013).

Additionally, various countries generate their own reports, addressing unique educational concerns. For instance, in Australia, a prominent issue revolves around the educational progress of Indigenous students, who consistently perform better than non-Indigenous peers. To gain a more comprehensive understanding of the achievement rates of Indigenous students, Australia conducts targeted analyses, yielding reports like "The Achievement of Australia's Indigenous Student 2000–2006" (Bortoli & Thomson, 2009).

2.20.9 Summary reports

Although international and national reports can provide many details about test results, they are also necessary for the production of short papers and focused reports. Information presented in this manner is typically readily comprehensible to the reader. This style of reporting is user-friendly and can concentrate on just one or two specific issues.

2.20.10 Access to databases

The IEA Data Center has developed software applications, namely the IEA International Database Analyzer and IEA Data Visualizer, to simplify the analysis and presentation of data obtained from IEA studies within a comprehensive international database. You can access and download these applications directly from the IEA website.

2.20.11 Analysis services and support

Many LSAs primarily offer a database of information gathered during student assessments and from various sources, including principals, teachers, and parents. Some of this resulting information is accessible for free on social networking platforms, enabling individuals with an interest to conduct their own research. The generated data is huge. For instance, the PISA 2012 website offers a comprehensive dataset that includes a whopping 350 variables for each student. These variables include not just the mean performance scores of students but also their responses to all questionnaire items, scores within structured indices, and information related to sample weights. Likewise, the TIMSS database offers an additional resource called the International Database (IDB) Analyzer, which is available for free download. This tool aids researchers in using SPSS to accurately analyze data by applying the appropriate weighting and aggregating information collected from students, their educators, and educational institutions.

2.20.12 Policy briefs for ministry

Policy matters can find resolution through concise reports. SACMEQ Policy Problems Series is the best example for this type of policy briefs which contains topics such as what are the levels and styles of learning and mathematical success? (Makuwa, 2010). Unwritten papers focus on a specific target group - for example, in a teacher's research paper, they can explore particular aspects of the framework and delve into the challenges that learners encounter most frequently.

2.20.13 Interactive Web Sites

Simplifying the exchange of information regarding test results can be achieved quite effortlessly though the establishment of a dedicated website for this specific purpose. Users would only need to input the necessary details, typically from a menu, and the system would generate the response. These websites can prove highly valuable as they will be established for the purpose of analyzing data through methods grounded in test construction. PISA, for example, generates common errors using a standard frequency (BRR) An intricate analytical approach that doesn't require user intervention. Website users need not be concerned, as they will obtain results consistent with those presented in the international report. The OECD maintains a website (OECD, 2012) that enables users to choose various preferences for a specified set of countries or regions participating in PISA. Subsequently, an outcome image is generated, enabling users to easily make straightforward comparisons between different countries. It also provides the capability to assess the robustness of the relationship between variability and student performance. This proves highly beneficial for individuals tasked with producing a report on test results or examining specific issues. For instance, an education authority might wish to compare gender disparities in their students' outcomes with those from similar or geographically proximate nations. Nonetheless, complex multivariate analyses and in-depth investigations still fall under the purview of researchers who delve into publicly accessible datasets for such purposes, facilitated by analytical agencies. Given the complexity of data sets and the intricate nature of using the information within them, user guidelines and analysis instructions are made available to cater to the needs of those accessing the public data set.

2.20.14 Capacity building workshops

In certain countries, there might be a lack of access to the necessary technology required to conduct the analysis needed for generating report items. In this regard, countries can find organizations that can donate to conduct workshops that focus on analysis and report writing. The purpose of these workshops is to boost the skills of the country's workforce to a level where they can effectively conduct on-site analyses.

This primarily centers on the SACMEQ project, which orchestrates workshops in various countries. These interviews are strategically crafted to empower participants from each country with the necessary expertise to effectively apply, evaluate, and document the outcomes of significant academic assessments.

2.20.15 Manual

Aside from the workshops, a data analysis manual is also developed. This manual will provide analysts with a comprehensive foundation in the precise methods required for them to independently perform their own analyses. Example of this type of publication can be TIMSS 2007 User Guide for the International Database (Foy & Olson, 2009). The OECD has additionally released the PISA Data Analysis Manual (OECD, 2009) with the aim of aiding researchers in crafting their own analyses of the process. This resource enables them to duplicate the techniques employed in generating the official report's analysis.

2.21 At National Level Reporting and Disseminating

2.21.1 Reporting and Dissemination Model Used By NAEP

Table 2.9 Reporting and Dissemination of NAEP

Audience: At Federal level, State level, Local district level, local school level, general public, Member of press, Education research personnel					
Strategy					
What to Report?	Interest	Utility	Understandable	Validly interpret	Preference
How to report?	Interest	Utility	Understandable	Validly interpret	Preference
Dissemination vehicle	Accessibility	Regular Use	Preference	Feasibility	…

The National Assessment of Educational Progress (NAEP) is an exemplary survey that provides the American public with accurate information about academic achievement. NAEP is administe red by administrators from the National Center for Education Statistics (NCES), under the directi on of the bipartisan, independent National Assessment Governing Council. The Governing Coun cil is responsible for improving the teaching process, improving the quality and effectiveness of NAEP results, and overseeing the initial publication of NAEP data. Its results are frequently presented on an easily navigable website, showcasing key insights, links, and outcomes. This website format is employed to display outcomes for various countries, states, and territories.

The Board plays a pivotal role in contributing insights about the structure, accessibility, and content of these publications. The NCES Director, in consultation with the Commission, holds responsibility for overseeing the assessment process, ensuring data accuracy, compiling information, and releasing NAEP results. The Board is tasked with enhancing the quality and dissemination of NAEP outcomes, coordinating scheduling, and managing the initial public disclosures.

Guidelines are adhered to when crafting a NAEP report. These reports are publicly accessible, communicated in a manner suitable for diverse stakeholders, and conform to the NCES data quality criteria. They embody objectivity, precision, transparency, and impartiality. The report encompasses outcomes categorized by country, state, school district, and school type, while subgroup outcomes are provided where reliable. Group findings are shared for public review, though they don't modify the results. Results pertaining to students with disabilities and English language learners are presented separately, taking into account demographic shifts in the assessed student population.

The NAEP report presents a comprehensive representation of student achievements, presenting results based on achievement levels, grade averages, and percentile distributions. Varied data sets are included when dependable and available. The report's format may evolve with time, in accordance with technological advancements and input from the Commission and NCES. Charts encompass visual aids such as videos, diagrams, charts, and graphs to illuminate key discoveries.

Valid and reliable NAEP data is accessible via NAEP Data Tools at the initial release, except data from specialized samples and surveys. The committee reviews the draft and collaborates with NCES to refine the proposed process. Ongoing feedback from the committee assists in enhancing the initial release, to which NCES responds accordingly.

The Board is in charge of the strategic planning and execution of the primary unveiling, determining the timing, date, and strategy. Collaborating with NCES, the Council forges an embargo policy. The Board also orchestrates initial public relations endeavors, including press releases, conferences, and content surveys. A comprehensive information dissemination approach is pursued, with support from cultural, social, national, state, and local entities.

The debut edition may be accompanied by supplementary resources, technical documentation, and secondary analysis. For result publication, initial information is shared through diverse channels, encompassing news platforms and organizations. Schools and districts engaged in NAEP receive guidance on accessing outcomes. Highlights and geographical representations are disseminated across media and social platforms, along with key points crafted for provinces and urban regions.

2.21.1.1 Reporting and Dissemination of ANA Result

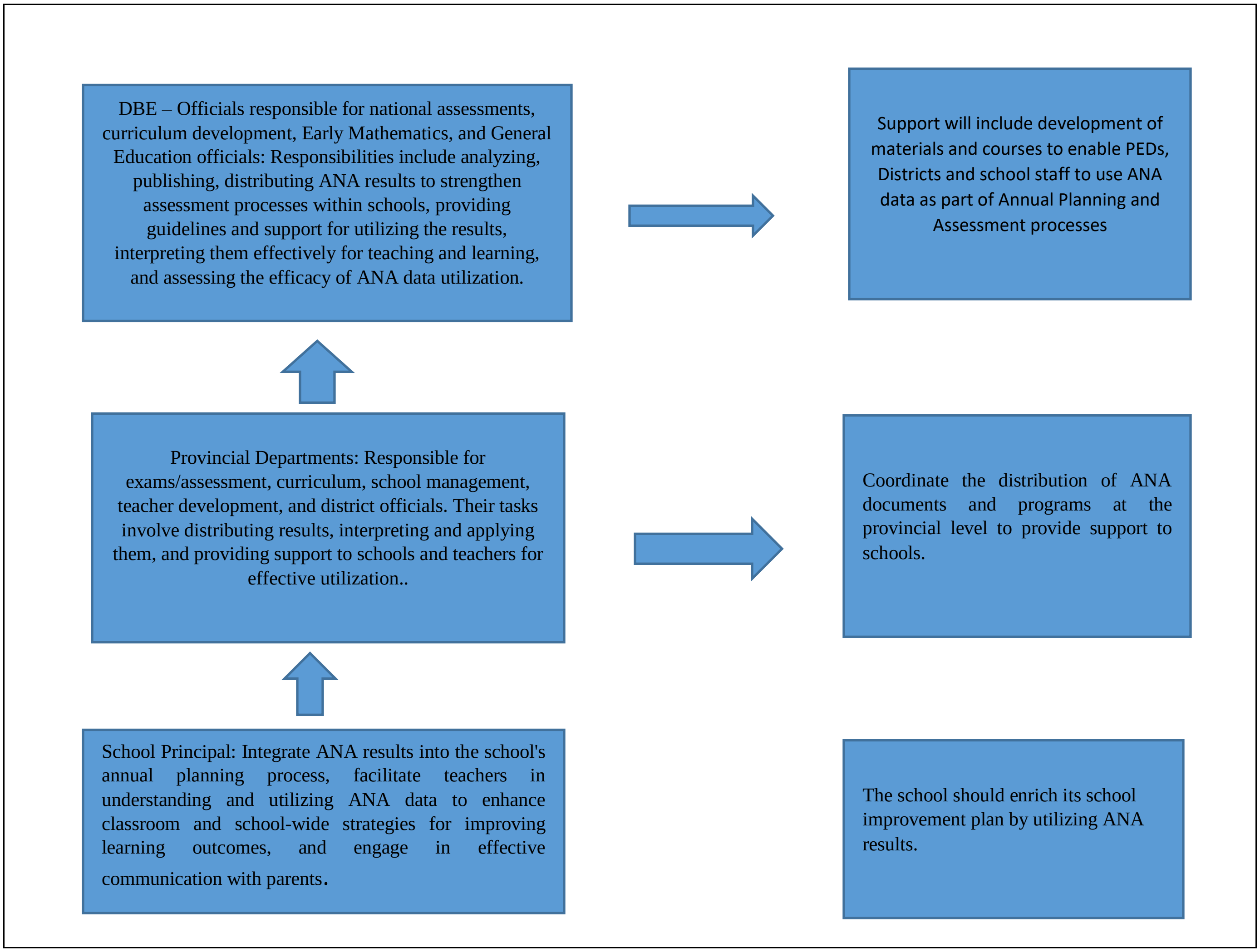


Figure 2.3 Reporting and Dissemination of ANA

The ANA findings provide the system with useful data to assess literacy and numeracy skills and serve as a gauge for tracking progress over time towards the accomplishment of predetermined goals. Educators and administrators are equipped with tools to evaluate the data and utilize it to create interventions to improve underperforming regions. The results clearly indicate any supplementary educational materials that may need to be created by teachers or subject matter experts to improve in-class teaching and assessment in areas where students exhibit notable weaknesses. Based on the findings, districts create targeted interventions for the schools that require them the most. Currently, 60% of students can demonstrate acceptable levels of literacy and numeracy according to national criteria. Schools set yearly goals with respect to this criterion. Learners receive information on the skill areas in which they excel and those in which they struggle. Teachers are directed towards specific teacher development courses using information from ANA, and principals are significantly engaged in discussions about the issues faced by the schools. Districts, parents, and schools identify which schools, students, and instructors most urgently require help through the ANA findings, which provide a consistent source of information (Makola, 2018). Schools are informed of issues such as programs that might not be operating as planned or the need to fill crucial teaching positions to ensure they receive the entire allocation of resources in accordance with the funding strategy. However, it is important to note that ANA results do not serve as a basis for allocating additional resources to a school than usual, as this might lead to a scenario where schools purposefully perform poorly to draw in more funding. Instead, the existing strategy of giving disadvantaged school communities greater resources to make up for home background disadvantages continues to be implemented (Heever, 2016). By the end of the academic period in which the DBE announces the ANA results, the school principal ensures that parents receive them. It is advised that the learners' quarterly report cards include the ANA findings. Parents are informed about and given a discussion about the report card with the ANA results. In alignment with the principles outlined in Section 16A of the South African Schools Act, it is mandatory to present a summary of the results along with a school development plan during the governing body meeting following the release of the ANA results. This demonstrates how the school intends to use its resources wisely to address the report's criticisms of its performance. The School Management Team (SMT) and School Governing Body (SGB) lead this nationwide awareness campaign to ensure that all parents are well-informed about the Assessment of National Achievement (ANA), its importance, and the overarching strategy to improve student academic achievement. Reports on learner performance that highlight conceptual and skill development challenges and the type of assistance that students require are given to parents (Christie & Potterton, 2011).

By law, school administrators inform the department head about a student's academic achievements. This obligation involves the creation and submission of an Annual Academic Report to the Head of the Department for assessment. This responsibility is stipulated in Section 16A of the South African Schools Act of 1996, as amended, which mandates that school principals compile and deliver an annual academic report. The report includes the efficient use of available resources in addition to the academic performance of that school according to the minimum outcomes, standards, and assessment processes established by the Minister. The District office receives the annual report and end-of-year schedules, which are examined to identify underperforming schools. The designated low-performing school is subsequently informed in writing of the situation. The designated schools formulate and present a plan that clearly outlines how resources will be distributed and offers a detailed framework of strategies designed to enhance academic performance within the school. The Head of the Department approves the Academic Improvement Plan. The District office receives the annual report and end-of-year schedules, which are examined to identify underperforming schools. The designated low-performing school is subsequently informed in writing of the situation. The strategies for improving the academic performance of the school include closely monitoring the ANA results (APIP). Provinces, districts, and educational institutions that set targets utilize their present performance as a baseline before establishing improvement goals. Targets must be both relevant and reasonable, including existing resources and capabilities to be realistic. They should not be too high or too low for them to imply anything. External IQMS moderators conduct regular visits to feeder schools, including those that are under-performing and located in districts with specific challenges. During these visits, their primary aim is to assess not only the implementation of IQMS but also the various factors that influence the quality of teaching and learning. Additionally, they visit District offices to see how well they are performing in terms of supporting their schools and how effectively they are doing so. Primarily, this intervention assists in acquainting both the schools they visit and the district offices with the understanding and interpretation of the content contained within this guideline document. Additionally, moderators are responsible for keeping an eye on planned interventions and the timely completion and submission of reports. This makes it possible to find troubled schools and help them. The monitoring reports help the DBE determine how these ANA findings are being used as well as how much corrective action is being done by the district and schools (Department of Basic Education, 2012).

SUMMARY

This chapter highlights the previous literature regarding the assessment, the historical evolution of Large scale Assessments (LSAs), the diverse objectives they serve such as promoting self-regulated learning, the various perspectives on assessment, and an examination of assessment challenges within the context of Pakistan have been explored. Moreover, the researches regarding large scale assessment and particularly large-scale assessment in Pakistan e.g. NAT and its framework. The second part of this deals with the quality of the test item. In which teacher perception, item analysis, and CTT are briefly explained. The third part of the chapter addresses the global landscape of reporting and dissemination in large-scale assessment. It discusses various frameworks utilized at both international and national levels for this purpose.

CHAPTER 3

RESEARCH METHODOLOGY

This section offers a summary of the methodologies and procedures employed in the present research endeavor. The study adopted a mixed-methodology approach, and this chapter encompasses key aspects such as the study's target population, the methodology for selecting the study sample, the tools and instruments utilized, the pilot testing process, the validation of research instruments, and the outlined procedure for collecting data. This study was designed to analyze the development of National achievement test of two year 2016 and 2019 through item analysis and to identify the teacher's perception regarding NAT and its application. For item analysis, there are two approaches used classical test theory (CTT) and item response theory (IRT). The CTT was employed to identify the psychometric properties of the test items of the National Achievement Test 2016 and 2019 such as item difficulty, item discrimination. The second questionnaire to identify the teacher perception regarding NAT and its application was conducted through survey. The data was collected from the teachers which participated in National Achievement Test 2016 and 2019. And also the document analysis was conducted for identifying the gaps in reporting and dissemination of NAT. This chapter also discusses method used for data collection and data interpretation.

3.1 PHILOSOPHICAL PARADIGM

This study adopts the pragmatism paradigm, which is particularly suited to educational research that seeks to address complex, real-world issues through practical and context-sensitive inquiry. Pragmatism, as a philosophical orientation, is grounded in the belief that the value of knowledge lies in its practical consequences and its usefulness in informing action. It rejects the rigidity of purely positivist or interpretivism positions and instead supports the use of multiple methods, perspectives, and data sources to understand and solve problems.

In the context of this study, which investigates the development, application, and impact of the National Achievement Test (NAT) in Pakistan, pragmatism provides an appropriate lens because it aligns with the study's dual emphasis on empirical analysis and contextual interpretation. The research aims not only to examine the technical aspects of the NAT such as its structure, validity, and alignment with curriculum standards but also to explore stakeholders' perceptions and the broader educational implications of its implementation.

A key tenet of pragmatism is its focus on what works in a given context. Therefore, this study is problem-centered and outcome-oriented, seeking to generate practical recommendations that can enhance the effectiveness of the NAT as a tool for improving teaching, learning, and policymaking. This includes identifying gaps in test design, challenges in implementation, and mismatches between assessment outcomes and educational objectives.

Furthermore, the pragmatist paradigm justifies the use of a mixed-methods approach, which combines quantitative data (e.g., performance trends, statistical indicators) with qualitative insights (e.g., interviews with educators, curriculum experts, and policymakers). This methodological pluralism enables a more holistic understanding of how the NAT functions within Pakistan's educational landscape and how it can be improved.

Ultimately, the adoption of a pragmatist paradigm allows this research to maintain flexibility, relevance, and applicability, ensuring that its findings contribute meaningfully to educational assessment practices, curriculum development, and evidence-based policy decisions in Pakistan.

3.2 RESEARCH DESIGN

Utilizing an observational approach, this study retrospectively examines pre-existing data from the National Achievement Test (NAT) administered in 2016 and 2019 across public and private schools in Pakistan. Specifically focusing on Grades 4 and 8 Mathematics and Science, the study aims to evaluate the consistency and effectiveness of the NAT assessment instruments in measuring intended criteria. Additionally, the research incorporates a correlational analysis to understand the relationship between NAT implementation and teacher perceptions. The survey includes a combination of open-ended and closed-ended questions to gather both quantitative and qualitative insights. Furthermore, a checklist through document analysis was utilized to assess the reporting of NAT results, enhancing the rigor and comprehensiveness of the analysis.

3.3 RESEARCH METHODOLOGY

This study employs an ex-post facto method to critically analyze the development and application of the National Achievement Test (NAT) based on existing data from NAT administrations in 2016 and 2019, without researcher intervention. Utilizing Classical Test Theory (CTT), quantitative analyses focus on assessing key psychometric properties of the NAT, including Difficulty Index (DIF) and Discrimination Index (DI), providing insights into test reliability and validity across diverse contexts. Concurrently, qualitative investigation delves into the causal repercussions and potential impacts of NAT administration on stakeholders such as students, teachers, and policymakers, exploring how NAT results influence educational policies, curriculum development, and instructional practices. This qualitative aspect involves a survey questions administered first to collect quantitative data on specific aspects of teacher perceptions, providing structured insights into their attitudes and behaviors related to NAT implementation. Subsequently, open-ended survey questions are utilized to capture nuanced perspectives from teachers regarding NAT application, offering insights into their perceptions and utilization of NAT results in instructional strategies. Complementing these analyses, document analysis examines the reporting and dissemination of NAT results, shedding light on how assessment outcomes are communicated and utilized within the educational system with the support of a checklist.

3.4 POPULATION

The Population for the study was the subject teacher and students of Math’s & Science of Grade 4th and Grade 8th which was the part of previous NAT 2016 and 2019 in Islamabad and AJK.

Table 3.1:

The Total Numbers of Teacher Participated in NAT 2016

Grade	Area	Subject	Location	Gender	Teacher
Grade 4	AJK		Rural:37	Female:30	44
		Mathematics	Urban:07	Male:14	
	ICT		Rural:09	Female:20	28
			Urban:19	Male:8	
Grade 8	AJK	Science	Rural:40	Female:20	48
			Urban:08	Male:28	
	ICT		Rural:11		34
			Urban:23	Female:16	
				Male:18	

Note: From National Achievement Report, 2016.

The Table 3.1 identify the population of NAT 2016 for this research includes teachers in Grade 4 and Grade 8 from two areas: AJK (Azad Jammu and Kashmir) and Islamabad, which represents the Islamabad Capital Territory (ICT). The subjects of focus are Mathematics and Science for Grade 4& 8. The teachers are further categorized based on their location, distinguishing between rural and urban areas within AJK and Islamabad.

For Grade 4 Mathematics, there are 37 teachers from rural areas in AJK and 7 teachers from urban areas. Among them, there are 18 female teachers and 24 male teachers, making a total of 44 teachers. In Grade 4 ICT, there are 9 teachers from rural areas and 19 teachers from urban areas.

The gender distribution for Grade 4 ICT includes 20 female teachers and 8 male teachers, making a total of 28 teachers.

Moving to Grade 8, for Science, there are 40 teachers from rural areas in AJK and 8 teachers from urban areas. Among them, there are 28 female teachers and 20 male teachers, making a total of 48 teachers. For Grade 8 ICT, there are 11 teachers from rural areas and 23 teachers from urban areas. The gender distribution includes 17 female teachers and 17 male teachers, making a total of 34 teachers.

Table 3.2: ***The Total Numbers of Teacher Participated in NAT 2019***

Grade	Area	Subject	Location	Gender	Teacher
Grade 4	AJK		Rural:10	Female:07	14
		Science	Urban:04	Male:07	
	ICT		Rural:02	Female:08	16
			Urban:06	Male:08	
Grade 8	AJK	Mathematics	Rural:09	Female:07	14
			Urban:05	Male:07	
	ICT		Rural:02	Female:04	08
			Urban:06	Male:04	

Note: From National Achievement Report, 2019.

The table 3.2 identify the population of the study consisted of teachers from Grade 4 and Grade 8 in two areas: AJK (Azad Jammu and Kashmir) and ICT (Islamabad Capital Territory). The subjects of focus were Science for Grade 4 and Mathematics for Grade 8. The teachers were further categorized based on their location, distinguishing between rural and urban areas within AJK and ICT.

For Grade 4 Science, there were 10 teachers from rural areas in AJK and 4 teachers from urban areas. Among them, there were 7 female teachers and 7 male teachers, making a total of 14 teachers. In Grade 4 ICT, there were 2 teachers from rural areas and 6 teachers from urban areas. The gender distribution for Grade 4 ICT included 8 female teachers and 8 male teachers, making a total of 16 teachers.

Moving to Grade 8, for Mathematics, there were 9 teachers from rural areas in AJK and 5 teachers from urban areas. Among them, there were 7 female teachers and 7 male teachers, making a total of 14 teachers. For Grade 8 ICT, there were 2 teachers from rural areas and 6 teachers from urban areas. The gender distribution included 4 female teachers and 4 male teachers, making a total of 8 teachers. These teachers were the participants who took part in NAT 2019.

Table 3.3
Numbers of Students Participated In NAT 2016

Grade	Area	Subject	No of Students	Total School
Grade 4	AJK		880	59
	ICT	Mathematics	560	28
Grade 8	AJK	Science	960	48
	ICT		680	34

Source: National Achievement Report, 2016.

The Table 3.3 provides information about the student data from two areas, AJK (Azad Jammu and Kashmir) and ICT (Islamabad Capital Territory), who participated in the National Achievement Test (NAT) from Grade 4 and Grade 8. The subjects of focus were Mathematics and Science for Grade 4 & 8.

For Grade 4 Mathematics in AJK, there were a total of 880 students from 59 schools who participated in the NAT. In Grade 4 ICT, there were 560 students from 28 schools. These students from AJK represent the population that took part in the NAT for these subjects.

Moving to Grade 8, for Science in AJK, there were 960 students from 48 schools who participated in the NAT. In Grade 8 ICT, there were 680 students from 34 schools. These students from AJK represent the population that took part in the NAT for these subjects.

Table 3.4

Numbers of Students Participated In NAT 2019

Grade	Area	Subject	Location	No of Students	Total School
Grade 4	AJK	Mathematics	Rural:10 Urban:04	78	14
	ICT		Rural:02 Urban:06	516	08
Grade 8	AJK	Science	Rural:10 Urban:04	113	14
	ICT		Rural:02 Urban:06	242	08

Note: From National Achievement Report, 2019.

The table 3.4 presents the data of students who took part in the NAT 2019 from Grade 4 and Grade 8 in two areas: AJK (Azad Jammu and Kashmir) and ICT (Islamabad Capital Territory). The subjects assessed in the NAT included Mathematics and ICT for Grade 4, and Science and ICT for Grade 8. In Grade 4 Mathematics, there were 78 students from AJK, distributed across 14 schools. For Grade 4 ICT, there were 516 students from 8 schools in AJK.

Moving to Grade 8, in AJK, there were 113 students who participated in the NAT for the Science subject, distributed across 14 schools. In ICT, which represents Islamabad Capital Territory, there were 242 students from 8 schools who took part in the NAT 2019 for Grade 8.

3.5 SAMPLING

Regarding the pilot study's sample size is a subject of debate in the literature, with no clear consensus on the appropriate number of participants (Thabane et al., 2010). However, Hertzog (2008) underscored that when the primary aim of the pilot test is to assess the internal consistency of an instrument, a small sample size would be inadequate. In this study, the sampling design employed stratified sampling. Primary and secondary level schools in Islamabad and AJK were selected accordingly. As previously mentioned, the researcher conducted item analysis using SPSS, implying a quantitative analysis of the responses. Utilizing SPSS typically necessitates a large dataset to yield statistically meaningful results. Thus, the overall student population was included in this study to ensure a robust dataset for analysis and enhance the reliability of research findings.

Regarding teachers, a sample of 100 was selected out of 206. The selection of the sample was guided by research advisor (2006), which were based on the Krejcie & Morgan sampling method from 1970. Due to constraints such as budget and time, it was not feasible to include all teachers.

An important criterion in the selection of primary schools was their participation in the National Achievement Test. The sample assessment test was obtained from the coordinator of science and mathematics in the national assessment system for education. Demographic data of the representative sample was collected from the national director of NEAS.

Table 3.5

Sample size for the study

		2016		2019	
Grade	**Area**	**Students**	**Teachers**	Students	**Teachers**
Grade 4	AJK	78	20	78	11
	ICT	516	10	516	13
Grade 8	AJK	113	20	113	11
	ICT	242	10	242	05

The table 3.5 represents the sample size of students and teachers who participated in the National Achievement Test (NAT) in the years 2016 and 2019 for Grade 4 and Grade 8 in the AJK (Azad Jammu and Kashmir) and ICT (Islamabad Capital Territory) areas.

For Grade 4 in AJK, there were 78 students and 20 teachers who participated in the NAT in 2016. In 2019, the same number of students (78) took part in the NAT, but the number of teachers decreased to 11.In Grade 4 ICT, the sample size was larger with 516 students and 10 teachers in

2016. Similarly, in 2019, there were 516 students who participated in the NAT, but the number of teachers increased to 13.

Moving to Grade 8 in AJK, the sample size was 113 students and 20 teachers in 2016. In 2019, the number of students (113) remained the same, but the number of teachers decreased to 11.In Grade 8 ICT, there were 242 students and 10 teachers in 2016. In 2019, the number of students remained the same (242), but the number of teachers decreased to 5.

3.6 METHODS AND TOOLS OF DATA COLLECTION

The study utilized an ex-post facto method, enabling researchers to analyze pre-existing data of National achievement test and to assess whether the instrument employed maintained consistency and effectively measured its intended constructs. The purpose of this study was to analysis the development of National Achievement Test and the perceptions teachers hold regarding the application of NAT. Teacher's perceptions of NAT were identify through the adapted questionnaire. For analysis of development NAT, the item analysis technique was applied on the National achievement test of science and mathematics NAT 2016 and 2019.

3.6.1 For Item analysis

In this study, for analyses of development of NAT the students" performance scores were gotten for science and math cognitive achievement tests which were designed for NAT 2016 &2019. The student and teacher demographics information and the teacher's data were collected from the National Coordinator of National Assessment Wing (NAW). The reason why this achievement test is selected for this study is that it the only National test conducted in Pakistan. Therefore, the findings will be applicable at the national level. The achievement test which items was used in this study contained objective items. For item analysis, the correct response was coded as 1 and 0 as the wrong response.

Table 3.6

Description of the Achievement Test 2016 and 2019

Subject	Grade	Objective Items	Content
Mathematics	Grade 4	50	Test contains topics such as numbers, factors, multiples, common and decimal fractions, Measurement and information handling.
Mathematics	Grade 8	31	Test contains topics such as numbers, factors, multiples, common and decimal fractions, Measurement and information handling.
Science	Grade 4	38	Test contains topics related to Life Science, Physical Science and Earth and Space Science have been included in this test
Science	Grade 8	44	Test contains topics related to Life Science, Physical Science and Earth and

	Space Science have been included in this test

The table 3.6 highlighted the achievement test used in the study includes two subjects: Mathematics and Science, administered to students in Grade 4 and Grade 8. The Mathematics test for Grade 4 consists of 50 objective items that cover various topics such as numbers, factors, multiples, common and decimal fractions, measurement, and information handling. Similarly, the Mathematics test for Grade 8 includes 31 objective items covering similar topics.

For the Science subject, the Grade 4 test comprises 38 objective items that assess students' knowledge in areas of Life Science, Physical Science, and Earth and Space Science. Likewise, the Grade 8 Science test consists of 44 objective items covering the same science domains.

3.6.2 For Teacher Perception

The second questionnaire to evaluate the teacher perception regarding NAT and its application contained two major sections: demographics (6 items), and NAT perception related item (24 items) which further subdivided into the following accountability (9 items), monitoring (5) instructional methodology and gate keeping (6+4) The questionnaire is distributed among a subject teacher of Maths, of grade 4th and grade 8th science which were the part of previous NAT to identify their perception about the national achievement test and its application relationship.

Teachers Questionnaire Item Distribution

Dimension	**Items**
Accountability	9

Monitoring	5
Instructional	
Methodology	6
Gatekeeping	4
Total number of items	24
Rating scale	Ordinal frequency Five-point

The second section comprised 24 items aimed at gauging teachers' perceptions of NAT. It utilized a five-point rating scale, with "agree" on the left and "disagree" on the right.

The study centered on participants' outlooks regarding NAT (National Achievement Test). These outlooks are crafted within an individual's cultural milieu and serve as a means for individuals to comprehend events, like NAT, within their surroundings. The significance of these perceptions lies in their role in shaping behaviors, as behaviors often stem from how individuals perceive reality or what they regard as normal or commendable. Numerous factors, including personal interests, life experiences, and proximity, can influence these perceptions. Consequently, when delving into perceptions about NAT, it's plausible that their proximity to the education system plays a substantial role in shaping these perceptions.

The third section comprised of 10 open-ended questions. This section was intended to examine teacher's perception regarding NAT application. Researcher utilized thematic analysis, a commonly employed qualitative interpretation method (Guest et al., 2012), to analyze the data. The study followed the thematic analysis principles as defined by Braun and Clarke (2006). This approach involved a repetitive process, including (a) assessing the data, (b) generating codes (i.e., identifying significant data segments), and (c) identifying and reviewing overarching themes. (i.e., categorizing codes into explicit categories). This approach allowed us to report the emerged concepts (codes) and patterns (themes) within the data (Braun & Clarke, 2006).

Throughout the analysis process, there was a constant iterative loop between the raw data collected and the developed codes and themes, enabling the identification of patterns directly from participant quotes, avoiding vague interpretations (Boyatzis, 1998). Each code represented a distinct meaning, although they may not be entirely mutually exclusive, as the smallest unit of interpretation (codes) could converge into associated themes. To maintain consistency, meticulous attention was given to the details in participant quotes. The quotes included in the results section represent a carefully chosen set that aims to capture the core essence of the identified codes or themes, rather than attempting to cover all possible examples exhaustively.

3.7 For Reporting and Dissemination

Furthermore, to analyze the reporting and dissemination aspects of the NAT for the years 2016 and 2019, various documents related to NAT administration and reporting were examined after filling the checklist by 20 test administrator in NAT **(Annex-III).** The objective of this document analysis is to gain insights into the current framework for reporting and dissemination of NAT results. For document analysis, study considered four criteria presented by Flick (2018).

3.8 Pilot Testing

Pilot testing involves the preliminary assessment of research instruments on a smaller scale to prepare for the main study (Polit et al., 2001). Its primary purpose is to pre-test and evaluate the validity of research instruments. Typically, a sample size ranging from 10% to 20% of the intended main study sample size is considered suitable for conducting a pilot study. It's essential to note that a pilot study doesn't guarantee the success of the research instruments, but it substantially increases the likelihood of their effectiveness (Baker, 1994). Pilot testing is instrumental in identifying and rectifying inappropriate or misleading survey questions, ensuring the proper utilization of research instruments, and maintaining data consistency. Importantly, data from the pilot study should not be incorporated into the final research outcomes (Lancaster et al., 2004). The overarching goal of a pilot study is to furnish valid and reliable information, which profoundly influences the overall success or failure of the primary research project. Additionally, pilot studies save valuable time, effort, and resources, preventing potential losses that may arise from the unforeseen failure of a major research study (De Vaus, 1993).

3.9 Pilot Study of research instrument

In the current research study, an adapted research instrument was employed, which was used in a study exploring LSAs in Canada (Klinger, Rogers, Miller & DeLuca, 2008). However, given its application in the Pakistani educational context, this research instrument required validation. The purpose of the pilot testing phase was to evaluate the research instrument's validity and reliability. Pilot testing serves as a critical step in identifying and addressing various potential issues related to the research instrument. For this research study, a pilot study was a sample of 30 experts was selected for participation, following the criterion of 10 experts per checklist item, as recommended by Nunnally and Bernstein (1994).

3.10 Validity of Research Instruments for teacher: Validity refers to the extent to which research instruments accurately measure the aspects the researcher intends to assess through the questionnaire. It is crucial to establish both reliability and validity before administering research instruments. In this study, content validity for all research instruments was assessed through face validity **(Annex-V)** and factor analysis, which involved examining factor loadings, communalities, eigenvalues, and variance percentages, as shown in the table below.

Table 3.7

Factor Loading, Communalities and Variance for Gate Keeping Related Item

Factor	Item Code	Loading	Communalities	Eigen Values	Variance %
Gate keeping	GQ21	.887	.786	2.868	71.708
	GQ22	.947	.896		
	GQ23	.937	.877		
	GQ24	.556	.309		

The table 3.7 provided represents the results of a factor analysis conducted on a questionnaire related to the concept of gatekeeping. The factor analysis aims to identify underlying factors or dimensions that explain the relationships among a set of variables, represented by item codes GQ21 to GQ24.

The loading column indicates the strength of the relationship between each item and the underlying factor. Higher loading values indicate a stronger association. In this case, all the items show relatively high loading values, indicating a strong relationship with the underlying factor of gatekeeping. The communalities column represents the amount of variance in each item that can be explained by the underlying factors. Higher communalities suggest a better fit with the factors. In this analysis, the items have relatively high communalities, indicating a good fit with the gatekeeping factor.

The eigenvalues column shows the amount of variance explained by each factor. It indicates the importance of each factor in explaining the relationships among the variables. The higher the eigenvalue, the more influential the factor. In this case, the first factor (gatekeeping) has an eigenvalue of 2.868, suggesting it explains a substantial amount of variance.

The variance percentage column indicates the proportion of total variance explained by each factor. In this analysis, the gatekeeping factor explains 71.708% of the total variance.

Additionally, the Knowledge Management Organization (KMO) value of 0.702 is mentioned. The KMO is a statistical measure used to assess the overall suitability of data for factor analysis. A value above 0.7 generally indicates that the data is suitable for factor analysis.

The factor analysis of the questionnaire items related to gatekeeping indicates a strong relationship with the underlying factor. The loading values, communalities, and eigenvalues support the presence and importance of the gatekeeping factor. The KMO value of 0.702 suggests that the data is appropriate for factor analysis.

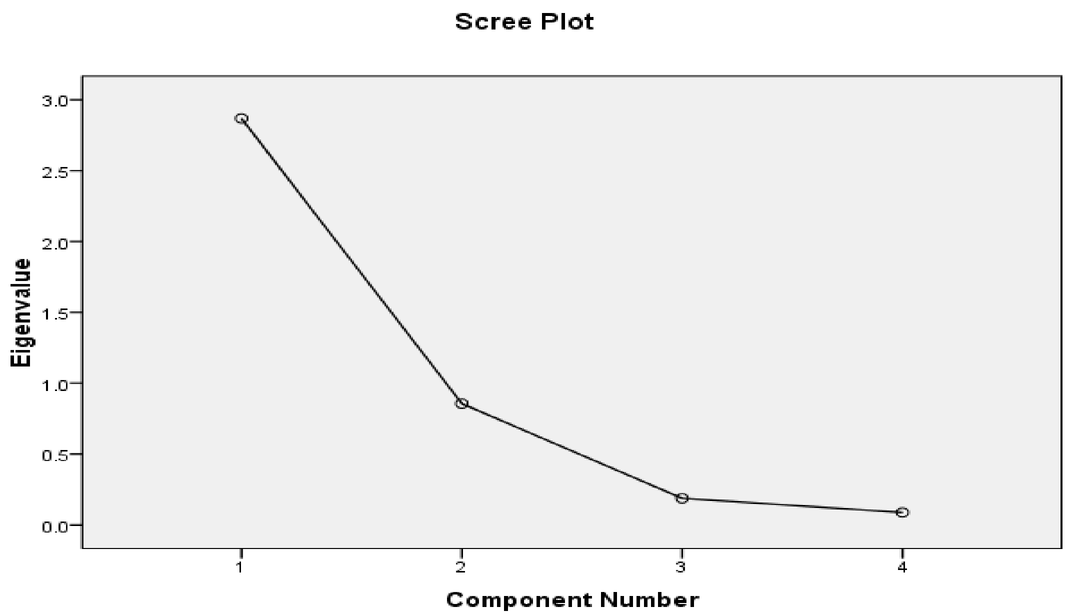


Figure 3.1 Scree Plot of gate keeping related items

Table 3.8 Factor Loading, Communalities and Variance for Accountability Related Items

Factor	Item Code	Loading	Communalities	Eigen Values	Variance %
Accountability	AQ1	.747	.619		
	AQ2	.558	.859		
	AQ3	.710	.721		
	AQ4	.869	.771		
	AQ5	.528	.240	2.882	54.660
	AQ6	.825	.794		
	AQ7	.867	.878		
	AQ8	.900	.867		
	AQ9	.508	.254		

The table 3.8 provided represents the results of a factor analysis conducted on a questionnaire. The factor analysis aims to identify underlying factors or dimensions that explain the relationships among a set of variables. In this case, the variables are represented by item codes (AQ1 to AQ9) related to the concept of accountability. The loading column indicates the strength of the relationship between each item and the underlying factor. Higher loading values indicate a stronger association. The communalities column represents the amount of variance in each item that can be explained by the underlying factors. Higher communalities suggest a better fit with the factors.

The eigenvalues column shows the amount of variance explained by each factor. It indicates the importance of each factor in explaining the relationships among the variables. The higher the eigenvalue, the more influential the factor. The variance percentage column indicates the proportion of total variance explained by each factor. In this case, the first factor (accountability) explains 54.660% of the total variance.

Additionally, the Knowledge Management Organization (KMO) value of 0.724 is mentioned. The KMO is a statistical measure used to assess the overall suitability of data for factor analysis. A value above 0.7 generally indicates that the data is suitable for factor analysis. Overall, the factor analysis of the questionnaire items related to accountability shows that these items are strongly associated with the underlying factor, as indicated by the high loading values. The communalities and eigenvalues further support the presence and importance of the underlying factor. The KMO value suggests that the data is appropriate for factor analysis.

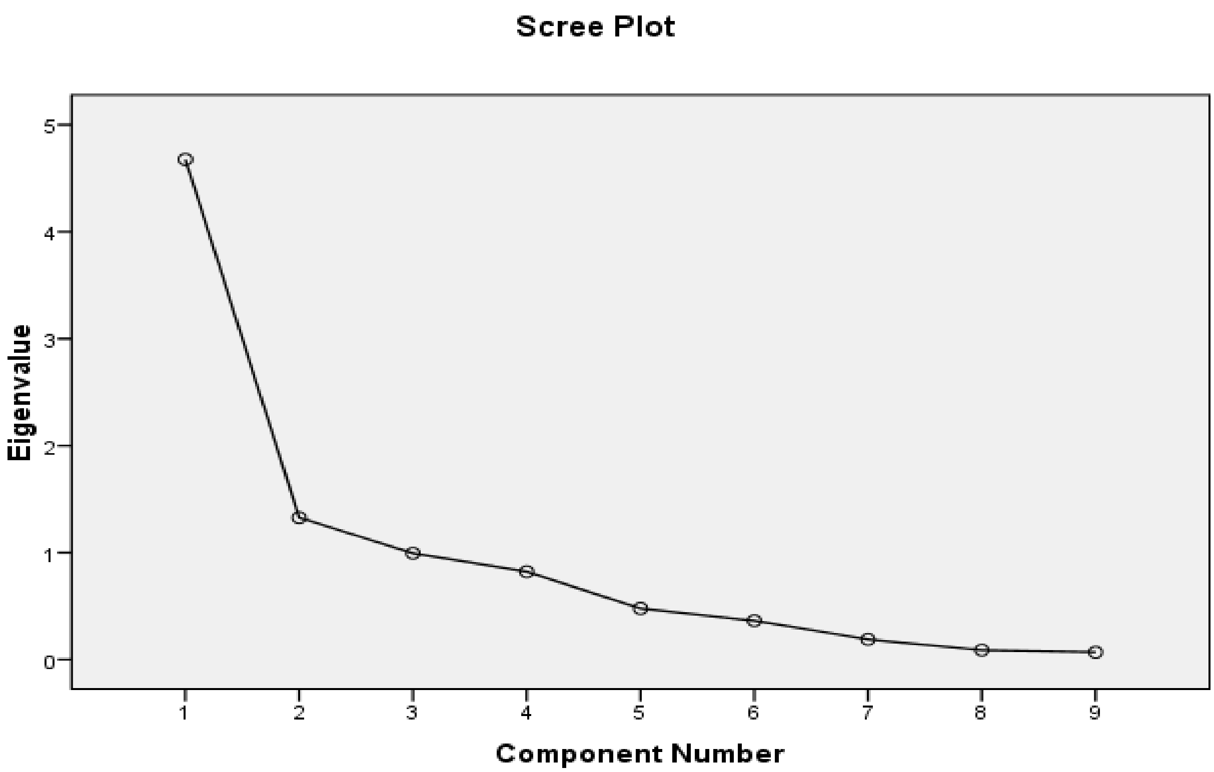


Figure 3.2 Scree Plot of accountability related item

Table 3.9

Factor loading, communalities and variance for Monitoring Related Items

Factor	Item Code	Loading	Communalities	Eigen Values	Variance %
Monitoring	MQ10	.762	.580		
	MQ11	.760	.577		
	MQ12	.844	.712	2.733	54.660
	MQ13	.838	.701		
	MQ14	.540	.462		

The provided table 3.9 represents the results of a factor analysis conducted on a questionnaire related to the concept of monitoring. The factor analysis aims to identify underlying factors or dimensions that explain the relationships among a set of variables, represented by item codes MQ10 to MQ14.

The loading column indicates the strength of the relationship between each item and the underlying factor. Higher loading values indicate a stronger association. In this analysis, items MQ10, MQ11, MQ12, and MQ13 have relatively high loading values, indicating a strong relationship with the underlying factor of monitoring.

The communalities column represents the amount of variance in each item that can be explained by the underlying factors. Higher communalities suggest a better fit with the factors. In this case, the items have relatively high communalities, indicating a good fit with the monitoring factor.

The eigenvalues column shows the amount of variance explained by each factor. It indicates the importance of each factor in explaining the relationships among the variables. The higher the eigenvalue, the more influential the factor. In this analysis, the first factor (monitoring) has an eigenvalue of 2.733, suggesting it explains a substantial amount of variance. The variance percentage column indicates the proportion of total variance explained by each factor. In this analysis, the monitoring factor explains 54.660% of the total variance. The KMO value provided for the factor analysis of the monitoring questionnaire items is 0.793. The KMO is a statistical measure used to assess the overall suitability of data for factor analysis. It ranges from 0 to 1, with higher values indicating better suitability for factor analysis.

In this case, the KMO value of 0.793 suggests that the data for the monitoring questionnaire items is considered to have a good level of suitability for factor analysis. This indicates that the variables included in the analysis are likely to have sufficient correlation and interrelatedness to extract meaningful factors.

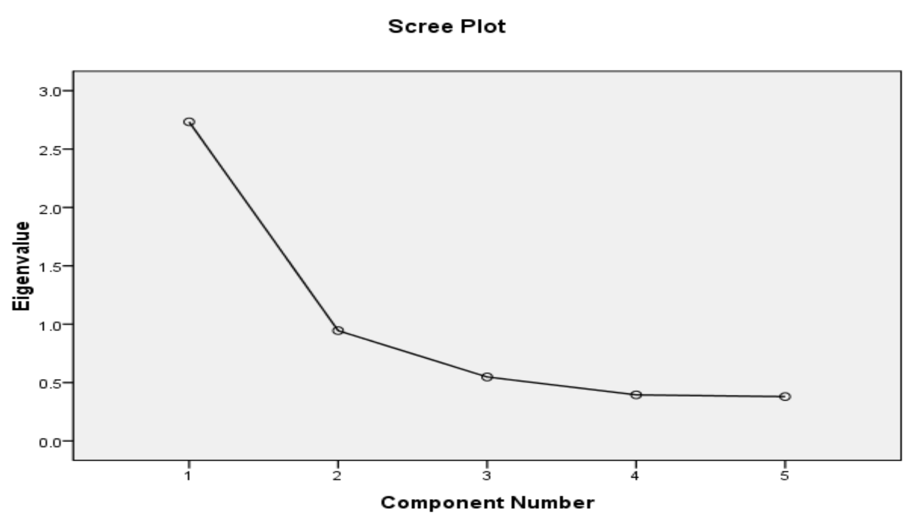


Figure 3.3 Scree Plot of monitoring related item

Table 3.10

Factor loading, Communalities and Variance for instructional methodology related items

Factor	Item Code	Loading	Communalities	Eigen Values	Variance %
Instructional methodology	IMQ15	.757	.614		
	IMQ16	.695	.692		
	IMQ17	.778	.699	3.160	52.672
	IMQ18	.877	.777		
	IMQ19	.821	.720		
	IMQ20	.877	.622		

The 3.10 table provided represents the results of a factor analysis conducted on a questionnaire related to the concept of instructional methodology. The factor analysis aims to identify underlying factors or dimensions that explain the relationships among a set of variables, represented by item codes IMQ15 to IMQ20

The loading column indicates the strength of the relationship between each item and the underlying factor. Higher loading values indicate a stronger association. In this analysis, items IMQ18, IMQ19, and IMQ20 have relatively high loading values, indicating a strong relationship with the underlying factor of instructional methodology.

The communalities column represents the amount of variance in each item that can be explained by the underlying factors. Higher communalities suggest a better fit with the factors. In this case, the items have relatively high communalities, indicating a good fit with the instructional methodology factor.

The eigenvalues column shows the amount of variance explained by each factor. It indicates the importance of each factor in explaining the relationships among the variables. The higher the eigenvalue, the more influential the factor. In this analysis, the first factor (instructional methodology) has an eigenvalue of 3.160, suggesting it explains a substantial amount of variance.

The variance percentage column indicates the proportion of total variance explained by each factor. In this analysis, the instructional methodology factor explains 52.672% of the total variance.

Additionally, the Kaiser-Meyer-Olkin (KMO) value of 0.724 is mentioned. The KMO is a statistical measure used to assess the overall suitability of data for factor analysis. A KMO value above 0.7 generally indicates that the data is suitable for factor analysis.

In summary, the factor analysis of the questionnaire items related to instructional methodology indicates a strong relationship with the underlying factor. The loading values, communalities, and eigenvalues support the presence and importance of the instructional methodology factor. The KMO value of 0.724 suggests that the data is appropriate for factor analysis.

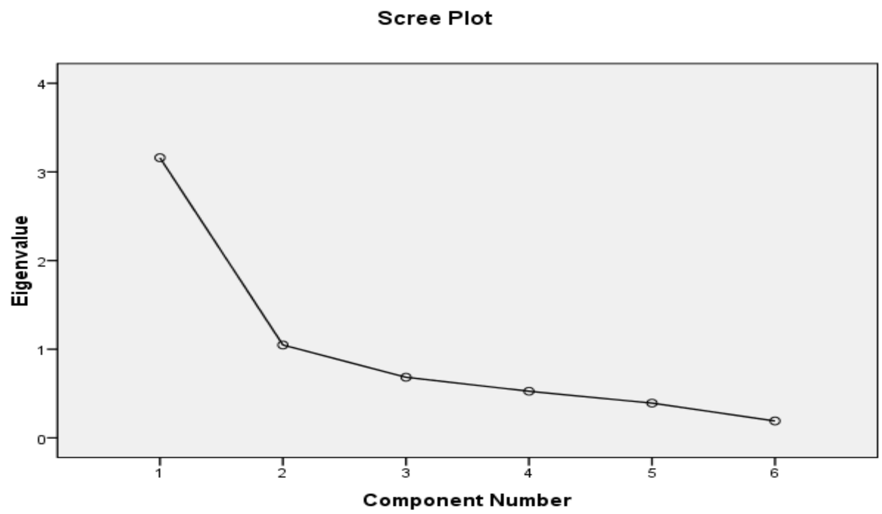


Figure 3.4 Scree plot of Instructional methodology related Items

3.11 Reliability of the Instruments

Reliability pertains to the steadiness and consistency of responses or behaviors within research studies (Creswell, 2007). In this study, the researcher assessed the internal consistency or reliability using measures such as Cronbach's alpha (Byrne, 2010). Cronbach's alpha is commonly used to test how well multiple indicators for a latent variable are related. According to Pedhazur and Schmelkin (1991), Cronbach's alpha is contingent on the inter-item correlations, and as the number of items increases, Cronbach's alpha generally tends to rise as well. Generally, a reliability coefficient (Cronbach's alpha) value above 0.7 is considered acceptable for diverse scales (Gliem et al., 2003). Certain researchers have proposed a threshold of Cronbach's alpha >0.6 as a benchmark for internal consistency (Chew, 2017).The findings of previous studies have shown satisfactory results using these cut-off values. Therefore, in the present study, these cut-off values were used, and if necessary, some items were eliminated to improve the reliability coefficient (Chew, 2017).

Table 3.11

Reliability of the Teacher related instrument

Reliability of the first set of questions

Board Question	Reliability
Does the National Achievement Test	**0.844**
Improve high academic standards?	
Improve students' achievement?	
Improve teachers' effectiveness?	
Perform as an accountability tool?	
Being a teacher, do you think that you/your	
Know how to interpret assessment results?	
Know how to use the test results?	
Principals know how to interpret the test results?	
Public know how to interpret the test results?	

Reliability of the second set of questions:

Board Question	Reliability
Does NAT monitor student progress by	**0.897**
Measuring student achievement?	
Aligning with classroom assessment instruments?	
Determining how well students are learning?	
Providing a "snap-shot" of what students can do?	
Helping students concentrate on learning effort?	

Reliability of the third set of questions:

Board Question	**Reliability**
Does the national achievement test impact...	**0.882**
Increasing teachers' assessment, knowledge, skills?	
Improving and enhancing teaching?	
Providing common measures for teachers?	
Evaluating the quality of students?	
Evaluating the quality of teachers?	
Evaluating the quality of schools?	

Reliability of the fourth set of questions:

Board Question	**Reliability**
Being a teacher, do you think that NAT	**0.834**
Opt the national assessment framework for assessment?	
Provide data for data-based decision-making at student level?	
Provide data for data-based decision-making at class level?	
Provide data for data-based decision-making at school level?	

For the open-ended questions in the questionnaire, a qualitative analysis approach was employed to analyze and interpret the responses. The responses were carefully read and reviewed to identify recurring themes, patterns, and key insights related to the National Achievement Test (NAT).

Inter-rater reliability was evaluated using Cohen's kappa statistic, a measure of agreement level between coders beyond chance. A kappa value of 0.85 was achieved, indicating substantial agreement between the coders. Any disagreements or discrepancies in coding were resolved through discussion and consensus between the coders.

The final coding scheme was then applied to the remaining responses, and the coders independently coded those responses accordingly. Regular meetings were held throughout the coding process to discuss any emerging themes, clarify coding guidelines, and address any questions or concerns. This rigorous approach to the analysis of open-ended responses helped ensure the reliability and consistency of the qualitative findings. The themes and insights derived from the analysis of the open-ended questions provide valuable qualitative data that complement the quantitative findings from the closed-ended questions.

3.12 Reliability and Validity Assessment of Checklist for Reporting and Dissemination of NAT

The checklist used in this study was developed to assess various aspects of reporting and dissemination practices related to NAT. Experts in educational assessment in Pakistan were involved in reviewing the checklist items to confirm their relevance and coverage of key aspects. A sample of 20 experts was selected for participation, following the criterion of 10 experts per checklist item, as recommended by Nunnally and Bernstein (1994).

To assess internal consistency reliability, Cronbach's alpha coefficient was calculated for each section of the checklist. Additionally, construct validity was examined through factor analysis to explore the underlying constructs represented by the checklist items. It was hypothesized that items within each checklist section would load onto distinct factors, providing evidence of construct validity.

Table 3.*12 Reliability and validity of checklist used for analysis of reporting and dissemination of NAT*

Aspect	Cronbach,s Alpha	Factor analysis
Cronbach,s Alpha Calculation	Expert Responses Agreed =80 percent Disagreed=20 percent	
Total Responses Agreed Disagree		
Calculated value/**overall mean**	Total number=0.7273	
Factor analysis		
Section 1		**Section 1**
For 1(a)		0.70
For 1 (b)		0.75
For 1(c)		0.72
Section 2		**Section 2**
For 2(a)		
For 2(b)		0.68
For 2(c)		0.73
		0.78
Overall Mean		0.725

)*/(

The table 3.12 revealed a Cronbach's alpha coefficient of 0.725, indicating a moderate to high level of internal consistency reliability for the checklist. This suggests that the checklist items are consistent in measuring the intended construct of reporting and dissemination practices related to NAT. Factor analysis further supported the construct validity of the checklist, as items within each checklist section loaded onto distinct factors. These findings provide evidence of the checklist's ability to effectively measure different aspects of reporting and dissemination practices.

3.13 Data Collection Procedure

The research aimed to analyze the development of the National Achievement Test (NAT) and identify teacher perceptions based on Nagy's (2000), framework for Large-Scale Assessments (LSAs), which includes accountability, gatekeeping, instructional improvement, and monitoring. Item analysis was conducted on the 2016 and 2019 Science and Mathematics Achievement Tests to determine their psychometric properties, focusing on objective items. Teacher perceptions were assessed using a standardized tool with three sections: demographics, NAT perception items divided into accountability (9 items), monitoring (5 items), instructional methodology (6 items), and gatekeeping (4 items), and 10 open-ended questions. The questionnaire was distributed to Grade 4 and Grade 8 Math and Science teachers who participated in NAT 2016 and 2019. Additionally, a comprehensive checklist (ACER-GEM & UIS, 2017) and document analysis was used to evaluate the reporting and dissemination process of LSAs, addressing policy, practice issues, and stakeholder information needs.

3.14 Ethical Consideration

As per Bryman (2012) emphasizes, obtaining ethics approval is a crucial step in research to ensure participant well-being, secure consent, and safeguard respondents against deception and breaches of privacy. In conducting this research on teacher perceptions regarding the National Achievement Test (NAT) and its application, utmost attention was given to ethical standards. Prior to administering the survey, informed consent was obtained from all participating teachers, ensuring they were fully aware of the study's purpose, procedures, and implications. Additionally, measures were implemented to guarantee the confidentiality of participants' responses, safeguarding their privacy and anonymity. The teacher questionnaire, serving as a standardized research tool, was utilized with permission from the author, Don Klinger, ensuring proper attribution and adherence to intellectual property rights. Furthermore, the acquisition of NAT data from 2016 and 2019 was done through formal requests to the Federal Ministry of Federal Education, ensuring compliance with legal and regulatory requirements regarding data access and usage. Throughout the survey process, ethical guidelines were adhered to, and all relevant ethical documents were considered, prioritizing the protection of participants' rights and welfare.

3.15 Data Interpretation

Upon data collection, the interpretation process was carried out using SPSS version 20.0. These techniques encompassed the following mentioned in the summary table 3.13 summary table.

3.13 *Summary Table*

Research Objective	Research Design	Data Analysis
To assess the psychometric characteristics of the National Achievement Test (NAT) and assess its effectiveness as an assessment tool by using the Classical Test Theory (CTT) Framework.	Quantitative method	Through CTT model

To assess the perception of teachers in the utilization of National Achievement Test (NAT) results and identify effective strategies for applying NAT data to improve student achievement.	Quantitative design	Means
To obtain objective and accurate information about the application of the National Achievement Test (NAT) by considering the four functions of Large Scale Assessment (LSA), namely accountability, gatekeeping, instructional identification, and monitoring of student progress through teacher's perception.	Qualitative design	Percentage method
To propose a comprehensive model for effectively disseminating NAT results to	Check list /observation	Document analysis

various stakeholders, facilitating their access to and understanding of the assessment outcomes.	

CHAPTER 4

DATA ANALYSIS AND INTERPERTATION

This chapter includes analysis techniques used for this study and interpretations of results for studying the analysis of the development and application of National achievement test. In this study, an ex-post facto method was employed, enabling researchers to scrutinize pre-existing data from the National Achievement Tests conducted in 2016 and 2019. The objective was to assess the consistency of the instrument used and its alignment with the intended measurement objectives. Apart from it, another standardized tool was adapted to identify the teacher perception regarding NAT and its application. Data was collected from teachers those who participated in NAT 2016 & 2019 from different school of AJK and Islamabad. Researcher utilized SPSS (Statistical Package for Social Sciences) 21 to analyze the data. The analysis and interpretations were aligned with the study objectives. This chapter is divided into three sections. The first section presents ex post facto analysis applied to the achievement test, addressing the study's initial objective. Section two based on objective two and three the response of teacher regarding their perception regarding NAT and its application. The last section deals with objective four of the study.

OBJECTIVE 1: To provide a comprehensive critically analysis psychometric analysis of NAT 2016 and 2019 by using IRT framework

4.1 Interpretation of Difficulty Index

The Item Difficulty Index, also known as "item easiness," quantifies the probability of obtaining a correct answer on a particular test and typically falls within a range of 0.0 to 1.0. A high difficulty score indicates that a significant portion of the sample answered the question correctly, whereas a lower difficulty score suggests that a smaller proportion of the sample comprehended and answered the question accurately (Boateng *et al*., 2018). The researchers have employed the following criteria to identify the different levels of individual performance on given questions and to analyze the difficulty index of each item.

Table 4.1 ***Item Difficulty Index of Sciences Test of Grade 8 for AJK and ICT in 2016***

Item no.	AJK	ICT	Item No.	AJK	ICT
GS Test item-1	.52	.54	**GS Test item-22**	.43	.51
GS Test item-2	.51	.55	**GS Test item-23**	.40	.45
GS Test item-3	.43	.46	**GS Test item-24**	.52	.52
GS Test item-4	.50	.54	**GS Test item-25**	.40	.39
GS Test item-5	.53	.59	**GS Test item-26**	.46	.59
GS Test item-6	.49	.53	**GS Test item-27**	.29	.45
GS Test item-7	.56	.61	**GS Test item-28**	.30	.35
GS Test item-8	.55	.65	**GS Test item-29**	.28	.34
GS Test item-9	.38	.48	**GS Test item-30**	.27	.42
GS Test item-10	.53	.58	**GS Test item-31**	.36	.36
GS Test item-11	.28	.36	**GS Test item-32**	.50	.49
GS Test item-12	.44	.48	**GS Test item-33**	.50	.53
GS Test item-13	.25	.31	**GS Test item-34**	.47	.53
GS Test item-14	.36	.39	**GS Test item-35**	.48	.46
GS Test item-15	.33	.34	**GS Test item-36**	.50	.49
GS Test item-16	.40	.46	**GS Test item-37**	.50	.53
GS Test item-17	.47	.52	**GS Test item-38**	.47	.53

GS Test item-18	.62	.70	**GS Test item-39**	.22	.21
GS Test item-19	.42	.41	**GS Test item-40**	.48	.46
GS Test item-20	.56	.61	**GS Test item-41**	.40	.39
GS Test item-21	.38	.40	**GS Test item-42**	.37	.45
			Overall Mean	0. 0.431	0. 0.475

Table 4.2 ***Evaluation of Difficulty Index for Test Score***

Difficulty Index	**No. of Items**		**Evaluation of Item**	**Decision**
	AJK	**ICT**		
<0.30	6	1	Difficult	Revise or Discard
0.30-0.70	36	41	Moderately Difficult/ Moderately Easy	Retain
>0.70	-	-	Easy	Revise/Discard

Table 4.1 and 4.2 presents the difficulty index of test items in the Science subject for Grade 8 students from AJK and ICT provinces. The difficulty index indicates the level of difficulty for each question. According to the index, question numbers 11, 13, 27, 29, 30, and 39 are found to be the most difficult for AJK students, suggesting a need for reconsideration or revision of these questions. On the other hand, only question number 39 falls within the range of difficulty index (<0.30) for ICT students, indicating that this particular question is difficult or unclear and should be reviewed or removed from the test. However, the remaining 36 items included in the Science test for AJK students are considered moderately difficult or moderately easy and should be retained. Similarly, for ICT students, 41 items fall within the range of difficulty index (0.30-0.70), indicating moderate difficulty or ease. Overall, the difficulty index provides insights into the appropriateness of test items and suggests potential revisions or removal of certain questions for improved assessment

Table 4.3 ***Evaluation of discarded item domain wise***

Item no	**Difficulty Found in AJK and ICT**	**Domain**
11,13,27,30,39	AJK	Physical Science
29	AJK	Life Science
39	ICT	Physical Science

The Table 4.3 shows that domain wise discarded test item of Science of grade 8 from AJK and ICT The question no. 11, 13, 27,30 and 39 which are supposed to be discarded belong to “Physical Sciences” domain and test item number 29 belongs to Life Science. The difficulty Index of these questions which belong from different domains lies under the criteria <0.30, therefore these questions of Science test should be revised or discarded from Item bank

Table 4.4. ***Item Difficulty Index of Mathematics Test of Grade 8 for AJK and ICT 2019***

Item no.	**AJK**	**ICT**	**Item no.**	**AJK**	**ICT**
Math Test item-1	.65	.66	Math Test item-18	.38	.40
Math Test item-2	.19	.27	Math Test item-19	.32	.26
Math Test item-3	.56	.64	Math Test item-20	.23	.32
Math Test item-4	.53	.57	Math Test item-21	.35	.37
Math Test item-5	.46	.55	Math Test item-22	.45	.49
Math Test item-6	.44	.60	Math Test item-23	.37	.48
Math Test item-7	.46	.53	Math Test item-24	.41	.53
Math Test item-8	.61	.61	Math Test item-25	.29	.33
Math Test item-9	.38	.42	Math Test item-26	.27	.25
Math Test item-10	.47	.52	Math Test item-27	.35	.44
Math Test item-11	.46	.44	Math Test item-28	.34	.37
Math Test item-12	.43	.45	Math Test item-29	.26	.30
Math Test item-13	.27	.31	Math Test item-30	.40	.36
Math Test item-14	.28	.26	Math Test item-31	.42	.49

Math Test item-15	.39	.41	Math Test item-32	.25	.25
Math Test item-16	.43	.56			
Math Test item-17	.38	.40	**Over all mean**	**0.45**	**0.49**

Table 4.5 ***Evaluation of Difficulty Index for Test Score***

Difficulty Index	No. of Items		Evaluation of Item	Decision
	AJK	ICT		
<0.30	8	5	Difficult	Revise or Discard
0.30-0.70	24	27	Moderately Difficult/ Moderately Easy	Retain
>0.70	-	-	Easy	Revise/Discard

Table 4.4 and 4.5 shows that test item no. 2,13,14,20,25,26,29 and 32 are fell into difficulty criteria of difficult index range. The table 4.5 highlights that 8 Items of mathematics is difficult for the AJK student. While, test items no. 2,14,19,26 and 32 are lie under the range of difficulty index (>0.30) for student of ICT which indicates that 5 items of mathematics test for ICT students are difficult questions, thus, these 5 items should be revised or discarded from questionnaire. Whereas rest 24 items of mathematics test for AJK students are moderately difficult or moderately easy questions. However, under the difficulty index range (0.30-0.70) 27 questions for student of ICT are moderately difficult or moderately easy questions and high proportion of sample gave answers to these questions correctly. Therefore, these questions must be included in questionnaire.

4.6 ***Evaluation of Discarded item domain wise from Difficulty Index of Mathematics Test Grade 8***

Item No	Difficulty found in AJK & ICT	Domain
14,29	Both	Algebra
2,32	Both	Geometry & Measurement
13,20,25	AJK	Numbers and operation
26	Both	Information Handling

Table 4.6 demonstrates that the discarded items belong to different mathematical domains, indicating areas where students faced significant challenges. Specifically, item numbers 2 and 32 fall under the **"**Geometry & Measurement" domain, highlighting difficulties in understanding spatial concepts and measurements. Item number 13,20,25 belongs to the "Numbers & Operation" domain, suggesting gaps in numerical computations and operations. Similarly, item number 14,29 is categorized under "Algebra"**,** reflecting challenges in algebraic reasoning. Additionally, items 26 and 28 are from the "Information Handling" domain, indicating issues with data interpretation and statistical concepts. Since these items fall within the difficulty index range of (<0.30), they need to be revised and ultimately discarded from the item bank to ensure the test maintains an appropriate level of fairness and validity.

Table 4.7 ***Item Difficulty Index of Mathematics Test of Grade 4 for AJK and ICT in 2016***

Item no.	AJK	ICT	Item no.	AJK	ICT
Math Test item-1	.65	.66	Math Test item-26	.35	.44
Math Test item-2	.19	.27	Math Test item-27	.34	.37
Math Test item-3	.56	.64	Math Test item-28	.26	.30
Math Test item-4	.53	.57	Math Test item-29	.40	.36
Math Test item-5	.46	.55	Math Test item-30	.42	.49
Math Test item-6	.44	.60	Math Test item-31	.25	.25

Math Test item-7	.46	.53	Math Test item-32	.32	.36
Math Test item-8	.61	.61	Math Test item-33	.25	.27
Math Test item-9	.38	.42	Math Test item-34	.25	.30
Math Test item-10	.47	.52	Math Test item-35	.30	.30
Math Test item-11	.46	.44	Math Test item-36	.24	.34
Math Test item-12	.43	.45	Math Test item-37	.36	.34
Math Test item-13	.27	.31	Math Test item-38	.26	.38
Math Test item-14	.28	.26	Math Test item-39	.37	.40
Math Test item-15	.39	.41	Math Test item-40	.25	.34
Math Test item-16	.43	.56	Math Test item-41	.20	.22
Math Test item-17	.38	.40	Math Test item-42	.35	.45
Math Test item-18	.32	.26	Math Test item-43	.22	.30
Math Test item-19	.23	.32	Math Test item-44	.24	.26
Math Test item-20	.35	.37	Math Test item-45	.18	.19
Math Test item-21	.45	.49	Math Test item-46	.25	.35
Math Test item-22	.37	.48	Math Test item-47	.28	.34
Math Test item-23	.41	.53	Math Test item-48	.49	.46
Math Test item-24	.29	.33	Math Test item-49	.15	.20
Math Test item-25	.27	.25	Math Test item-50	.34	.23
			Overall Mean	**0.6076**	**0.6504**

Table 4.8 ***Evaluation of Difficulty Index for Test Score***

Difficulty Index	No. of Items		Evaluation of Item	Decision
	AJK	ICT		
<0.30	20	11	Difficult	Revise or Discard
0.30-0.70	30	39	Moderately Difficult/ Moderately Easy	Retain
>0.70	-	-	Easy	Revise/Discard

Table 4.7 and 4.8 shows that test item no. 2,13,14,19,24,25,28,31,33,34,36,38,40,41,43,44,45,46,47, and 49 are fell into difficulty criteria of difficult index range. in table 4.8 , shows that 20 questions for AJK students of mathematics of

grade 4 are difficult questions, thus, these should not be included in questionnaire. While, 2,14,18,25,31,33,41,44,45,49 and 50 are lie under the range of difficulty index (<0.30) for student of ICT province indicates that 11 questions of mathematics test for ICT students are difficult questions, thus, these 11 questions should be revised or discarded from questionnaire. Whereas rest of the 30 questions of mathematics test for AJK students are moderately difficult or moderately easy questions. However, under the difficulty index range (0.30-0.70) 39 questions for student of ICT are moderately difficult or moderately easy questions and high proportion of sample gave answers to these questions correctly. Therefore, these questions must be included in questionnaire.

Table 4.9 ***Evaluation of Discarded item domain wise from Difficulty Index of Mathematics Test Grade 4***

Item No	Difficulty Index found in AJK & ICT	Domain
2,41	Both	Geometry & measurement
19,36,38	AJK	Geometry & Measurement
50	1CT	Geometry & Measurement
14,25,31,44,45,49	Both	Number & Operation
13,24,43,46	AJK	Number & Operation
18	ICT	Number & Operation
28,34,40	AJK	Information Handling
33	Both	Information Handling

Table 4.9 provides an evaluation of discarded items from the Mathematics Test for Grade 4, categorized by domain according to their Difficulty Index. The table outlines the item number, the respective Difficulty Index found in Azad Jammu and Kashmir (AJK) and Islamabad Capital

Territory (ICT), and the domain associated with each item. Items 2 and 41 were discarded due to difficulty index concerns and were identified in both AJK and ICT regions, specifically in the domain of Geometry & Measurement. Additionally, items 19, 36, and 38 were discarded in AJK, also falling under the Geometry & Measurement domain. Item 50, discarded in ICT, was linked to the Geometry & Measurement domain as well. Items 14, 25, 31, 44, 45, and 49 were discarded due to difficulty index issues, found in both AJK and ICT, and belonged to the Number & Operation domain. Furthermore, items 13, 24, and 43, 46 were discarded in AJK within the Number & Operation domain, while item 18 was discarded in ICT in the same domain. Regarding the Information Handling domain, items 28, 34, and 40 were discarded in AJK, while item 33 was discarded in both AJK and ICT due to difficulty index concerns.

Table 4.10 ***Item Difficulty Index of Sciences Test of Grade 4 for AJK and ICT 2019***

Item no.	AJK	ICT	Item no.	AJK	ICT
GS test item -1	.62	.65	GS test item -19	.29	.36
GS test item -2	.46	.43	GS test item -20	.32	.34
GS test item -3	.49	.49	GS test item -21	.29	.30
GS test item -4	.45	.59	GS test item -22	.33	.40
GS test item -5	.46	.49	GS test item -23	.29	.38
GS test item -6	.35	.69	GS test item -24	.45	.36
GS test item -7	.35	.31	GS test item -25	.31	.29
GS test item -8	.52	.32	GS test item -26	.35	.27
GS test item -9	.40	.39	GS test item -27	.32	.31
GS test item -10	.40	.42	GS test item -28	.30	.23
GS test item -11	.41	.42	GS test item -29	.34	.21
GS test item -12	.31	.23	GS test item -30	.25	.20
GS test item -13	.37	.38	GS test item -31	.20	.27
GS test item -14	.35	.46	GS test item -32	.31	.29
GS test item -15	.35	.47			
GS test item -16	.38	.40	**Total mean**	**0.359**	**0.379**
GS test item -17	.29	.40			
GS test item -18	.18	.10			

Table 4.11 ***Evaluation of Difficulty Index for Test Score***

Difficulty Index	**No. of Items**		**Evaluation of Item**	**Decision**
	AJK	**ICT**		
<0.30	8	10	Difficult	Revise or Discard
0.30-0.70	24	22	Moderately Difficult/ Moderately Easy	Retain
>0.70	-	-	Easy	Revise/Discard

Table 4.10 and 4.11 shows the difficulty index of the test item of Science's students. The index shows that Items no. 17,18,19,21,23,28,30, and 31 are most difficult question for AJK students of Sciences, consequently, the question should be discarded or revised. While test items no. 12,18,21,25,26,28,29,30,31, and 32 are lie under the range of difficulty index (>0.30) indicates that 10 questions of test for ICT students are difficult questions, therefore these 10 questions of general science test should be reconsidered to revised or discarded from test. However, rest 24 questions included of science test for AJK students are moderately difficult or moderately easy questions, thus, these 24 questions should be retained in test. However, under the difficulty index range (0.30-0.70) there are 22 questions for ICT students are moderately difficult or moderately easy questions and high proportion of sample gave answer to these questions correctly.

4.12 ***Evaluation of Discarded item from Difficulty Index of Sciences Test Grade 4***

Items	**Difficulty Index found in AJK/ICT**	**Domain**
18,30,31,28,32	Both	Physical Science
12,29	ICT	Physical Science
17,23	AJK	Physical Science
21	Both	Life Science

19	AJK	Life Science
25,26	ICT	Life Science

The **Table 4.12** shows the evaluation of discarded items from the difficulty index of the Sciences Test for Grade 4 highlights variations in student performance across AJK and ICT. Several items from the **Physical Science** domain (18, 30, 31, 28, and 32) were found to be difficult in both regions, while items 12 and 29 were problematic only in ICT, and items 17 and 23 posed challenges specifically for AJK students. In the **Life Science** domain, item 21 was difficult for students in both regions, whereas item 19 was discarded only in AJK. Additionally, items 25, 26, were removed in ICT but not in AJK, indicating region-specific difficulties. These findings suggest that students struggled more with Physical Science concepts compared to Life Science, and that regional differences in understanding may necessitate a more tailored approach to assessment design.

Table 4.13 ***Comparison of NAT 2016 and 2019 Science Test of Grade 4 and 8 based on Difficulty Index Results***

Analysis	**2016 Science Test Grade 8**	**2019 Science Test Grade 8**
Overall Mean Difficulty Index	AJK: 0.431 ICT: 0.475	AJK: 0.359 ICT: 0.379
Item Difficulty Distribution Evaluation of Difficulty	Range: 0.00 - 0.70 Difficult: 6 items AJK, 1 item ICT	Range: 0.10 - 0.69 Difficult: 8 items AJK and 10 ICT

	Moderately Difficult/Moderately Easy: 36 items AJK ,41 ICT	Moderately Difficult/Moderately Easy: 24 items AJK,27 items ICT
	Easy: None	Easy: None
Recommendations for Test Development	Improvement/revision needed	Improvement/revision needed

Table 4.13 showed the comparison between the 2016 and 2019 Science Test Grade 8 reveals notable similarities and differences. In 2016, the overall mean difficulty index for AJK was 0.431, and for ICT, it was 0.475 and in 2019, for ICT, it was 0.359 AJK and 0.379 ICT indicating a moderate level of difficulty for both regions. The item difficulty distribution ranged from 0.00 to 0.70, suggesting variability in difficulty levels. Similarly, in 2019, the item difficulty distribution ranged from 0.10 to 0.69, reflecting similar variability but with a slightly shifted range compared to 2016. While both years saw recommendations for improvement and revision in test development, the evaluation of difficulty showed some distinctions. In 2016, there were 6 items ajk and 1 item ICT classified as difficult and 36 items and 41 ICT as moderately difficult/moderately easy, while in 2019, there were 8 AJK and 10 ICT difficult items and 24 AJK and 22 ICT moderately difficult/moderately easy items, indicating a slight increase in difficult items in 2019. These differences highlight potential shifts in test construction and emphasize the ongoing need for refinement to ensure consistent and reliable assessment in Science for Grade 8 students across AJK and ICT.

Table 4.14 ***Comparison of NAT 2016 and 2019 Mathematics test of Grade 4 and 8 based on Difficulty Index Results***

Analysis	2016 Mathematics Test Grade 4	2019 Mathematics Test Grade 8
Overall Mean Difficulty Index	AJK: 0.607 ICT: 0.650	AJK 0.451 ICT O.496
Item Difficulty Distribution	Range: 0.15 - 0.65	Range: 0.19 - 0.65
Evaluation of Difficulty	Difficult: 20 items AJK, 11 items ICT Moderately Difficult/Moderately Easy: 30 items AJK and 39 items ICT Easy: None	Difficult: 8 items AJK,5 items ICT Moderately Difficult/Moderately Easy: 24 items AJK and 27 items ICT Easy: None
Recommendations for Test Development	Improvement/revision needed	Improvement/revision needed

Table 4.14 shows the comparison between the 2016 Mathematics Test Grade 4 and the 2019 Mathematics Test Grade 8 reveals several important insights into the evolution of these assessments over time. Firstly, the overall mean difficulty index for both AJK and ICT increased from 2016 to 2019, indicating that the tests became slightly more challenging over this period. Specifically, in 2019, the overall mean difficulty index was 0.45 for AJK and 0.49 for ICT, whereas in 2016, it was 0.6076 for AJK and 0.6504 for ICT. This suggests a shift towards higher difficulty levels in the later assessment. Additionally, while the range of item difficulty remained consistent between the two years, spanning from 0.15 to 0.65 in 2016 and 0.19 to 0.65 in 2019, there were notable differences in the evaluation of difficulty. In 2016, there were 20 items AJK and 11 ICT classified as difficult and 30 items and 39 ICT as moderately difficult/moderately easy, whereas

in 2019, there were 8 AJK and 5 items ICT difficult items and 24 AJK and 27 ICT moderately difficult/moderately easy items. This indicates a decrease in the number of difficult items and a more balanced distribution of difficulty levels in the later assessment.

4.2 Interpretation of Discriminant Index

The discrimination index of items also known as an "item-effectiveness test" measures the extent to which an item accurately distinguishes between high-achieving and low-achieving students. (Rezigalla, 2022). The item discriminant index can be assessed by conducting a correlation analysis between the overall criterion and performance on a particular item, as indicated by Whiston in 2016. This analysis can be performed using either the point biserial correlation coefficient (rpb) or the phi coefficient (Sim & Rasiah, 2006). Under the Classical Test Theory (CTT) framework, Item discrimination represents a slope parameter that gauges the rate at which the likelihood of providing a correct response changes with an increase in proficiency or the measured attribute (Bazaldua, 2017). To analyze and interpret the discriminant index the study used the following table.

Table 4.15 ***Item Discriminant Index of Mathematics Test of Grade 8 for AJK and ICT 2019***

Item no.	AJK		ICT	
	r_{pb}	Sig	r_{pb}	Sig
Math test item-1	$.344^{**}$	0.000	$.298^{**}$	0.000
Math test item-2	$.161^{**}$	0.000	$.249^{**}$	0.000
Math test item-3	$.333^{**}$	0.000	$.248^{**}$	0.000
Math test item-4	$.356^{**}$	0.000	$.412^{**}$	0.000
Math test item-5	$.421^{**}$	0.000	$.443^{**}$	0.000

Math test item-6	$.338^{**}$	0.000	$.389^{**}$	0.000
Math test item-7	$.433^{**}$	0.000	$.439^{**}$	0.000
Math test item-8	$.217^{**}$	0.000	$.300^{**}$	0.000
Math test item-9	$.344^{**}$	0.000	$.291^{**}$	0.000
Math test item-10	$.402^{**}$	0.000	$.438^{**}$	0.000
Math test item-11	$.407^{**}$	0.000	$.417^{**}$	0.000
Math test item-12	$.374^{**}$	0.000	$.453^{**}$	0.000
Math test item-13	$.473^{**}$	0.000	$.519^{**}$	0.000
Math test item-14	$.208^{**}$	0.000	$.388^{**}$	0.000
Math test item-15	$.286^{**}$	0.000	$.357^{**}$	0.000
Math test item-16	$.361^{**}$	0.000	$.372^{**}$	0.000
Math test item-17	$.400^{**}$	0.000	$.434^{**}$	0.000
Math test item-18	$.463^{**}$	0.000	$.463^{**}$	0.000
Math test item-19	$.121^{**}$	0.001	$.285^{**}$	0.000
Math test item-20	$.312^{**}$	0.000	$.438^{**}$	0.000
Math test item-21	$.375^{**}$	0.000	$.476^{**}$	0.000
Math test item-22	$.412^{**}$	0.000	$.535^{**}$	0.000
Math test item-23	$.397^{**}$	0.000	$.413^{**}$	0.000
Math test item-24	$.365^{**}$	0.000	$.418^{**}$	0.000
Math test item-25	$.282^{**}$	0.000	$.361^{**}$	0.000
Math test item-26	$.321^{**}$	0.000	$.405^{**}$	0.000
Math test item-27	$.309^{**}$	0.000	$.414^{**}$	0.000
Math test item-28	$.313^{**}$	0.000	$.448^{**}$	0.000
Math test item-29	$.347^{**}$	0.000	$.365^{**}$	0.000
Math test item-30	$.353^{**}$	0.000	$.416^{**}$	0.000
Math test item-31	$.347^{**}$	0.000	$.365^{**}$	0.000
Math test item-32	$.346^{**}$	0.000	$.365^{**}$	0.000

Table no.4.15 indicates that question item no. 2, and 19 of Mathematics test are marginal item but not discriminant therefore these items should be revise or discarded from questions list. Under the range of 0.20-0.29, 4 items from AJK and 5 items from ICT are moderately discriminant or fair item that should be considered. Subsequently, there are 17, 8 items of AJK and ICT, respectively, lie under the range of 0.30-0.39, are good or discriminant items. Furthermore, under the criteria of ≥0.40, 8, 18 items from AJK and ICT, respectively, are very good discriminant or items.

Table 4.16

Item Discriminant Index Evaluation of mathematics 2016

Discriminant Index (r_{pb})	**No. of Items**		**Evaluation of Item**	**Decision**
	AJK	**ICT**		
Negative	0	0	Worst/ Defective Item	Discard
<0.20	2	0	Marginal Item, not Discriminant	Revise/ Discard
0.20-0.29	4	5	Moderately Discriminant/ Fair item	Retain
0.30-0.39	17	08	Discriminant Item, Good item	Retain
≥0.40	8	18	Very Discriminating, very good Item	Retain

Table 4.16 indicates that 2 question items from AJK and 0 from ICT in the Grade 8 Mathematics **t**est are marginal items (Discriminant Index < 0.20) and do not effectively distinguish between high- and low-performing students. Therefore, these items should be revised or discarded**.**

For items with a Discriminant Index between 0.20-0.29, a total of 4 items from AJK and 5 items from ICT fall into the moderately discriminating (fair) category**.** These items should be considered for retention with possible improvements. Additionally, 17 items from AJK and 8 items from ICT fall within the 0.30-0.39 range**,** classifying them as good or discriminant items. These items should be retained as they effectively differentiate between student performance levels.

Furthermore, under the ≥0.40 category**,** 8 items from AJK and 18 items from ICT are classified **as** highly discriminating and very good items**.** These items are the most effective in assessing student performance and should be retained in the questionnaire without modification**.**

Table 4.17 ***Item Discriminant Index of G. Sciences Test of Grade 8 for AJK and ICT 2016***

Item no.	AJK		ICT		Item no.	AJK		ICT	
	r_{pb}	Sig	r_{pb}	Sig	Sig	r_{pb}	Sig	r_{pb}	sig
GS test item-1	.417**	0.000	.367**	0.000	**GS test item-22**	.399**	0.000	.401**	0.000
GS test item-2	.401**	0.000	.301**	0.000	**GS test item-23**	.478**	0.000	.459**	0.000
GS test item-3	.298**	0.000	.299**	0.000	**GS test item-24**	.344**	0.000	.328**	0.000
GS test item-4	.480**	0.000	.542**	0.000	**GS test item-25**	.315**	0.000	.337**	0.000
GS test item-5	.480**	0.000	.498**	0.000	**GS test item-26**	.479**	0.000	.455**	0.000
GS test item-6	.460**	0.000	.406**	0.000	**GS test item-27**	.429**	0.000	.417**	0.000
GS Test item-7	.352**	0.000	.468**	0.000	**GS test item-28**	.450**	0.000	.476**	0.000
GS Test item-8	.360**	0.000	.317**	0.000	**GS test item-29**	.356**	0.000	.358**	0.000
GS test item-9	.414**	0.000	.439**	0.000	**GS test item-30**	.426**	0.000	.485**	0.000
GS test item-10	.443**	0.000	.424**	0.000	**GS test item-31**	.246**	0.000	.355**	0.000
GS test item-11	.153**	0.000	.190**	0.000	**GS test item-32**	.433**	0.000	.509**	0.000
GS test item-12	.427**	0.000	.384**	0.000	**GS test item-33**	.415**	0.000	.477**	0.000

GS test item-13	.377**	0.000	.341**	0.000	**GS test item-34**	.484**	0.000	.499**	0.000
GS test item-14	.440**	0.000	.395**	0.000	**GS test item-35**	.452**	0.000	.518**	0.000
GS test item-15	.416**	0.000	.364**	0.000	**GS test item-36**	.433**	0.000	.509**	0.000
GS test item-16	.268**	0.000	.191**	0.000	**GS test item-37**	.415**	0.000	.477**	0.000
GS test item-17	.294**	0.000	.319**	0.000	**GS test item-38**	.485**	0.000	.499**	0.000
GS test item-18	.298**	0.000	.366**	0.000	**GS test item-39**	.176**	0.000	.189**	0.000
GS test item-19	.178**	0.001	.141**	0.000	**GS test item-40**	.452**	0.001	.518**	0.000
GS test item-20	.261**	0.000	.256**	0.000	**GS test item-41**	.315**	0.000	.337**	0.000
GS test item-21	0.019	0.000	-0.002	0.000	**GS test item-42**	.398**	0.000	.465**	0.000

Notes; **. Correlation is significant at the 0.01 level (2-tailed). *. Correlation is significant at the 0.05 level (2-tailed).

Table 4.18 ***Item Discriminant Index Evaluation of General Science 2016***

Discriminant Index (r_{pb})	No. of Items		Evaluation of Item	Decision
	AJK	**ICT**		
Negative	0	1	Worst/ Defective Item	Discard
<0.20	4	4	Marginal Item, not Discriminant	Revise/ Discard
0.20-0.29	6	3	Moderately Discriminant/ Fair item	Keep
0.30-0.39	9	14	Discriminant Item, Good item	Keep
≥0.40	23	20	Very Discriminating, very good Item	Keep

Table no. 4.17 & 4.18 demonstrate that question item no. 23 of General Science test is worst item therefore this item should be discarded. Under the range of <0.20, 4 items from AJK and 4 items from ICT are marginal item or not discriminant these should be revised or discarded from questionnaire. Subsequently, there are 6 and 3 items of AJK and ICT, respectively, lie under the range of 0.20-0.29, are moderate discriminant or fair items therefore these items should be maintained in questionnaire. Furthermore, 9 items from AJK list and 14 from ICT questionnaire list lie under the criteria of good item of discriminant 0.30-0.39. While, for ≥0.40 range of very good discriminant, 23 items from AJK and 20 items from ICT are very good discriminant or items.

Table 4.19 ***Item Discriminant Index of Math's Test of Grade 4 for AJK and ICT in 2016***

Item no.	**AJK**		**ICT**		**Item no.**	**AJK**		**ICT**	
	r_{pb}	**Sig**	**r_{pb}**	**Sig**		**r_{pb}**	**Sig**	**r_{pb}**	**sig**
Math test item-1	.336**	0.000	.282**	0.000	**Math test item-26**	.312**	0.000	.394**	0.000
Math test item-2	.135**	0.000	.219**	0.000	**Math test item-27**	.299**	0.000	.416**	0.000
Math test item-3	.308**	0.000	.233**	0.000	**Math test item-28**	.321**	0.000	.457**	0.000
Math test item-4	.321**	0.000	.381**	0.000	**Math test item-29**	.339**	0.000	.373**	0.000
Math test item-5	.383**	0.000	.396**	0.000	**Math test item-30**	.331**	0.000	.389**	0.000
Math test item-6	.314**	0.000	.365**	0.000	**Math test item-31**	.346**	0.000	.371**	0.000
Math test item-7	.381**	0.000	.417**	0.000	**Math test item-32**	.123**	0.000	.319**	0.000
Math test item-8	.217**	0.000	.275**	0.000	**Math test item-33**	.403**	0.000	.456**	0.000
Math test item-9	.302**	0.000	.273**	0.000	**Math test item-34**	.312**	0.000	.325**	0.000
Math test item-10	.366**	0.000	.410**	0.000	**Math test item-35**	.248**	0.000	.351**	0.000
Math test item-11	.370**	0.000	.400**	0.000	**Math test item-36**	.265**	0.000	.270**	0.000
Math test item-12	.338**	0.000	.406**	0.000	**Math test item-37**	.169**	0.000	.358**	0.000

Math test item-13	.438**	0.000	.483**	0.000	**Math test item-38**	.277**	0.000	.354**	0.000
Math test item-14	.202**	0.000	.386**	0.000	**Math test item-39**	.234**	0.000	.409**	0.000
Math test item-15	.249**	0.000	.349**	0.000	**Math test item-40**	.370**	0.000	.451**	0.000
Math test item-16	.330**	0.000	.344**	0.000	**Math test item-41**	.142**	0.000	.216**	0.000
Math test item-17	.366**	0.000	.408**	0.000	**Math test item-42**	.281**	0.000	.240**	0.000
Math test item-18	.468**	0.000	.485**	0.000	**Math test item-43**	.337**	0.000	.331**	0.000
Math test item-19	.106**	0.001	.273**	0.000	**Math test item-44**	.145**	0.001	.330**	0.000
Math test item-20	.302**	0.000	.421**	0.000	**Math test item-45**	.300**	0.000	.440**	0.000
Math test item-21	.370**	0.000	.464**	0.000	**Math test item-46**	.158**	0.000	.264**	0.000
Math test item-22	.411**	0.000	.499**	0.000	**Math test item-47**	.209**	0.000	.352**	0.000
Math test item-23	.358**	0.000	.365**	0.000	**Math test item-48**	.138**	0.000	.214**	0.000
Math test item-24	.344**	0.000	.415**	0.000	**Math test item-49**	.255**	0.000	.389**	0.000
Math test item-25	.255**	0.000	.353**	0.000	**Math test item-50**	0.031	0.000	.322**	0.000

Table 4.20 ***Item Discriminant Index Evaluation of mathematics 2016***

Discriminant Index (r_{pb})	**No. of Items**		**Evaluation of Item**	**Decision**
	AJK	**ICT**		
Negative	0	0	Worst/ Defective Item	Discard
<0.20	9	0	Marginal Item, not Discriminant	Revise/ Discard
0.20-0.29	12	11	Moderately Discriminant/ Fair item	Keep

0.30-0.39	25	22	Discriminant Item, Keep Good item
≥0.40	04	17	Very Discriminating, Keep very good Item

The tables 4.19 and 4.20 provide a detailed analysis of the item discriminant index evaluation for the Grade 4 Mathematics test in AJK and ICT. These tables offer valuable insights into the effectiveness of each test item in differentiating between high and low-performing students.

For AJK, the items with the highest discriminant power include item numbers 13, 18, and 22, with correlations of .438**, .468**, and .411**, respectively. These values indicate strong discrimination and statistical significance at the 0.01 level, making these items highly effective in distinguishing between students of different ability levels. Consequently, these items should be retained in the test without modifications.

Conversely, item numbers 19, 32, 41, 44, and 46 exhibited weaker discriminant power, with correlations of .106**, .123**, .142**, .145**, and .158**, respectively. While these items are statistically significant, their lower values suggest they are less effective in differentiating student performance. These items should be reviewed carefully, revised to improve their discriminatory ability, or discarded if necessary.

For ICT, the most discriminating items include item numbers 13, 18, 22, 27, 28, and 40, with correlations of .483**, .485**, .499**, .416**, .457**, and .451**, respectively. These items have very high discriminant power, making them crucial for maintaining the integrity of the test and should be retained.

However, items with lower discriminant values, such as item numbers 32, 41, 42, 44, and 48, with correlations of .319**, .216**, .240**, .330**, and .214**, respectively, demonstrate relatively weaker discrimination. While some of these values still fall within an acceptable range, careful revision is recommended to enhance their effectiveness. Items with a discriminant index lower than .20 should be considered for revision or removal to ensure the overall quality of the test.

Based on the summary from Table 4.20, AJK had 9 items with a discriminant index of less than 0.20, suggesting they are marginal and need revision or removal. In contrast, ICT had no items in this category, indicating a generally stronger test performance. The number of moderately discriminant items (0.20-0.29) was slightly higher in AJK (12) compared to ICT (11), while ICT had a significantly higher number of very good discriminators (≥0.40) with 17 items compared to just 4 in AJK.

Overall, the findings suggest that while the test demonstrates strong discriminant validity, there is room for improvement, particularly in the AJK version of the test. It is recommended to refine weaker items by reviewing their content, response options, and difficulty levels to enhance their ability to distinguish student proficiency effectively.

Table 4.21 ***Item Discriminant Index of Science Test of Grade 4 for AJK and ICT 2019***

Item no.	**AJK**		**ICT**	
	r_{pb}	**Sig**	**r_{pb}**	**Sig**
GS test item-1	$.147^{*}$	0.159	$.256^{**}$	0.000
GS test item-2	$.419^{**}$	0.000	$.394^{**}$	0.000
GS test item -3	$.453^{**}$	0.000	$.348^{**}$	0.000
GS test item -4	$.401^{**}$	0.000	$.303^{**}$	0.000
GS test item -5	$.348^{**}$	0.001	$.206^{**}$	0.000

GS test item -6	.560**	0.000	.337**	0.000
GS test item -7	.333**	0.001	.315**	0.000
GS test item -8	.364**	0.000	.290**	0.000
GS test item -9	.385**	0.000	.229*	0.000
GS test item -10	.467**	0.000	.335**	0.000
GS test item -11	.432**	0.000	.243**	0.000
GS test item -12	-.249*	0.016	.103*	0.021
GS test item -13	.253*	0.014	.239**	0.000
GS test item -14	.432**	0.000	.243**	0.000
GS test item -15	.372**	0.000	.338**	0.000
GS test item -16	.116	0.269	.212**	0.000
GS test item -17	.400**	0.000	.223**	0.000
GS test item -18	.182	0.081	.172**	0.000
GS test item -19	.348**	0.001	.293**	0.000
GS test item -20	.284**	0.006	.340**	0.000
GS test item -21	.176	0.091	.255**	0.000
GS test item -22	.412**	0.000	.334**	0.000
GS test item -23	.228*	0.028	.330**	0.000
GS test item -24	-0.012	0.911	.081	0.071
GS test item -25	.296**	0.004	.443**	0.000
GS test item -26	.298**	0.004	.369**	0.000
GS test item -27	.324**	0.002	.303**	0.000
GS test item -28	.386**	0.000	.227**	0.000
GS test item -29	.221*	0.034	.223**	
				0.000
GS test item -30	.052	0.622	.056	0.214
GS test item -31	.158	0.130	.159**	0.000
GS test item -32	.296**	0.004	.443**	0.000
Overall Mean	**0.298**		**0.369**	

Table 4.22 ***Item Discriminant Index Evaluation of Science 2019***

Discriminant Index (r_{pb})	No. of Items		Evaluation of Item	Decision
	AJK	ICT		
Negative	2	0	Worst/ Defective Item	Discard
<0.20	6	5	Marginal Item, not Discriminant	Revise/ Discard

0.20-0.29	7	13	Moderately Discriminant/ Fair item	Keep
0.30-0.39	8	12	Discriminant Item, Good item	Keep
≥0.40	9	2	Very Discriminating, very good Item	Keep

Table no.4.21 and table 4.22 demonstrate that question item no. 12 and 21 of Science test are worst items therefore these items should be discarded. Under the range of <0.20 6 items from AJK and 5 items from ICT are marginal item or not discriminant these should be revised or discarded from questionnaire. Subsequently, there are 7 and 13 items of AJK and ICT, respectively, lie under the range of 0.20-0.29, are moderate discriminant or fair items therefore these items should be maintained in questionnaire. Furthermore, 8 items from AJK list and 12 from ICT questionnaire list lie under the criteria of good item of discriminant 0.30-0.39. While, for ≥0.40 range of very good discriminant, 9 items from AJK and 2 items from ICT are very good discriminant or items.

Table 4.23 Comparison of NAT 2016 and 2019 of General Science for Grade 4 and 8 based on Discrimination Index

Aspect	**2019 Science Grade 4 (AJK & ICT)**	**2016 Science Grade 8 (AJK & ICT)**
Negative Discriminant Index	2	1
Marginal Item (rpb < 0.20)	AJK: 6, ICT: 5	AJK: 4, ICT: 4
Moderately Discriminant (0.20-0.29)	AJK: 8, ICT: 13	AJK: 6, ICT: 2
Discriminant Item (0.30-0.39)	AJK: 8, ICT: 13	AJK: 9, ICT: 14
Very Discriminant (≥0.40)	AJK: 9, ICT: 2	AJK: 23, ICT: 20

Table 4.23 compares the discrimination quality of science assessment items for Grade 4 students in 2019 and Grade 8 students in 2016 across AJK and ICT. The 2019 assessment showed a slightly higher number of items with poor or marginal discrimination, including 2 items with negative indices and more marginal items in both regions (AJK: 6, ICT: 5) compared to 2016 (AJK: 4, ICT: 4). Moderately discriminant items were notably higher in 2019 ICT (13) than in 2016 ICT (2), while discriminant items (rpb 0.30–0.39) were fairly consistent across both years. However, the 2016 Grade 8 assessment stood out with a significantly greater number of very discriminant items (AJK: 23, ICT: 20), indicating a higher proportion of high-quality items compared to 2019 (AJK: 9, ICT: 2).

Table 4.24 Comparison of NAT 2016 and 2019 of Mathematics for Grade 4 and Grade 8 based on the Discrimination Index

Aspect	2019 Mathematics Grade 8 (AJK & ICT)	2016 Mathematics Grade 4 (AJK & ICT)
Negative Discriminant Index	0	0
Marginal Item (rpb < 0.20)	AJK: 4, ICT: 3	AJK: 9, ICT: 0
Moderately Discriminant (0.20-0.29)	AJK: 11, ICT: 10	AJK: 12, ICT: 11
Discriminant Item (0.30-0.39)	AJK: 24, ICT: 20	AJK: 25, ICT: 22
Very Discriminant (≥0.40)	AJK: 11, ICT: 17	AJK: 4, ICT: 17

The table 4.24 compares the discriminant indices of the 2019 Mathematics Grade 8 assessments for AJK and ICT with those of the 2016 Mathematics Grade 4 assessments for the same regions. Notably, both the 2019 and 2016 assessments showed no negative discriminant index, indicating no severely flawed items. However, in terms of marginal items (rpb < 0.20), AJK had fewer marginal items in 2019 compared to 2016, whereas ICT maintained consistency. Moderately discriminant items (0.20-0.29) exhibited similar patterns between AJK and ICT across both years. In the category of discriminant items (0.30-0.39), AJK and ICT showed a slight decrease in 2019 compared to 2016. Interestingly, the number of very discriminant items (≥0.40) increased notably for ICT in 2019, indicating substantial improvement, while AJK maintained a relatively stable number. Overall, the data suggests variations in the quality and distribution of discriminant items across different categories between the two assessments years, with improvements observed particularly in ICT in 2019.

OBJECTIVE 2

To generate objective and accurate information about the application of NAT through teacher perceptions

Teacher's Perception about National Achievement Test (NAT) from 2019 and 2016 Survey

Table 4.25 ***Overall Teacher's Responses on NAT from ICT and AJK from survey of 2016***

Teacher Response	Mean of ICT	Standard Deviation	Mean of AJK	Standard Deviation
1) Do you think NAT improves				
Academic standards	4.18	0.98	4.66	0.50
Students' achievement	4.27	0.46	4.22	0.44
Teacher's effectiveness	3.90	0.94	4.22	0.44
Performance as accountability tool	3.90	0.94	3.44	0.72
Accuracy of data	3.45	1.12	3.55	0.52
Overall Mean	**3.94**		**4.02**	

2) Do you think that NAT monitor the progress of student by				
Measuring student achievements	3.90	0.70	4.33	0.5
Classroom assessment instruments	3.90	0.70	3.88	0.78
Determining Intended curriculum	4.27	0.46	4.33	0.70
Providing "snap-shot"	4	0.77	4.44	0.72
Helping Students in Effort in learning	4.09	0.70	4.44	0.52
Overall Mean	**4.03**		**4.28**	
3) Do you think that national achievement test impact the instructional methodology by				
Increasing teachers' assessment, knowledge, and skills	4	1	4.22	0.44
Improving and enhancing teaching	4.09	0.94	4.22	0.83
Providing Common measures	4.45	0.68	4.11	0.92
Evaluating the quality of students	4.54	0.52	4.33	0.70
Evaluating the quality of teachers	4.27	0.90	4.33	0.70
Evaluating the quality of schools	4.27	0.90	4.11	0.60
Overall Mean	**4.27**		**4.22**	
4) Being a teacher, do you think that National Achievement Test				
Opt the national assessment framework for assessment	4	0.77	3.66	0.86
Provide data for data-based decision-making at student level	4.09	0.83	4.22	0.66
Provide data for data-based decision-making at class level	4.09	0.94	4.22	0.44
Provide data for data-based decision-making at school level	4.27	0.46	4	0.5
Overall Mean	**4.11**		**4.02**	

5) Being a teacher, do you think that you/your				
Know how to interpret assessment results.	4.09	0.83	3.77	0.66
Know how to use the test results	3.90	0.70	3.88	0.78
Principals know how to interpret the test results	3.63	0.92	3.77	0.83
Public know how to interpret the results	3	1	3.22	0.83
Overall Mean	**3.65**		**3.66**	

The table 4.25 provided data presents the mean scores and standard deviations of teacher responses on a 5-point Likert scale related to the National Achievement Test (NAT) in Pakistan. The findings indicate areas of agreement and disagreement among teachers. Teachers widely agreed that NAT contributes to the improvement of academic standards and enhances students' achievements. They also acknowledged NAT's effectiveness in monitoring student progress and its positive impact on instructional methodologies. However, there were areas of disagreement, particularly concerning NAT's role as an accountability tool and teachers' level of confidence in interpreting and using test results. Additionally, teachers expressed mixed views on the understanding of principals and the public in interpreting test results. These findings highlight the importance of addressing the identified issues to enhance the effectiveness and utilization of NAT for educational improvement in Pakistan. Further research and targeted interventions may be required to address the concerns and utilize NAT data more effectively for driving positive educational outcomes.

Table 4.16: ***Overall Teacher's Responses on NAT from ICT and AJK from survey of 2019***

Teacher Responses	**Mean of ICT**	**Standard Deviation**	**Mean of AJK**	**Standard Deviation**
1) Do you think NAT improves				
Academic standards	3.5	1.5	3.0	1.2
Students' achievement	4.0	1.7	4.2	1.3
Teacher's effectiveness	4.2	1.4	4.0	1.5
Performance as Accountability tool	4.1	1.6	3.5	1.4
Accuracy of data				
	4.3	1.5	4.1	1.3
Overall Mean				
2) Do you think that NAT monitor the progress of student by	**4.02**		**3.76**	
Measuring student achievements				
Classroom assessment instruments	3.7	1.4	3.5	1.3
Determining Intended curriculum	3.9	1.5	3.8	1.4
Providing "snap-shot"	3.8	1.3	3.9	1.5
Helping Students in Effort in learning				
Overall Mean	3.6	1.6	3.4	1.3
3) Do you think that national achievement test impact the instructional methodology by	3.9	1.4	4.0	1.6
Increasing teachers' assessment, knowledge, and skills	**3.78**		**3.52**	
Improving and enhancing teaching				
Providing Common measures	4.0	1.4	3.7	1.3
Evaluating the quality of students	3.8	1.5	3.8	1.3
Evaluating the quality of teachers	3.6	1.4	3.9	1.5
Evaluating the quality of schools	3.5	1.5	3.8	1.3

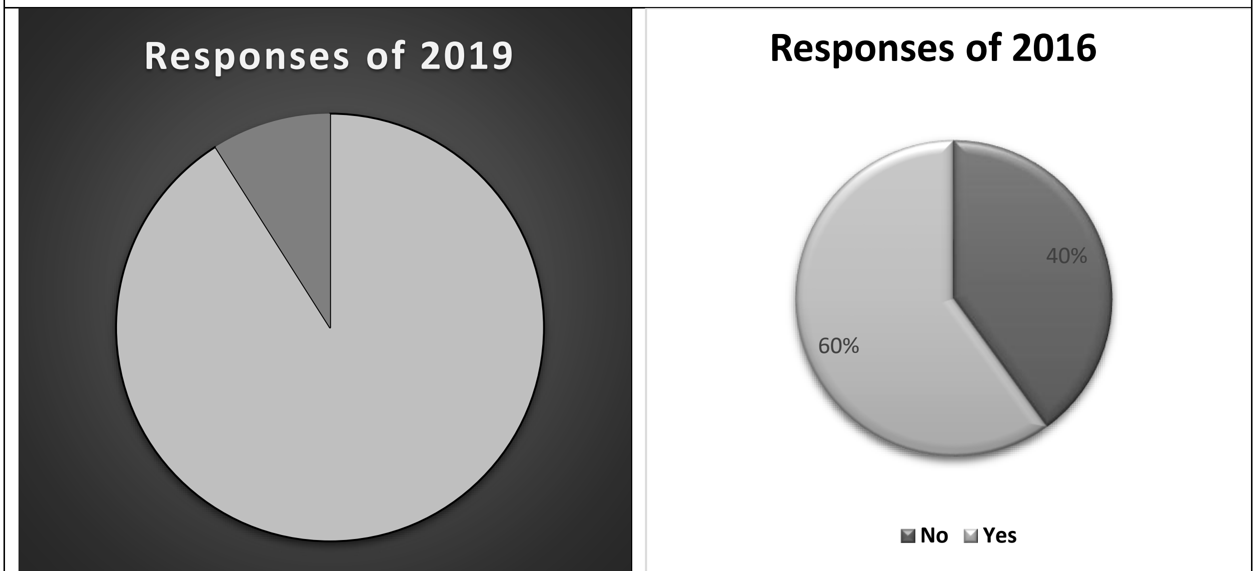
Overall Mean				
4) Being a teacher, do you think that National Achievement Test	3.7	1.3	3.6	1.4
	3.38		**3.78**	
Opt the national assessment framework for assessment	3.8	1.4	3.9	1.5
Provide data for data-based decision-making at class level	3.8	1.4	3.6	1.3
Provide data for data-based decision-making at school level	3.5	1.4	3.7	1.3
Overall Mean	**3.72**		**3.64**	
5) Teacher's understanding				
Know how to interpret assessment results	3.2	1.4	3.4	1.3
Know how to use the test results	2.1	1.5	3.0	1.4
Principals know how to interpret the test results	2.8	1.3	2.2	1.4
Public know how to interpret the results	0.9	1.3	1.2	1.4
Overall Mean	**2.5**		**2.45**	

The table 4.26 presents the perceptions of teachers in two regions, ICT and AJK, regarding the National Achievement Test (NAT) using a 5-point Likert scale. The mean values and standard deviations for each perception statement indicate the level of agreement or disagreement among teachers. In terms of NAT improvements, teachers generally agree that NAT contributes to enhanced academic standards in both regions, but there is higher agreement in AJK. Similarly, teachers perceive NAT positively impacting students' achievements, with AJK showing slightly higher agreement. However, opinions vary on NAT's role as an accountability tool, with higher agreement in ICT. Teachers in both regions express mixed views on the accuracy of data provided by NAT. Regarding NAT's monitoring of student progress, teachers believe it effectively measures student achievements, but the "snap-shot" feature receives lower agreement. In terms of NAT's impact on instructional methodology, teachers perceive it positively influencing teachers' assessment knowledge and skills in both regions. While improving and enhancing teaching has moderate agreement, providing common measures receives lower agreement. Teachers show mixed perceptions of NAT's impact on evaluating student, teacher, and school quality. Finally, teachers generally have a moderate understanding of interpreting assessment results, with slightly lower scores in AJK. Understanding the test results and interpretations by principals and the public are also moderately perceived. It is essential to consider these perceptions while exploring strategies to enhance the effectiveness and utilization of NAT in improving educational outcomes in both regions.

Objective 3: To assess the perceptions, experiences, and challenges faced by teachers in the utilization of National Achievement Test (NAT) results and identify effective strategies for applying NAT data to improve student achievement.

Figure 4.1 Usage of NAT Results in Institutes

Do you actually use the results of NAT being a teacher?



The survey included open-ended questions to gather additional insights, and the responses are represented in the pie charts provided above. Figure 4.1 presents the responses from 2019, indicating that only 9% of the respondents reported using the NAT results based on their experience, while a significant majority of 91% denied using them. On the other hand, the responses from the 2016 survey show that 60% of the respondents reported using the NAT results based on their experience, while 40% of the faculty denied doing so. This suggests that NAT results were more frequently utilized in 2016 compared to 2019.

Figure 4.2 NAT Value in 2016 and 2019

Do you feel the national achievement test have value?					
2016			2019		
Quality of the student	**Quality of the teacher**	**Quality of the School**	**Quality of the student**	**Quality of the teacher**	**Quality of the school**
70%	20%	10%	61.9%	23.8%	14.3%

Figure 4.2 illustrates the perceived value of the National Achievement Test (NAT) in evaluating the quality of students, teachers, and schools based on responses from 2016 and 2019. In 2019, 61.9% of respondents believed that NAT is valuable for evaluating student quality, followed by 23.8% for school quality, and 14.3% for teacher quality. In contrast, the 2016 responses show a stronger belief in the value of NAT, with 70% recognizing its value in assessing students, 20% for teachers, and 10% for schools. While overall agreement was higher in 2016, a small minority still expressed disagreement, with 6% disagreeing and 4% strongly disagreeing with NAT's value. This comparison suggests a decline in the perceived usefulness of NAT, particularly in evaluating teacher and school quality over time.

Table 4.27 *Teacher Perspectives on NAT as an Indicator of the Education System*

Theme	Responses of 2019	Responses of 2016
Assessment and Evaluation Assessing the Education System, Giving Effective Results, Helping Identify and Analyze, Achievement Level Variations	20%	15%
Accountability and Monitoring Good Indicator for Accountability of any Education System Serving as an Accountability Tool	40%	70%
Improvement and Weakness Identification Overcoming Weaknesses to Improve the Education System Identifying Weak Areas	30%	10%
Relevance and Impact Best Indicator for Analyzing the Education System, Essential for Attaining Quality Education, Contributes to Quality Education	10%	5%

This table 4.27 presents the perspectives of teachers regarding the National Achievement Test (NAT) as an indicator of the education system. The responses from 2019 and 2016 are categorized into four themes, highlighting the teachers' viewpoints on different aspects related to NAT. In

2019, teachers emphasized that NAT is a valuable tool for assessing the education system. They believe that NAT provides effective results (20%) and helps in identifying and analyzing variations in achievement levels among students (15%). Teachers also acknowledged the role of NAT in identifying weaknesses and areas for improvement within the education system (30%). Additionally, they recognized NAT as the best indicator for analyzing the education system and its contribution to achieving SDG 4 (10%).

In 2016, teachers expressed strong views regarding NAT as a good indicator for accountability and monitoring in the education system (70%). They believed that NAT played a crucial role in holding educational institutions and stakeholders accountable. Similarly, teachers recognized NAT as a tool for assessing the education system, providing effective results (15%) and assisting in identifying and analyzing variations in achievement levels (20%). They also acknowledged NAT's role in identifying weaknesses and areas for improvement (10%). Additionally, teachers highlighted its importance as the best indicator for analyzing the education system and its contribution to achieving SDG 4 (5%).

Figure 4.3 ***NAT alignment in Institutes***

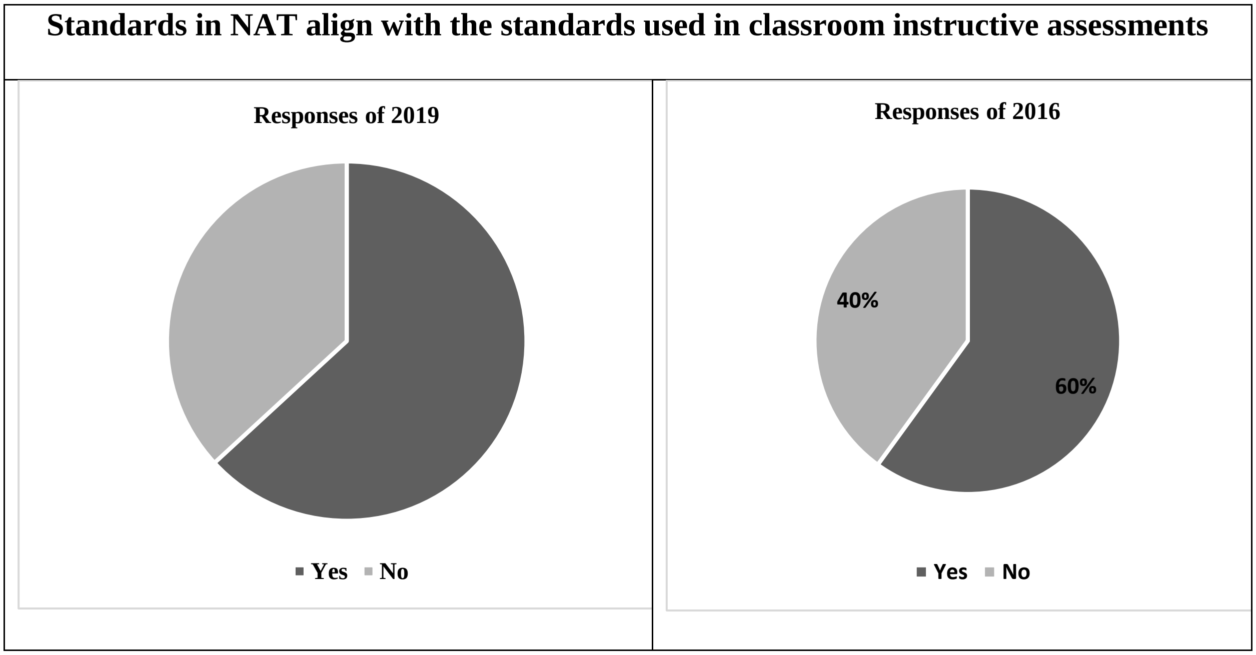


Figure 4.3 presents the responses of teachers from ICT and AJK in 2016 regarding the alignment of standards in the National Achievement Test (NAT) with the standards used in classroom instructive assessments. The findings indicate that 60% of teachers agreed that NAT aligns with the standards used for classroom instructive assessment. On the other hand, 40% of teachers held diverse opinions on this alignment. These responses provide insights into the perceptions of teachers from ICT and AJK regarding the consistency and compatibility of NAT standards with the standards implemented in classroom assessments. In 2019, the finding indicate 63% agreed and 37% held diverse opinions on this alignment.

Table 4.28: ***Teacher Perspectives on Factors Influencing Poor Student Achievement in NAT Assessment.***

Responses	**Responses 2019**	**Response 2016**
Training and Professional Development	20%	15%
Teaching and Learning Approaches	45%	50%
Socio-economic Factors	10%	13%
Curriculum and Resources	25%	22%

The table 4.28 highlighted that in 2019, teachers identified various factors that influenced poor student achievement in NAT. Theme 1 focused on training and professional development, with teachers highlighting the lack of training as a significant factor (20%). Theme 2, centered around teaching and learning approaches, encompassed concerns related to creativity and learning styles, lack of interests and guidance, and the presence of huge strength and weak concepts (45%). Theme 3 delved into socio-economic factors, specifically highlighting students' socio-economic background as a contributing factor to poor achievement (10%). Theme 4 centered on curriculum and resources, with teachers emphasizing the need for improvement in skills, full incorporation of the curriculum into textbooks, availability of audio/visual aids, and effective monitoring (25%).

In 2016, teacher’s shared similar perspectives on the factors influencing poor student achievement in NAT. Theme 1 emphasized the lack of training as a significant factor (15%). Theme 2, related to teaching and learning approaches, highlighted concerns such as the lack of creativity and learning styles, lack of interests and guidance, and the presence of huge strength and weak concepts (50%). Theme 3, focused on socio-economic factors, noted the influence of students' socio-economic background on achievement (13%). Theme 4, centered on curriculum and resources, pointed out the need for improvement in skills, full incorporation of the curriculum into textbooks, and effective monitoring (22%).

Overall, teachers recognized the importance of addressing factors such as training, teaching and learning approaches, socio-economic background, and curriculum and resources to improve student achievement in NAT. Their perspectives in both 2019 and 2016 emphasized the significance of these factors in shaping students' performance in NAT assessments

Figure 4.4 ***Satisfaction with NAT***

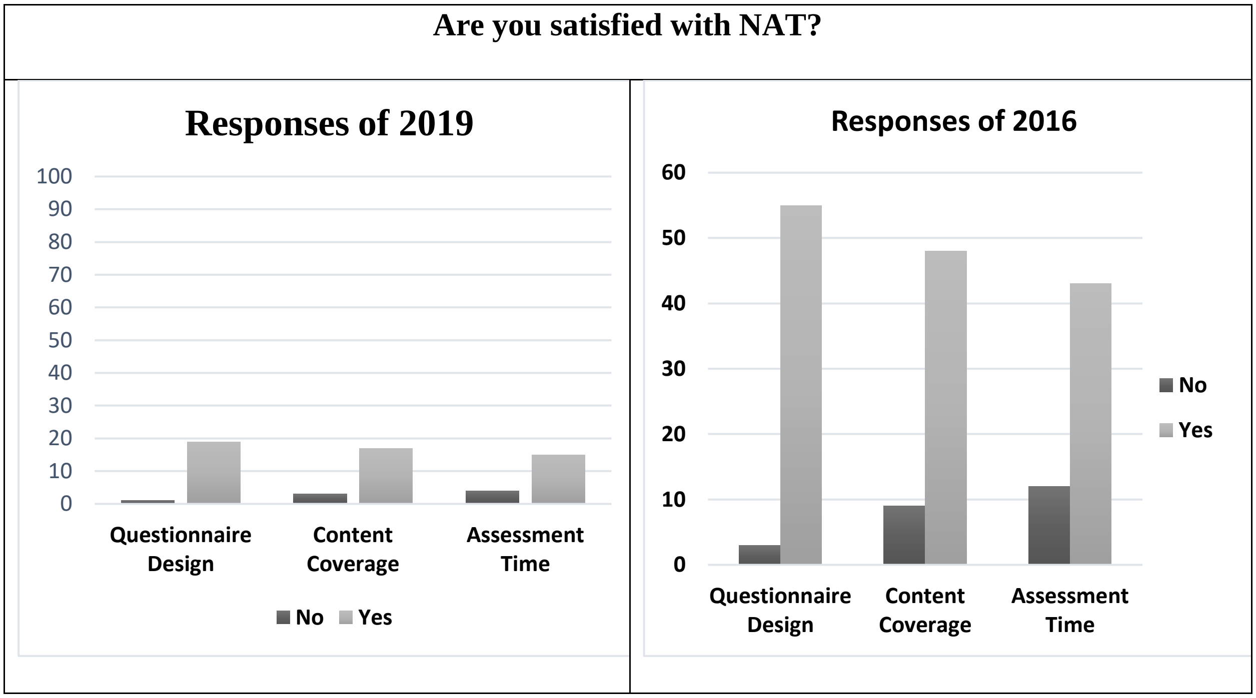


Table 4.4 presents the satisfaction levels of teachers with NAT, specifically focusing on questionnaire design, content coverage, assessment time, and overall results. The findings indicate that the majority of teachers in both 2019 and 2016 expressed satisfaction with NAT on these aspects. The average proportion of teachers who reported dissatisfaction with NAT was relatively low. These results suggest that teachers generally perceive NAT positively in terms of questionnaire design, content coverage, assessment time, and overall outcomes. The findings highlight a consistent trend of satisfaction among teachers regarding these aspects of NAT, as observed in both years of the study.

Figure 4.5 ***Data dissemination of NAT result***

Is there any mechanism for data dissemination of NAT result at your school level?

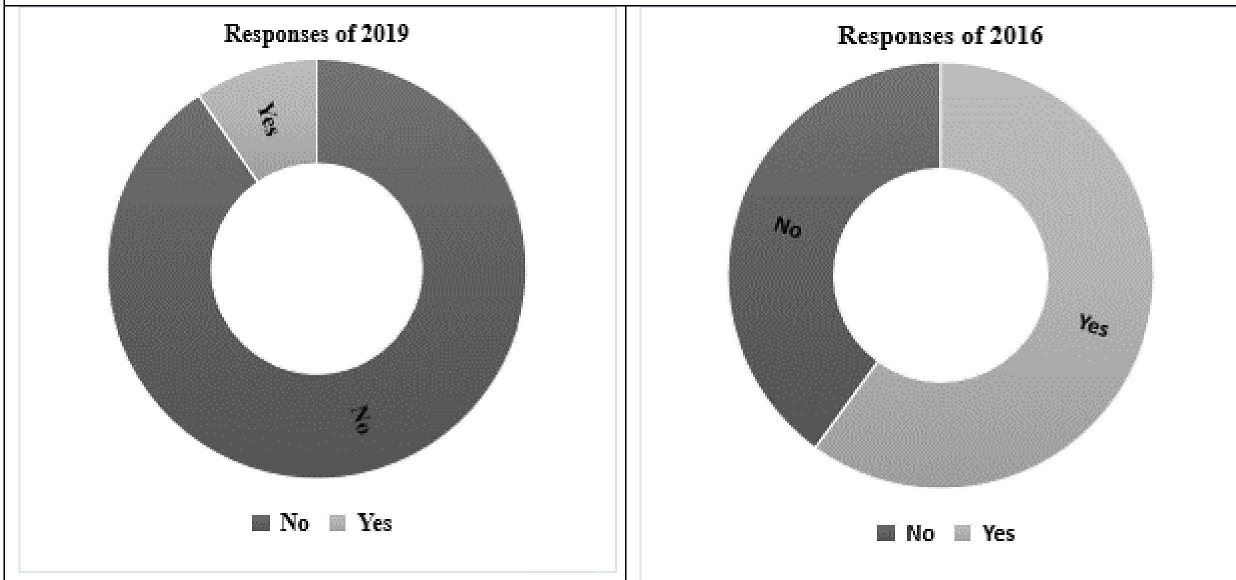


Figure 4.5 illustrates the school-level mechanisms for data dissemination of NAT results in 2019 and 2016. The findings reveal that in 2019, the majority of schools surveyed (91%) did not have any established mechanism for disseminating NAT results. Only a small percentage of schools (9%) reported having a mechanism in place for data dissemination. In contrast, the responses from 2016 indicate that approximately 40% of the surveyed schools lacked a data dissemination mechanism, while 60% of the schools reported having a mechanism in place. These results suggest a significant difference in the presence of data dissemination mechanisms between the two years, with a higher proportion of schools lacking such mechanisms in 2019.

Table 4.29 ***Challenges that teachers are facing in application of NAT in an institutes***

Response	**2019**	**2016**
Teacher-related challenges:	40%	25%
Institutional challenges:	25%	35%
Student-related challenges:	20%	25%
Other factors:	15%	15%

The table 4.29 illustrated that in 2019, teachers identified various challenges they face in the application of NAT. The theme of teacher-related challenges received a higher response rate of 40%. Teachers emphasized the lack of efficiency, preparation, language barriers, and the need for proper training. This highlights the importance of addressing these areas to improve the application of NAT. In the theme of institutional challenges, which received a response rate of 25%, teachers expressed concerns about the cooperation of management, limited time for application or preparation, low socio-economic backgrounds of students, lack of facilities and resources, and inadequate awareness about NAT. These challenges indicate the need for better institutional support and resources to enhance the application of NAT. The theme of student-related challenges received a response rate of 20%, with teachers highlighting the importance of training, conceptual learning, addressing unprepared students, and reducing the undue burden on students. This suggests the need for student-centered approaches and support systems. Finally, in the theme of other factors, 15% of teachers pointed out the challenges arising from rapid changes in textbooks and the burden of a large number of subjects on students.

In 2016, the challenges identified by teachers in the application of NAT were similar to those in 2019. The theme of teacher-related challenges received a response rate of 25%, indicating concerns such as lack of efficiency, preparation, language barriers, and the need for proper training. The theme of institutional challenges had a higher response rate of 35% in 2016, with teachers expressing concerns about the cooperation of management, limited time for application or preparation, low socio-economic backgrounds of students, lack of facilities and resources, and inadequate awareness about NAT. The theme of student-related challenges received a response rate of 25%, highlighting the importance of training, conceptual learning, addressing unprepared students, and reducing the undue burden on students. Similarly, in the theme of other factors, 15% of teachers mentioned the challenges posed by rapid changes in textbooks and the burden of a large number of subjects on students

Figure 4.6 ***Best ways to effectively applying NAT results as a teacher***

Table 4.30: ***Best ways to effectively applying NAT results as a teacher***

Theme-wise Response	2019	2016
Professional Development and Awareness -	60%	70%
Support and Resources	40%	25%
Language Barriers	---	5%

The table 4.30 presents the theme-wise responses of teachers in 2019 and 2016 regarding the best ways to effectively apply NAT results as a teacher. In the theme of Professional Development and Awareness, 60% of the teachers in 2019 and 70% in 2016 expressed the importance of continuous learning and improvement in their skills and knowledge. This indicates a consistent recognition of the need for professional development among teachers.

In the theme of Support and Resources, 40% of the teachers in 2019 and 25% in 2016 highlighted the significance of having adequate support and resources to effectively apply NAT results. This includes access to materials, technology, and a conducive working environment. The increase in the response rate from 2016 to 2019 suggests a growing awareness of the importance of resources in improving the application of NAT results. Regarding Language Barriers, there is no data available for 2019, but 5% of the teachers in 2016 identified language barriers as a factor to consider in effectively applying NAT results. This indicates a relatively lower emphasis on addressing language obstacles in that particular year.

Overall, the findings suggest that teachers recognize the importance of continuous professional development, the availability of resources, and addressing language barriers in effectively applying NAT results. The variations in response rates between 2019 and 2016 highlight the evolving priorities and potential shifts in the focus areas for teachers. These insights can inform strategies and interventions to support teachers in maximizing the application of NAT results in their teaching practices.

Figure 4.7*: Responses of teacher regarding the resources required for improving student performance*

In comparing the responses of teachers and school administrators in 2016 and 2019 regarding resources to enhance student performance in NAT, several key themes emerged. Notably, the importance of professional development and training significantly increased from 10% in 2016 to 30% in 2019, indicating a growing recognition of the need for skills enhancement. Concept-based learning remained consistent at 10% in both years. Effective monitoring, which garnered a substantial 30% in 2016, saw a slight decrease to 20% in 2019. Syllabus and guidelines remained a consistent focus with 20% in 2016 and 15% in 2019. The emphasis on resources and infrastructure shifted from libraries in 2016 (10%) to digital classrooms in 2019 (5%). Lastly, the significance of activity-based learning remained steady at 20% in both years. These variations reflect evolving educational priorities and contexts over time.

Objective 4: To propose a step by step approach for effectively disseminating NAT results to various stakeholders, facilitating their access to and understanding of the assessment outcomes.

Assessment activities aim to gather data on students' skills and knowledge. After developing and testing assessment tools and analyzing the data, the focus shifts to effective reporting and dissemination. Traditionally, large-scale assessments involve producing lengthy reports filled with graphs and tables, often in print, which tend to collect dust on education officials' shelves. This outdated approach fails to cater to stakeholders' needs, as many struggle to interpret complex numerical data, impeding its potential to enhance teaching and learning. There's a growing recognition of the necessity to design more user-friendly reporting and dissemination strategies that acknowledge the diverse requirements of stakeholders. These stakeholders, including government officials, policy makers, education administrators, teachers, parents, students, the general public, and the media, have varying preferences for data detail, presentation, and summaries. All should receive concise written summaries and data visualizations highlighting key findings, along with recommendations for improving education policies and practices.

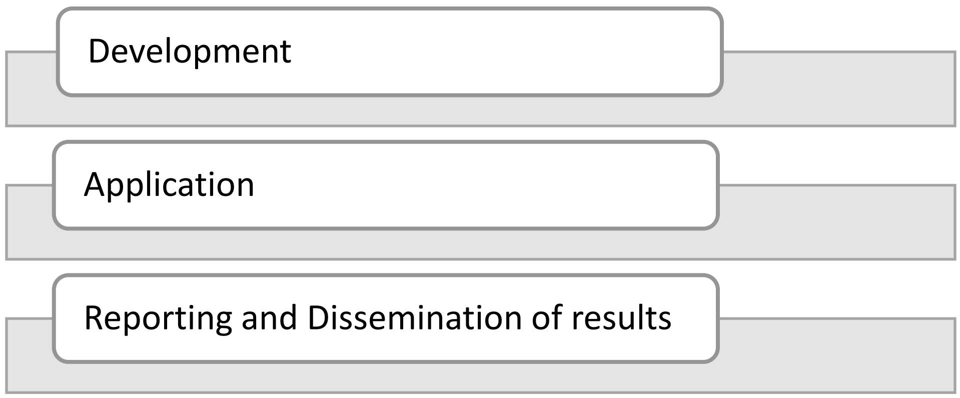


Fig 4.8 Current Model for Reporting and Dissemination of National Achievement Test

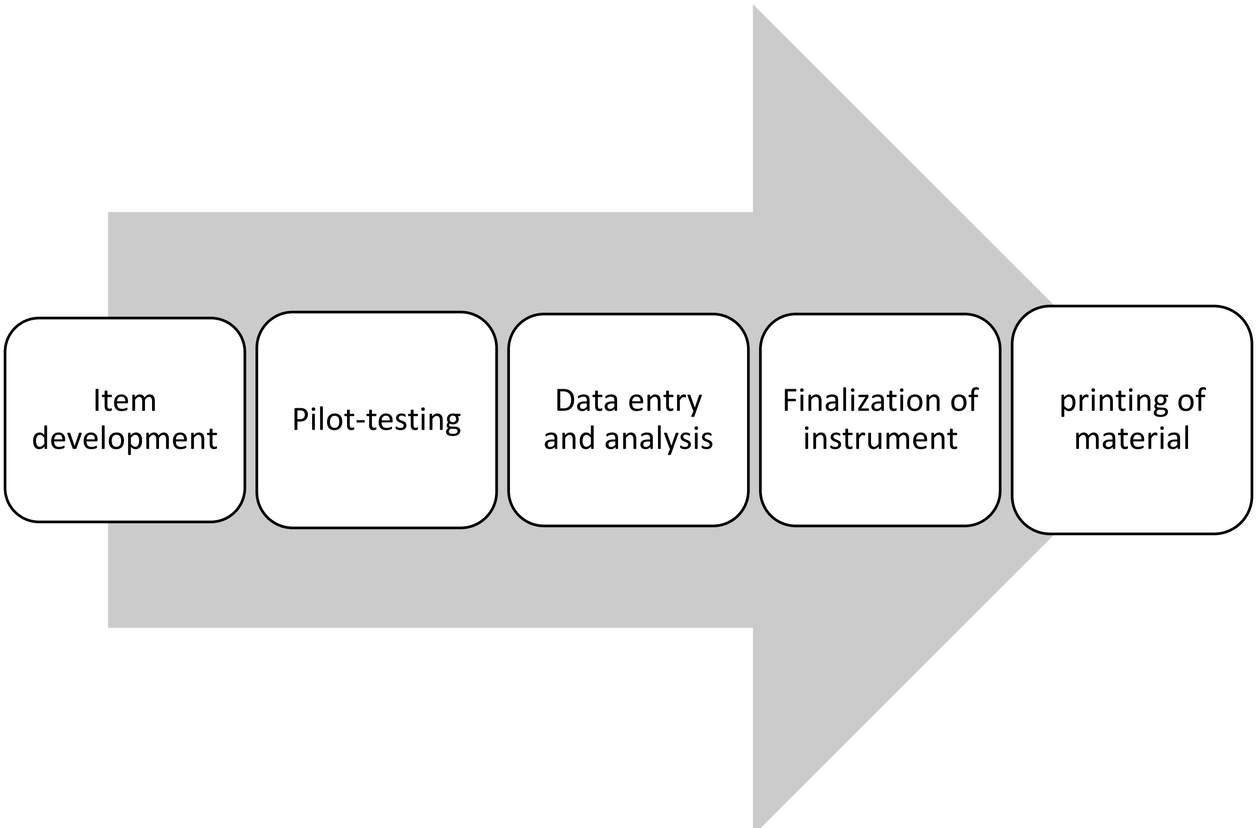


Fig 4.9 Current Practice for Development of Assessment Instrument

NAW follows a procedure for instrument development that starts with assessment framework development/review, followed by item writing by the subject experts, pilot testing for selecting more statistically fit test items, data analysis of the items, review by the experts, and final test development in each tested subject. On each stage, quality is ensured by thorough discussions in addition to meetings of the experts on review of the test items. The items for the subjects of English, Social Studies and Science for fourth grade and Urdu/ Sindhi and Mathematics for Eighth grade were written by the subject specialists and other experts in the field. These items were used to develop pilot testing booklets (NAT Report, 2016)

One of the most important steps in any assessment is test development. The assessment booklets contained the items, accepted after the data analysis of pilot test. In addition to the newly developed and piloted items, some anchor items from previous national assessment studies were also included in the assessment booklets. For each subject at least two parallel booklets namely version A and version B (Test booklet A and B) were developed. Before assembling the booklets, Tables of Specification (ToS) for each subject booklet was prepared. International practices of sequencing the items were used while assembling the booklets based on these ToS. After test development, the next important stage was printing of the test instruments. Keeping in view specialized nature of work, all printing jobs of the National Achievement Test (NEAS, 2016)

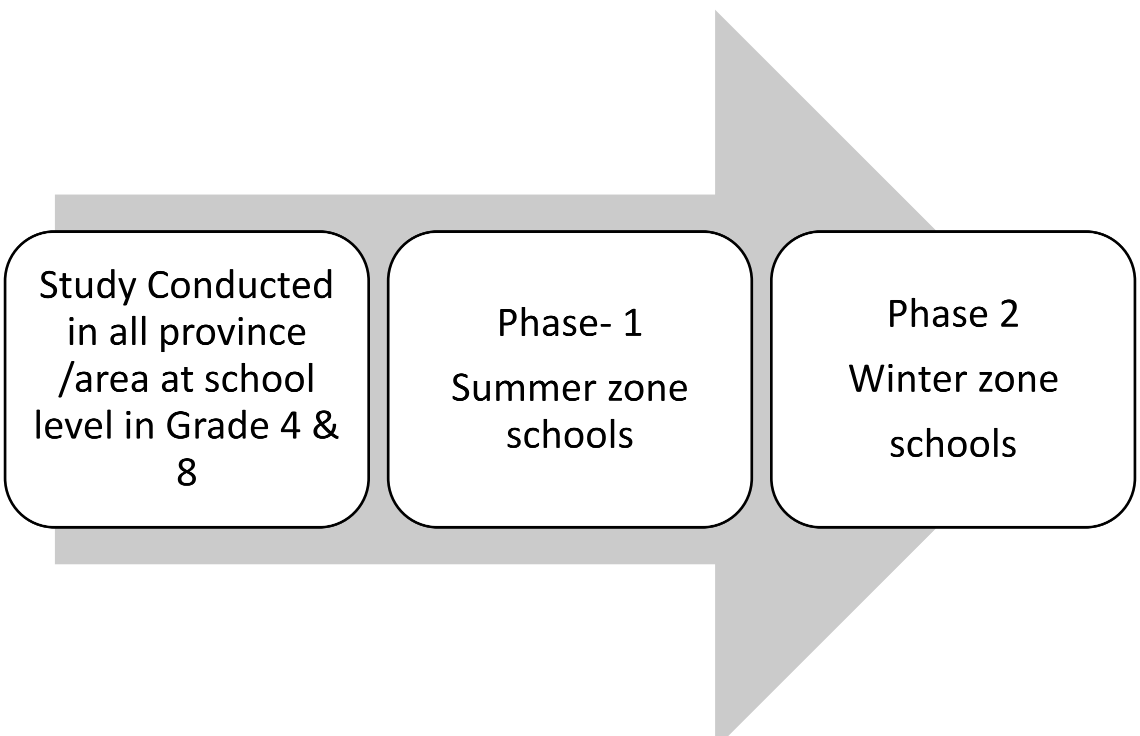


Fig 4.10 Current practice in PIE Application of National Achievement Test

After review and finalization of test items final test administration was scheduled according to this given plan across different area.

4.3 Reporting and dissemination of National achievement test

Assessment activities primarily gather data on students' skills and knowledge. After creating, testing, and analyzing assessment tools, the critical step is effective reporting and dissemination.

Traditional reporting in initiatives like PIE involves extensive reports filled with graphs and tables, often in print. Unfortunately, this outdated approach doesn't meet stakeholders' diverse needs. Many struggle with complex data, hindering the use of assessment data for improving education (Bhuta and Rizvi, 2022)

Thus, there's a growing recognition of the need for user-friendly reporting strategies tailored to different stakeholders. These include government officials, policy makers, education administrators, teachers, parents, students, the public, and the media. Each group has specific preferences for data detail, presentation, and summaries.

It's crucial that stakeholders receive written summaries and data visualizations highlighting key findings and suggesting improvements for educational policies and practices.

4.4 Analysis Based on Checklist:

Through the checklist, it was found that the present system of reporting and dissemination of national achievement the following gaps were found:

Table 4.31 Gaps found in National Achievement Test

Significant Gap	NAT 2016	NAT 2019
1.Targeted and a tailored approach (consideration for the specific needs of the audience or stakeholder)		
a. Report for teacher (that provide insight into student performance at group level.	somehow	somehow
b. Simplified reports for Parents of children overall performance	No	No
c. Community stakeholder access to aggregated data such as local organizations or NGOs may require access to aggregated data to identify educational gaps and prioritize intervention	somehow	No
2. Clear communication channels and mechanism for sharing data with relevant stakeholders		
a. Developing user-friendly online platforms	No	No
b. Organizing workshop for data usage	No	No
c. Training sessions to enhance data interpretation skill and creating data summaries and info graphics for wider dissemination	No	No

National Achievement Test (NAT) assessments conducted in 2016 and 2019, specifically related to the targeted approach, tailored communication, and data sharing with stakeholders. The gaps identified are as follows:

4.5 Targeted and Tailored Approach:

a. **Report for Teachers:** Both in NAT 2016 and NAT 2019, there is a gap in providing reports to teachers that offer insights into student performance at the group level. These reports would help teachers better understand their students' strengths and weaknesses, enabling them to make informed instructional decisions.

b. **Simplified Reports for Parents:** Similarly, in both assessments, there is a gap in providing simplified reports to parents regarding their children's overall performance. Such reports would help parents understand their child's progress and identify areas that may require additional support or attention.

c. **Community Stakeholder Access to Aggregated Data:** Another gap identified in both NAT 2016 and NAT 2019 is the lack of access for community stakeholders, such as local organizations or NGOs, to aggregated data. Providing these stakeholders with access to aggregated data would enable them to identify educational gaps and prioritize interventions in their respective communities.

4.5.1 Clear Communication Channels and Mechanisms for Data Sharing:

a. **Developing User-Friendly Online Platforms**: Both in NAT 2016 and NAT 2019, there is a need for the development of user-friendly online platforms that facilitate efficient data sharing. These platforms would provide stakeholders with easy access to assessment results, enabling them to interpret and analyze the data effectively.

b. **Organizing Workshops:** There is a gap in organizing workshops in both assessments, which could help stakeholders understand the assessment process, data interpretation, and utilization. Workshops would provide a platform for clarifying any doubts or queries and promoting effective communication and engagement among stakeholders.

c. **Training Sessions for Data Interpretation and Dissemination**: Both NAT 2016 and NAT 2019 lack training sessions aimed at enhancing stakeholders' skills in interpreting assessment data and creating data summaries and infographics for wider dissemination. Such training sessions would empower stakeholders to effectively analyze and present assessment results to drive evidence-based decision-making.

Addressing these gaps would lead to a more targeted and tailored approach, ensuring that the specific needs of different stakeholders are met. It would promote clear communication channels and mechanisms for data sharing, enabling stakeholders to access and utilize assessment results effectively. By bridging these gaps, the education system can enhance stakeholder engagement, promote evidence-based practices, and facilitate targeted interventions to improve student learning outcomes.

Phased Approach based Reporting and Dissemination for National Achievement Test

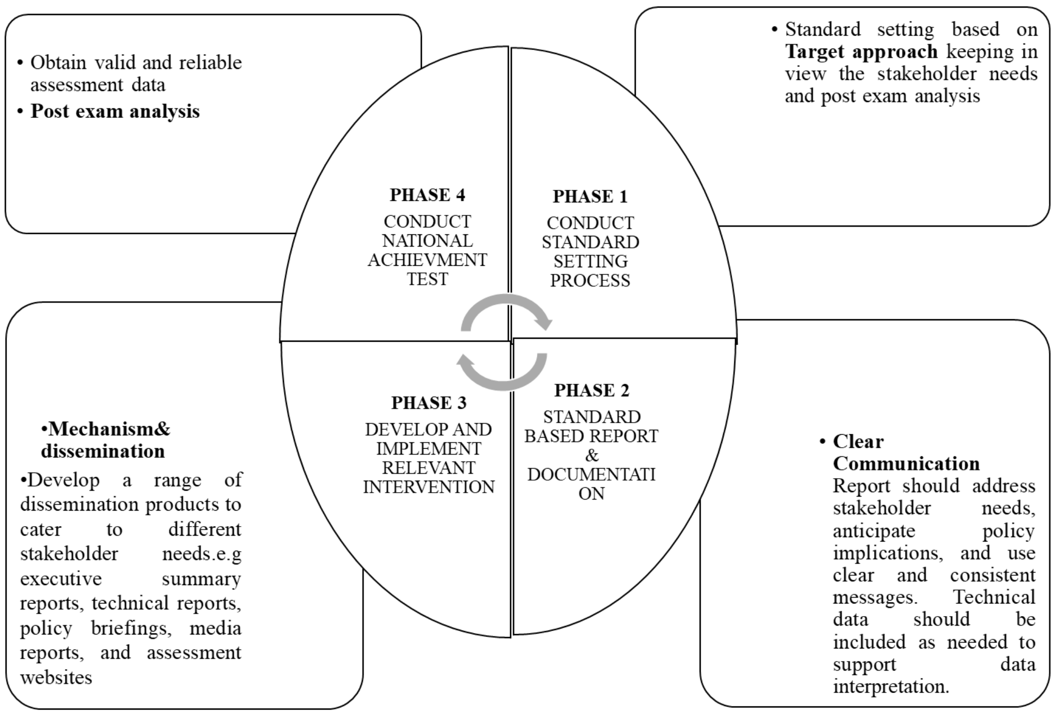


Fig 4.11 Phased Approach Based Reporting and Dissemination for National Achievement Test

The significant gaps identified in the proposed approach include the need for a targeted and tailored approach, clear communication channels, and mechanisms for sharing data with relevant stakeholders.

Targeted and tailored approach: This refers to the consideration of specific needs and requirements of different audiences or stakeholders. For example, a report designed for teachers should provide insights into student performance at the group level, enabling them to assess and address educational needs effectively. Simplified reports for parents should provide an overview of their children's overall performance, allowing them to understand and support their child's education. Community stakeholders, such as local organizations or NGOs, may require access to aggregated data to identify educational gaps and prioritize interventions.

Clear communication channels: To ensure effective sharing of data, it is essential to establish clear communication channels. This can include developing user-friendly online platforms where stakeholders can access relevant data. Additionally, organizing workshops and training sessions can help enhance stakeholders' data interpretation skills, enabling them to make informed decisions. Creating data summaries and infographics for wider dissemination can also aid in effectively communicating key findings and insights.

These gaps highlight the importance of tailoring information and providing accessible platforms for stakeholders to access and interpret data. By addressing these gaps, the proposed approach can facilitate better understanding, collaboration, and decision-making among stakeholders, ultimately leading to improved educational outcomes.

SUMMARY

The Teacher ‘perception regarding NAT 2016 and 2019 measured through the survey conducted in Islamabad (ICT) and Azad Jammu and Kashmir (AJK) collected overall 60 and 40 responses, respectively, from different schools of ICT and AJK and analyzed through the Likert scale and some open-ended question. In general, the findings depict that NAT plays an important role in building institutional ability and performance of both students and teachers. There is diversity in responses given by teachers in both surveys, however, the responses of 2019 show that the performance of NAT 2019 is not as much progressive as in 2016.

CHAPTER 5
SUMMARY, FINDING, CONCLUSION, DISCUSSION, AND RECOMMENDATION

5.1 Introduction

In the concluding chapter of this thesis, following a comprehensive data analysis, the focus shifts towards presenting the primary discoveries of this study, emphasizing deficiencies in the development and utilization of the National Achievement Test. Additionally, this chapter imparts valuable insights for deriving conclusions from the study's findings and extends recommendations aimed at guiding future researchers in their endeavors.

5.2 Summary

The main objectives of this research study were to evaluate the psychometric properties of the National Achievement Test (NAT) using the Classical Test Theory (CTT) framework, obtain objective information about NAT's application through teacher perceptions, assess teachers' experiences and challenges in utilizing NAT results, and propose an effective phased approach for disseminating NAT results to different stakeholders.

To achieve these objectives, mixed methodology approach was adopted, and an adapted questionnaire based on the conceptual model of Large Scale Assessment (LSA) was utilized. The study's first chapter provided background information, objectives, and an introduction to the methodology, while the second chapter explored literature related to large scale assessment, diverse perspectives on such assessments, and worldwide reporting and dissemination models for large scale assessment. The third chapter is based on methodology whereas the fourth chapter is based on study findings, conclusion and recommendations. The findings highlighted specific test items in Science and Mathematics that proved difficult for Grade 4 and Grade 8 students in AJK and ICT provinces, emphasizing the need for item revisions. Additionally, teachers' perceptions of NAT improved over time, with higher agreement on its positive impact in enhancing academic standards and monitoring student progress. Challenges in applying NAT and the best ways to utilize its results were also identified, emphasizing the importance of continuous professional development, availability of resources, and addressing language barriers. Furthermore, there was a gap in the reporting and dissemination of NAT results, necessitating targeted approaches and user-friendly platforms for effective data sharing with stakeholders.

5.3 FINDINGS:

5.3.1 Findings from the Objective 1: To evaluate the psychometric properties of National Achievement Test (NAT) and assess its effectiveness as an assessment tool by using the Classical Test Theory (CTT) Framework

1. The analysis of the difficulty index of General Science test items of 2016 for Grade 8 students from AJK and ICT provinces in table 4.1 reveals some significant findings. Among the test items, question numbers 11, 13, 27, 30, 39, were found to be difficult and should be revised or discarded. Additionally, question number 29, was also identified as difficult. These findings indicate the need for careful consideration and potential revisions in the General Science test items. Furthermore, it is worth noting that there are some differences between AJK and ICT in terms of difficulty index. While the majority of test

items fall within the moderately difficult/moderately easy range for both provinces, question number 39 stands out as particularly challenging for both AJK and ICT students. These findings highlight the importance of ensuring fairness and appropriateness in the assessment process, emphasizing the need for targeted improvements in specific test items to enhance the accuracy and reliability of the evaluation.

2. From the Evaluation of Discarded item domain wise from Difficulty Index of General Sciences Test Grade 8, 2016 in table 4.3. The difficulty index was found for AJK students in item no. 11, 13, 27, 30, and 39 are belong to "Physical Sciences" domain for AJK students of General Sciences. Whereas for the ICT students, only 1 question no. 39 of test for ICT students is belongs to "Life Sciences" domain. However, these questions from two different domains are difficult questions (lies under the criteria <0.30), therefore these questions of general science test should be reconsidered to revised or discarded from test.
3. The analysis of the difficulty index for Mathematics test items of 2016 for Grade 4 of students from AJK and ICT provinces in 2016 in table 4.7 and 4.8 reveals some significant findings. Among the test items, question numbers 2, 13, 14, 19, 24, 25, 28, 31, 33, 34, 36, 38, 40, 41, 43, 44, 45, 46, 47, and 49 were identified as difficult for AJK students and should be revised or discarded. On the other hand, for ICT students, question numbers 2, 14, 18, 25, 31, 33, 41, 44, 45, 49, and 50 were found to be challenging and require reconsideration. A closer examination of the findings reveals that 20 questions in the Mathematics test for AJK students and 11 questions for ICT students were considered difficult. These questions should be revised or excluded from the questionnaire to ensure a

fair assessment. However, it is important to note that the majority of the test items for both provinces fall within the moderately difficult/moderately easy range, suggesting an appropriate level of difficulty. When comparing the findings between AJK and ICT provinces, it can be observed that there are some variations in the specific questions identified as difficult. For example, question numbers 13, 18, and 50 were found to be challenging for ICT students but not for AJK students. Conversely, question numbers 19, 24, 36, 38, 40, 43, 46, and 47 were identified as difficult for AJK students but not for ICT students. These differences highlight the variations in the level of difficulty experienced by students from different regions, possibly indicating differences in curriculum implementation, teaching methodologies, or student preparation.

4. The evaluation of discarded items domain wise from the Grade 4 Mathematics Test in table 4.9 shows that Geometry & Measurement and Number & Operation were the most challenging domains. Several items were difficult for both AJK and ICT students, while some posed region-specific challenges, such as items 19, 36, and 38 in AJK and item 50 in ICT for Geometry & Measurement. Similarly, AJK struggled with items 13, 24, 43, and 46, while ICT found item 18 difficult in Number & Operation. In Information Handling**,** items 28, 34, and 40 were problematic in AJK, while item 33 was difficult for both regions. These findings suggest the need for targeted instructional support to address regional learning gaps.
5. The analysis of the Grade 8 Science test items for AJK and ICT in 2016, as presented in Table 4.17 and 4.18, revealed significant findings regarding the discriminant index of each

item. Item no. 21 was identified as the worst item and should be discarded from the questionnaire due to its poor performance. Additionally, four items from both AJK and ICT fell below the <0.20 range, indicating their marginality and lack of discriminant power, suggesting the need for revision or exclusion. Conversely, six items from AJK and three items from ICT were found to be moderately discriminant or fair within the 0.20-0.29 range, warranting their retention. Moreover, nine items from AJK and 14 items from ICT were identified as good discriminant items within the 0.30-0.39 range. Notably, a significant number of items, 23 from AJK and 20 from ICT, exceeded the criteria of ≥0.40, signifying their high discriminant power. These findings underscore the importance of carefully selecting and retaining items that effectively discriminate student performance in the Science test for Grade 8.

6. The item discriminant index evaluation for the Grade 4 Math's test 2016 in AJK and ICT in table 4.19 and table 4.20 revealed that no defective (negative discrimination) items were found in either AJK or ICT. However, 9 items in AJK had a poor discrimination index (<0.20), requiring revision or removal, while ICT had none in this category. Moderately discriminant (fair) items (0.20–0.29) were identified in both regions, with 12 in AJK and 11 in ICT. A significant number of items were classified as good discriminators (0.30–0.39), including 25 in AJK and 22 in ICT. Notably, ICT had more very good discriminating items (≥0.40), with 17 items, compared to just 4 items in AJK. Overall, the majority of test items in both regions were effective, but AJK had more items requiring revision or removal, indicating a need for improvement in test item quality.

7. The analysis of the difficulty index for Science test items of 2019 Grade 4 students from AJK and ICT provinces in table 4.10 reveals some notable findings. Among the test items, question numbers 17, 18, 19, 21, 23, 28, 30, and 31 were identified as the most difficult for AJK students and should be revised or discarded. Similarly, for ICT students, question numbers 12, 18, 21, 25, 26, 28, 29, 30, 31, and 32 were found to be challenging and require reconsideration. These findings suggest the need for improvement and potential revisions in these specific test items to ensure a fair and accurate assessment of students' knowledge in Science. Additionally, it is important to note that the majority of test items fall within the moderately difficult/moderately easy range for both provinces, indicating an appropriate level of difficulty.

8. From the Evaluation of Discarded item domain wise from Difficulty Index of science Test Grade 4, 2019 in table 4.12. It is found that Sciences test for AJK students**.** The question no. 17,18,23, 30 and 31 are belong to “Physical Sciences” domain, 19, 21 are from “Life Sciences” domains for AJK students of General Sciences. Whereas, in Science test for ICT students Consequently, item no. 18, 21, 25, 26, 32 of test for ICT students is belongs to “Life Sciences” domain, item no. 29,30,31, and 33 are from “Physical Sciences” domain. However, these questions from two different domains are difficult questions (lies under the criteria <0.30), therefore these questions of general science test should be reconsidered to revised or discarded from test.

9. The analysis of the difficulty index for 2019 for Grade 8 Mathematics test items in AJK and ICT provinces in 2019 reveals some important findings. Table 4.4 shows that test items 2, 13, 14, 19, 24, 25, 28, and 31 were identified as difficult, falling within the criteria of the difficulty index range. Further evaluation in Table 4.5 indicates that these difficult items need to be revised or discarded from the questionnaire. Specifically, for AJK students, 8 items were considered difficult, while for ICT students, 6 items required revision or exclusion. Upon comparing the findings between AJK and ICT, it can be observed that there are some differences in the specific items identified as difficult. Item numbers 13 and 24 were found to be challenging for AJK students, while item numbers 14, 19, and 31 were considered difficult for both AJK and ICT students. Additionally, item numbers 2, 25, and 28 were identified as difficult for ICT students but not for AJK students. These variations suggest discrepancies in the level of difficulty experienced by students from different provinces, highlighting potential differences in curriculum emphasis, instructional methods, or student preparedness.
10. From the Evaluation of Discarded item domain wise from Difficulty Index of Mathematics Test Grade 8, 2019 Mathematics test in table 4.6. It is found that the discard items are from different mathematical domains. Consequently, discarded items from Mathematics test of AJK students including item no. 2, 31 are from "Geometry & Measurement" domain, 13 from "Numbers & Operation" domain, 14, 19, and 24 are from Algebra, and item no. 25, 28 are from "Information Handling" domain. Whereas item no. 2, 31 from "Geometry & Measurement" domain, 14, 18 from "Algebra" domain, and item no. 25, 28 from

"Information Handling" domain. Therefore, 8 items for students of AJK and 6 items from mathematics test of ICT are from different domains and these questions lie under the range of (<0.30) difficulty index.

11. The analysis of the Grade 8 Mathematics test items in AJK and ICT provinces in 2019 revealed several significant findings. Table 4.15 and 4.16 provided insights into the discriminant index of each item, indicating their effectiveness in differentiating student performance. Two items, specifically item no. 2 and 19, were found to be marginal and non-discriminant, suggesting the need for revision or exclusion from the question list. Within the range of 0.20-0.29, 4 items from AJK and 5 items from ICT were moderately discriminant or fair, warranting consideration. Moreover, 17 items for AJK and 8 items for ICT fell within the 0.30-0.39 range, indicating their effectiveness as good discriminant items. Lastly, 8 items for AJK and 18 items for ICT exceeded the criteria of ≥0.40, demonstrating their high discriminant power and suggesting their retention in the questionnaire. These findings highlight the importance of selecting items that effectively discriminate student performance and contribute to the validity and reliability of the mathematics test.
12. The analysis of the Grade 4 Sciences test items for AJK and ICT in 2019, as presented in Table 4.21 and 4.22 revealed important findings regarding the discriminant index of each item. Item no. 12 and item no. 21 were identified as the worst items and should be discarded from the questionnaire due to their poor performance. Additionally, 6 items from AJK and 5 items from ICT fell below the <0.20 range, indicating their marginality and lack of

discriminant power, suggesting the need for revision or exclusion. Conversely, eight items from AJK and thirteen items from ICT were found to be moderately discriminant or fair within the 0.20-0.29 range, warranting their retention. Moreover, eight items from AJK and thirteen items from ICT were identified as good discriminant items within the 0.30-≥0.40, signifying their high discriminant power. These findings emphasize the significance of carefully evaluating and selecting items that effectively discriminate student performance in the Grade 4 Sciences test for both AJK and ICT.

13. The findings based on table 4.13 and 4.14 suggest a notable increase in difficulty levels across Mathematics Test Grade 4 and General Sciences Test from 2016 to 2019, particularly evident in AJK and ICT regions. There's a discernible trend towards a more balanced distribution of difficulty levels in the later assessments, with fewer difficult items and enhanced variability in question difficulty observed. Furthermore, table 4.23 and 4.24 showed that both regions exhibited deficiencies in item quality in 2019, with negative discriminant indices indicating areas for improvement, ICT showed notable improvements in Mathematics assessments, emphasizing the need for ongoing refinement in test construction strategies to ensure consistent and reliable assessment standards.

5.3.2 Findings from the Second Objective: To obtain objective and accurate information about the application of the National Achievement Test (NAT) by considering the four functions of Large Scale Assessment (LSA), namely accountability, gatekeeping, instructional identification, and monitoring of student progress through teacher's perception.

1. The comparison between teachers' perceptions of the National Achievement Test (NAT) of 2016 and 2019, as shown in Tables 4.25 and 4.26, reveals interesting trends. In 2016, teachers from both regions, ICT and AJK, generally showed higher mean scores compared to 2019, indicating a more positive outlook towards NAT in the later year. Teachers in 2016 demonstrated higher agreement on the positive impact of NAT in terms of enhancing academic standards and students' achievements, suggesting a growing recognition of NAT's role in improving educational outcomes. The perception of NAT as an effective tool for monitoring student progress also received higher agreement in 2016 indicating that teachers acknowledged the value of NAT data in understanding and tracking student growth.
2. The comparison between teachers' perceptions of the National Achievement Test (NAT) in 2016 and 2019 reveals interesting trends. In 2016, teachers from both regions, ICT and AJK, generally showed a more positive outlook towards NAT compared to 2019. They demonstrated higher agreement on the positive impact of NAT in terms of enhancing academic standards and students' achievements, suggesting a growing recognition of NAT's role in improving educational outcomes.
3. Regarding the understanding of teachers, principals and the public in interpreting test results, teachers had mixed views in both years. In 2016, teachers expressed slightly higher mean scores in this aspect compared to 2019, but the overall mean scores remained relatively low, indicating ongoing concerns. The Figure 4.1 also support the finding by identifying less percentage of teacher are using NAT results in their school.

5.3.3 Findings of the Third Objective: To assess the perceptions, experiences, and challenges faced by teachers in the utilization of National Achievement Test (NAT) results and identify effective strategies for applying NAT data to improve student achievement

1. The perspectives of teachers regarding the National Achievement Test (NAT) as an indicator of the education system were also explored in table 4.27. In 2019, 20% of teachers viewed NAT as providing effective results, while 15% emphasized its role in identifying and analyzing variations in student achievement levels. Additionally, 30% of teachers recognized NAT as a means of identifying weaknesses and areas for improvement within the education system. Furthermore, 10% of teachers believed that NAT was the best indicator for analyzing the education systems. In 2016, the majority of teachers (70%) expressed the view that NAT served as a good indicator for accountability and monitoring in the education system. Similar to 2019, teachers in 2016 recognized NAT's effectiveness in providing results (15%) and identifying variations in achievement levels (20%). They also acknowledged its role in identifying weaknesses and areas for improvement (10%), as well as its importance in analyzing the overall education system's (5%). Overall, the findings demonstrate that teachers recognize NAT as an important tool for evaluating and improving the education system, with slight variations in the percentage of teachers expressing certain perspectives between 2019 and 2016.

2. The major findings from the teacher responses in 2019 and 2016 regarding factors influencing poor student achievement in NAT reveal some important insights. In 2019, teachers emphasized the lack of training (20%) as a significant factor affecting student performance. They also expressed concerns about teaching and learning approaches, such as creativity, learning styles, interests, and weak concepts (45%). Socio-economic factors, particularly students' background, were identified as contributing to poor achievement (10%). Teachers also highlighted the need for curriculum improvement, full incorporation of the curriculum into textbooks, availability of resources like audio/visual aids, and effective monitoring (25%). Similarly, in 2016, teachers shared similar perspectives on these factors. The lack of training (15%) was identified as a key issue, along with concerns about teaching and learning approaches (50%) and socio-economic factors (13%). The need for curriculum improvement and effective monitoring (22%) was also emphasized. Overall, the major findings indicate the consistent recognition of the importance of training, teaching and learning approaches, socio-economic background, and curriculum and resources in influencing student achievement in NAT. While there are some variations in the response rates between 2019 and 2016, the identified factors remain crucial in addressing the challenges faced in improving student performance in NAT assessments.
3. The responses from teachers in 2019 and 2016 in table 4.29 regarding the challenges in applying NAT revealed slight variations in the percentage of certain

themes. In 2019, the theme of teacher-related challenges received a higher response rate of 40% compared to 25% in 2016. Similarly, the theme of institutional challenges had a lower response rate of 25% in 2019 compared to 35% in 2016. However, the themes of student-related challenges and other factors showed consistent response rates of 20% and 15% respectively in both years. These variations in response rates highlight the evolving priorities and potential shifts in focus areas for teachers, indicating the need for continued attention to these challenges in improving the application of NAT in education.

4. The major finding from the responses of teachers regarding the best ways to effectively apply NAT results in table 4.30 is that both in 2019 and 2016, there is a strong recognition among teachers of the importance of continuous professional development and awareness. This includes the need for ongoing learning and improvement in their skills and knowledge. Additionally, there is an emphasis on the significance of support and resources in effectively applying NAT results, with an increased response rate in 2019 compared to 2016. Teachers highlight the importance of having adequate resources, such as materials, technology, and a conducive working environment, to enhance the application of NAT results. However, in 2016, there is a lower emphasis on addressing language barriers as a factor in effectively applying NAT results, as only 5% of the teachers identified it. The variations in response rates between 2019 and 2016 suggest evolving priorities and potential shifts in focus areas for teachers. These findings underscore the

importance of continuous professional development, availability of resources, and the need to address language barriers in maximizing the application of NAT results.

5. The findings from the teacher responses indicate that certain resources are crucial for enhancing student performance in NAT assessments. These resources encompass professional development and training for teachers, concept-based learning approaches, effective monitoring and evaluation practices, well-defined syllabi and guidelines, as well as resources and infrastructure such as digitalized classrooms and libraries, alongside the incorporation of activity-based learning methods. Furthermore, when comparing the responses from 2016 and 2019, notable differences emerge. In 2016, there was a greater emphasis on the need for improvement in critical thinking and learning styles, guidance on utilizing NAT results effectively, and a proper understanding of the curriculum. Conversely, in 2019, there was an increased focus on teacher awareness programs regarding the use of NAT, the digitalization of classrooms, and addressing language barriers.

5.2.4 Finding of Fourth Objective: To propose a step by step approach for effectively disseminating NAT results to various stakeholders, facilitating their access to and understanding of the assessment outcomes.

The study's fourth objective aimed to propose a step by step approach for effectively disseminating NAT results to various stakeholders, facilitating their access to and understanding of the assessment outcomes. Figure 4.5 revealed a significant gap in the process, emphasizing the importance of a targeted approach in presenting assessment data to different stakeholders. Notably, the lack of specific reports for teachers hindered their ability to access group-level student performance insights crucial for informed instructional decisions. The study highlighted the necessity for clear communication channels and mechanisms for data sharing, stressing the need for user-friendly online platforms and training sessions to enhance stakeholders' skills in data interpretation and dissemination. In 2019, the majority of surveyed schools (91%) lacked an established mechanism for disseminating NAT results, compared to only 9% reporting such mechanisms. Conversely, in 2016, approximately 40% of schools lacked such mechanisms, while 60% had them in place, indicating a significant difference between the two years, with a higher proportion of schools lacking mechanisms in 2019.

5.3 DISCUSSION

Research by the Global Partnership for Education (2018) documented that the quality of assessment systems is a key component of strong education systems and serves as a vital means to monitor and improve learning outcomes. Hence, the objective of this study was to evaluate the National Achievement Test (NAT) and assess its effectiveness as an assessment tool using the Classical Test Theory (CTT) framework. The comprehensive psychometric analysis conducted for the years 2016 and 2019 which provided valuable insights into the difficulty indices and discriminating power of the test items. The findings from the analysis of item difficulty revealed specific test items that were challenging for students in general science test, especially in the "Physical Sciences" and "Life Sciences" domains. These difficult items were identified in both years, indicating the need for revisions or exclusions from the assessment to ensure a fair and accurate evaluation of student knowledge. Numerous studies was found that particularly focused on the factor that contributed to the poor performance of student in science. Another finding of this study from mathematics achievement test, was found that notably, certain questions from diverse mathematical domains, such as "Geometry and Measurement," "Numbers and Operations," "Algebra," and "Information Handling," were found challenging for student. Scholars suggest such as Tshabalala and Ncube (2016) have collectively highlighted the multifaceted nature of factors influencing mathematics performance, emphasizing that poor performance is a result of a complex interplay among student, teacher, and school-related variables. Within the realm of student factors, researchers such as Mohamed and Waheed (2011), and Ngussa and Mbuti (2017) consistently underline the pivotal role of attitude in shaping mathematical achievement. Attitude, in this context, denotes a learned inclination of individuals to react positively or negatively to objects, situations, concepts, or other individuals (Sarmah & Puri, 2014). It's important to note that attitudes are not static; they can evolve and transform over time (Syyeda, 2016). Once a positive attitude towards mathematics takes root, it can significantly enhance students' learning experiences (Akinsola & Olowojaiye, 2008; Mutai, 2011). Conversely, a negative attitude can act as a substantial impediment to effective learning, ultimately impacting academic performance (Joseph, 2013). Despite extensive research on this subject, there appears to be a dearth of specific domain wise comparison in mathematics linked with assessments conducted by researchers, warranting further investigation in this direction.

Furthermore, regional variations in item difficulty between students from AJK and ICT provinces was found in this study, which highlighted potential disparities in curriculum implementation and teaching methodologies.. The research conducted by Durry (1999) underscores a direct correlation between the curriculum and student achievement. It is widely believed that well-defined educational goals and objectives exert a positive influence on students' performance. In a separate investigation, Khan and Saeed (2010) revealed that an effective educational system necessitates a curriculum that equips children with the knowledge, skills, values, and attitudes required to thrive as successful citizens in today's world, characterized by communication and information technology. A parallel study by Cogan, Schmidt, and Wiley (2001) pinpointed the curriculum as the primary factor contributing to differential achievement among eighth-grade students in mathematics in the United States. Furthermore, Leighton (2004) ascertained that performance is intricately linked to the content coverage and the expected level of mastery within a specific domain, as well as the methods employed in teaching.

Teachers' perspectives on the National Achievement Test (NAT) reveal its significance in improving the education system, with it also highlight concerns regarding inadequate training and the necessity for better teaching and learning approaches. Young, McNamara, Brown, and O’Hara (2018) assert the importance of professional development for teachers in effectively utilizing data—collecting, analyzing, and strategically using it. Additionally, various research studies identify factors influencing poor student achievement in NAT, including lack of training, teaching and learning approaches, socio-economic background, and curriculum resources. Factors such as infrastructure, material availability, and efficient time management, both within and beyond the classroom, significantly influence students' learning outcomes. Cowie (2012) emphasizes that enhancing resource allocation must be coupled with improved resource utilization to better learning outcomes. Addressing these factors is crucial for improving student performance and promoting equity in education. Furthermore, the study explores the alignment of NAT with classroom instructive assessment standards, stressing the need for better integration with classroom practices. Swan and Burkhardt (2012) conclude that large-scale assessments influence classroom instruction and assessment practices. Teacher responses on NAT challenges underscore the prominence of teacher-related challenges over institutional ones. Continuous professional development is deemed critical, along with ample support and resources to effectively apply NAT results. Collins (2009) observes the recurring issue of limited teacher training in assessment, while Stiggins (2005) notes teachers' insufficient familiarity with fundamental testing and measurement concepts. Despite exposure to measurement guidelines during their courses, educators often fail to effectively apply and adhere to them.

It is crucial to recognize that assessment practices wield a significant influence on the overall quality of teaching and learning. Assessment data encompass diverse forms of information related to student achievement, spanning various assessment types like written and oral exams, standardized tests, portfolios, and report cards (Schildkamp & Lai, 2013). These data sources offer valuable insights into students' progress in relation to various curriculum objectives or standards, providing teachers with essential information that can inform and shape their instructional methods—a practice commonly referred to as formative assessment.

This study, keeping in view the gaps found in reporting and dissemination of NAT, proposed a comprehensive model for disseminating NAT results to stakeholders and highlighted the importance of a targeted and tailored approach. Research suggest that National assessment data can be released to and utilized by both internal and external stakeholders (DepED Order, 2009). “In numerous instances, the dissemination of information to the public and other stakeholders has been found lacking. The deficiency in many national assessment programs, as highlighted by Maligalig and Albert (2008), is evident in inadequate dissemination or when the provided data does not align with the actual needs of the stakeholders. Furthermore, insufficient efforts in sharing information with the public, coupled with misguided public expectations regarding assessment programs and their subsequent reforms, have led to the improper utilization of data in policymaking (Gutiérrez & Vázquez, 2008). This has sometimes discouraged the adoption of effective, long-term policies in favor of pursuing quick results (Meckes & Carrasco, 2010). In 1994, The Widmeyer Group assessed NAEP dissemination and found issues: data seen as complex, reports overwhelming, and a need for user-friendly elements like graphics and summaries. Utilizing assessment data is closely linked to what Webb (2002) defines as assessment literacy, which is the understanding of assessing students' knowledge and skills, interpreting assessment outcomes, and applying these insights to enhance student learning and program effectiveness. Rankin (2015) suggests that assessment developers should play a role in aiding educators to utilize assessment results effectively by organizing and providing comprehensive information in result communication. Boudett et al. (2008) stress the importance of result reporting and offer various strategies to assist teachers in result interpretation. Shiel, Kellaghan, and Moran (2010) found that many countries develop Test Reports in tandem with other pertinent assessment data to meet diverse stakeholder needs. To enhance stakeholders' proficiency in interpreting and utilizing assessment data, the establishment of user-friendly online platforms and the provision of training sessions are recommended. Rectifying the existing gaps in reporting and dissemination empowers stakeholders to make informed decisions, ultimately contributing to the enhancement of educational policies and practices. Research indicates that the accessibility of assessment data can significantly impact teachers' instructional approaches (McNaughton, 2012), underscoring the importance of this process. However, it is important to acknowledge the limitations of this study, such as the sample size and the generalizability of findings to other contexts. Future research should explore these aspects further and examine the long-term impact of implementing the proposed model for disseminating NAT results.

5.4 CONCLUSION:

This study highlights the pressing need to reevaluate our national achievement test, particularly in terms of its development and application. The analysis of the difficulty index and discrimination index reveals that, while many questions in the Mathematics and Science tests prove suitable and effective, there exist items in mathematics—specifically in the domains of geometry, measurement, and algebra—and in science—pertaining to physical and life sciences—that demand revision or removal due to their level of difficulty or poor discriminatory characteristics. This underscores the importance of ensuring comprehensive domain coverage, accounting for regional variations, and enhancing item construction to elevate the quality and precision of our assessments.

Moreover, there is widespread consensus among teachers that the National Achievement Test (NAT) serves as a crucial tool in the teaching-learning process, aiding in the identification of weaknesses and areas requiring improvement within the education system. However, our study has also brought to light certain challenges faced by teachers during the application of NAT within their schools. Teachers have emphasized the imperative need for ongoing professional development, sufficient resources, and support to effectively utilize NAT results and thereby enhance student achievement.

Furthermore, concerns have been raised regarding the reporting and dissemination of NAT results, necessitating measures to ensure that assessment outcomes are efficiently communicated to all stakeholders. To address this issue, our study recommends phased approach.

In conclusion, this research underscores the need for a comprehensive reevaluation of our national achievement test, emphasizing improvements in test item quality and addressing the challenges faced by teachers. By implementing these recommendations, we can enhance the effectiveness of the National Achievement Test and, in turn, improve the education system as a whole.

5.6 RECOMMENDATIONS

Based on the findings of this study, the following recommendations are proposed to improve the design, implementation, and impact of the National Achievement Test (NAT):

- Findings from the item analysis revealed that certain items from both NAT 2016 and 2019, particularly in Mathematics and Science, posed significant challenges for students. It is recommended that these items be thoroughly reviewed for clarity, curriculum alignment, and appropriateness. Items that are ambiguous or misaligned should be revised or removed in future assessments to ensure fairness and accurate measurement of student learning.

- During the development phase, significant variability was observed in item difficulty and discrimination indices across content areas. To ensure balance and consistency in test construction, it is recommended to incorporate Automated Test Assembly (ATA) methods. This will enable systematic selection and organization of test items based on predefined psychometric specifications.
- Teacher responses highlighted a lack of training in item development and assessment design. Targeted professional development programs should be initiated to build teacher capacity in instructional strategies, assessment literacy, and content knowledge. These trainings should be designed and delivered by relevant institutions such as provincial education departments, teacher training institutes, and curriculum development bodies.
- Item analysis results indicated specific areas of difficulty for students—namely, Physical and Life Sciences in Science, and Geometry and Algebra in Mathematics. Teachers should be supported to implement focused instructional interventions in these domains. This includes adopting effective teaching strategies, providing additional practice, and integrating resources to enhance students' conceptual understanding.
- The study found significant gaps in how NAT results are communicated to stakeholders. Currently, PIE disseminates raw and scaled scores; however, these lack actionable insights. It is recommended that PIE develop a comprehensive communication strategy that includes:

    - Tailored feedback for students, teachers, schools, and districts.
    - Clear interpretation of scores aligned with curriculum standards and learning outcomes (SLOs).
    - Accessible reporting formats in local languages.
    - A centralized web-based platform for long-term access to performance data, summaries, and reports.
    - Periodic meetings and communication with provinces, schools, and teacher education institutions.
- Low student performance in certain areas suggests a misalignment between the curriculum, teaching materials, and the NAT. Regular curriculum reviews should be undertaken to ensure alignment with assessment goals. Textbooks should comprehensively cover curriculum content, and incorporate audio-visual aids, digital classrooms, and libraries to enhance student engagement and access to learning resources.
- Teacher responses indicated that language barriers may impede both understanding of test items and utilization of test results. Therefore, it is essential to offer language support programs for both teachers and students to improve their comprehension and application of assessment data.
- Resource limitations were cited by teachers as a key barrier to successful NAT implementation. Schools should be provided with adequate materials, ICT tools, and supportive learning environments. Resource allocation must be needs-based, reflecting feedback from teachers and administrators gathered during this study.

- In the post-18th Amendment scenario where education is a provincial subject, improved coordination is essential for the smooth execution of national assessments. Revitalizing the NPCC is crucial for aligning efforts across provinces, ensuring consistent standards, and facilitating collaboration among stakeholders involved in the NAT.
- Given the limitations of the ex post facto design, it is important to interpret current findings with caution, acknowledging the inability to establish causal relationships. Future studies should explore longitudinal or experimental designs to better understand the causal factors affecting student performance in NAT, particularly in ICT and AJK regions.
- The systemic issues highlighted by this study call for collective action from policymakers, education planners, school administrators, and teacher educators. Targeted interventions, informed teacher training, and policy-level reforms must be prioritized to enhance teaching and learning outcomes, especially in Mathematics and Science.

5.6 ECOMMENDATION

5.7 Recommendation for Future Research:

To further enhance the understanding and effectiveness of the National Achievement Test (NAT) in Pakistan, a comprehensive study can be conducted focusing on the following areas:

1. Given the constraints of the ex post facto design, it is essential to interpret the findings cautiously, recognizing the inability to establish causal relationships between variables such as educational policies, teaching methodologies, or socioeconomic factors and NAT performance. Consequently, it is recommended to prioritize the exploration of alternative research methodologies, such as longitudinal studies or experimental designs, in future research endeavors. By adopting these more robust methodologies, researchers can enhance the validity of their findings and provide a more comprehensive understanding of the factors influencing student achievement in regions like ICT and AJK.
2. Investigate the process of item development and validation of NAT of several years to ensure the creation of valid and reliable assessment instruments. This includes exploring the psychometric properties of test items, such as difficulty index and discrimination index, and table refining item construction techniques.
3. Explore the perspectives and experiences of various stakeholders, including students, parents, policymakers, and education administrators, regarding NAT. This can provide valuable insights into their perceptions, challenges, and recommendations for improving the assessment system.
4. Examine the alignment between NAT and the national curriculum to ensure that the assessment accurately measures the intended learning outcomes. This involves assessing the content coverage of NAT items in relation to the curriculum standards and identifying any gaps or areas for improvement.

5. Investigate the impact of NAT on teaching and learning processes in classrooms. This can include examining how teachers utilize NAT results to inform their instructional practices, identify areas of improvement, and provide targeted support to students.
6. Assess the effectiveness of the current reporting and dissemination practices of NAT results. Explore ways to enhance the communication of assessment findings to different stakeholders and ensure that the information is accessible, understandable, and actionable.
7. Investigate the factors influencing student achievement in NAT, such as teaching methodologies, learning environments, socioeconomic factors, and curriculum implementation. Understanding these factors can guide targeted interventions and support systems to improve student performance.
8. Conduct a comparative analysis of NAT results with other regional or international assessments to benchmark student achievement and identify areas of improvement in the Pakistani education system.
9. Evaluate the effectiveness of training and professional development programs provided to teachers in relation to NAT. Assess the impact of these programs on teachers' understanding of assessment techniques, utilization of test results, and instructional practices.
10. Assess the implementation of assessment policies at the national, provincial, and school levels. Examine the challenges, successes, and areas for improvement in the execution of these policies and their impact on student learning outcomes.

11. By conducting a study focusing on these areas, policymakers and education stakeholders can gain valuable insights to inform evidence-based decision-making, improve the design and implementation of NAT, and enhance the overall quality of education in Pakistan.

REFERENCES

Abrams, L. M. (2007). Implications of high-stakes testing for the use of formative classroom assessment. In J. H. McMillan (Ed.), Formative classroom assessment: Theory into practice (pp. 79–98). New York/London: Teacher College, Columbia University.

Ahmed, S., Farooq, M. S., Ahmed, R., & Shah, S. Z. A. (2019). Assessment practices and their challenges in Pakistani higher education. Studies in Educational Evaluation, 61, 211-221.

Akinsola, M. K., & Olowojaiye, F. B. (2008). Teacher Instructional Methods and Student Attitudes towards Mathematics. International Electronic Journal of Mathematics Education, 3(1), 60-73. http://www. iejme.com/download/teacher-instructional-methods-and-student-attitudes-towards-mathematics.pdf

Alberta Education. (2009). Retrieved from http://www.education.gov.ab.ca

American Educational Research Association, American Psychological Association, & National Council of Measurement in Education. (1999). Standards for educational and psychological testing. American Educational Research Association.

Anastasi, A., & Urbina, S. (1997). Psychological testing (7th ed.). Prentice Hall. Brennan, R. L. (2001). Generalizability theory. Springer Science & Business Media.

Anastasi, A., & Urbina, S. (2007). Psychological testing. Dorling Kindersley (Pvt.) Ltd.

Anderson, B., & Postl, B. (2001, April). Using large data sets as a basis for school improvement. Paper presented at the Annual Meeting of the American Educational Research Association, Seattle, WA.

Anderson, J.A. (2005). Accountability in Education. Publisher.

Anderson, L.W., & Krathwohl, D.R. (Eds.). (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's* Taxonomy of Educational Objectives (complete edition). New York: Longman.

Argianti, A., & Retnawati, H. (2020). Characteristics of Math national-standardized school exam test items in junior high school: What must be considered? Jurnal Penelitian Dan Evaluasi Pendidikan, 24(2), 156–165. https://doi.org/10.21831/pep.v24i2.32547

ASER. (2020). *Annual Status of Education Report ASER-PAKISTAN 2019 Urban*. Retrieved from https://docs.edtechhub.org/lib/HBB5FK84

Asian Development Bank. (2019). Pakistan: School education sector assessment [PDF]. Retrieved from https://www.adb.org/sites/default/files/institutional-document/518461/pakistan-school-education-sector-assessment.pdf

Athanasou, J., & Lamprianou, I. (2005). A teacher's guide to assessment. Social Science Press.

Baird, J., Isaacs, T., Johnson, S., Stobart, G., Yu, G., Sprague, T., & Daugherty, R. (2011). Policy effects of PISA. Report Commissioned by Pearson UK. Retrieved from https://www.researchgate.net/profile/Talia_Isaacs/publication/262007629_Policy_effects_of_PISA._Pearson_UK/links/0046353b2b0f37aa20000000.pdf

Baird, J.-A., Johnson, S., Hopfenbeck, T. N., Isaacs, T., Sprague, T., Stobart, G., Yu, G. (2016). On the supranational spell of PISA in policy. Educational Research, 58, 121–138.

Baker, J. (2019, April 12). People are frightened of NAPLAN: Australia's testing dilemma. The Sydney Morning Herald. Retrieved from https://www.smh.com.au/education/people-are-frightened-of-naplan-australia-s-testing-dilemma-20190411-p51dcu.html

Ballard, K., & Bates, A. (2008). Making a connection between student achievement,

Bano, A., Saeed, M., & Hussain, T. (2017). School-based assessment practices: A study of 2teachers' perceptions in Pakistan. Assessment in Education: Principles, Policy & Practice, 24(2), 221-238.

Bazaldua, D. A. L., Lee, Y. S., Keller, B., & Fellers, L. (2017). Assessing the performance of classical test theory item discrimination estimators in Monte Carlo simulations. *Asia Pacific Education Review*, *18*(4), 585-598.

Beaton, A. E., & Johnson, E. G. (1992). Overview of the scaling methodology used in the National Assessment IRT. Journal of Educational Measurement, 29(2), 163-175.

Beck, K. (2020). Ensuring content validity of psychological and educational tests – the role of experts. Journal of Educational Measurement, 50(1), 1–73.

Bejar, I. I. (1983). Achievement Testing: Recent Advances. Sage Publications.

Benavot, A., & Kamens, D. H, (2011). National, regional and international learning assessments: trends among developing countries, 1960–2009. Globalization, Societies and Education, 9(3-4), https://doi.org/[DOI]

Benavot, A., and E. Tanner. 2007. The Growth of National Learning Assessments in the World, 1995-2006. Background Paper Prepared for the Education for All Global Monitoring Report. 2008. Paris: UNESCO

Bengal. (n.d.). Journal of Dental and Medical Sciences, 14(12), 47-52.

Bichi, A. A. (2016). CLASSICAL TEST THEORY: An Introduction to Linear Modeling Approach to Test and Item Analysis. International Journal for Social Studies, 2(9), 27. Retrieved from https://edupediapublications.org/journals

Birnbaum, A. (1968). Some latent trait models and their use in inferring an examinee's ability. In F. M. Lord & M. R. Novick (Eds.), Statistical theories of mental test scores (pp. 397-479). Reading, MA: Addison-Wesley.

Black P.& William, D. (2009). Developing the theory of formative assessment. Educational Assessment Evaluation and Accountability, 21.5-31 10.1007/s11092-008-9068-5.

Bliem, C. L. and Davinroy, K. H. (1997) *Teachers' Beliefs about Assessment and Instruction in Literacy*. CSE Technical Report 421. CRESST, Graduate School of Education & Information Studies, University of California. Available from https://www.cse.ucla.edu/products/reports/TECH421.pdf (accessed 26 July 2015).

Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: a primer. *Frontiers in public health*, *6*, 149.

Booth, N. (2017). What is formative assessment, why hasn't it worked in schools, and how can we make it better in the classroom? Impact: Journal of the Chartered College at Teaching. Retrieved from https://impact.chartered college/article/booth-what-formative-assessment-make-better-classroom

Boudett, K. P., City, E. A., & Murnane, R. J. (Eds.). (2008). Data Wise in Action: Stories of Schools Using Data to Improve Teaching and Learning. Cambridge, MA: Harvard Education Press.

Boyatzis, R. E. (1998). Transforming qualitative information: Thematic analysis and code development. Sage.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. Qualitative Research in Psychology, 3(2), 77-101.

Breakspear, S. (2012). The Policy Impact of PISA: An Exploration of the Normative Effects of International Benchmarking in School System Performance. OECD Education Working Papers, No. 71, New York: OECD. http://dx.doi.org/10.1787/5k9fdfqffr28-en

Breakspear, S. (2012). The policy impact of PISA: An exploration of the normative effects of international benchmarking in school system performance (OECD Education Working Papers, No. 71). Paris: OECD Publishing.

Breiter, A., & Light, D. (2004a). Decision Support Systems in Schools – from Data Collection to Decision

Breiter, A., & Light, D. (2004a). Decision Support Systems in Schools – from Data Collection to Decision Making. Paper presented at the America's Conference on Information Systems (AMCIS), New York, NY.

Brennan, R. L. (2000). (Mis) conceptions about generalizability theory. Educational Measurement: Issues and Practice, 19 (1), 5–10.

Broadfoot, P. (2007). An introduction to assessment. Continuum International Publishing Group. Darling-Hammond, L. (2010). The flat world and education: How America's commitment to equity will determine our future. Teachers College Press.

Brown, G.T.L. (2011) School Based Assessment Methods: Development and Implementation. Journal of Assessment Prodigious, 1, 30-32.

Bryman, A. (2012). Social research methods (4th ed.). Oxford: Oxford University Press.

Burgess Y. (2020). A Rasch investigation of type I error rates and power associated with item fit statistics under large-scale testing situations [Doctoral dissertation]. https://scholarcommons.sc.edu/etd/6005

Charlotte, NC: Information Age Publishing.

Cherry, K. (2020). Why Alfred Binet Developed IQ Testing for Students. Retrieved from https://www.verywellmind.com/history-of-intelligence-testing-2795581

Christie, P., & Potterton, M. (2011). 2010 Annual National Assessment of South African Learning Achievement at the Lower Grades.

Cimbricz, S. (2002). State-Mandated Testing and Teachers' Beliefs and Practice. Educational Policy Analysis Archives [On-line serial], 10(2). Available: http://epaa.asu.edu/epaa/v10n2.html

Clark, I. (2012). Formative assessment: Assessment is for self-regulated learning. *Educational Psychology Review, 24*(2), 205–249. https://doi.org/10.1007/s10648-011-9191-6

Coburn, C. E., Honig, M. I., & Stein, M. K. (2009). What's the evidence on districts' use of evidence? In J. D. Bransford, D. J. Stipek, N. J. Vye, L. M. Gomez, & D. Lam (Eds.), The Role of Research in Educational Improvement (pp. 67-86). Harvard Education Press

Collins, J. (2009, September 21). From good to great: What defines a level V leader? [Video]. Retrieved from http://www.youtube.com/watch?v=q-KyQ90XByY

Collins, N. (2009). Motivation and self-regulated learning. The Journal of Higher Education (Columbus, Ohio), 80(4), 476-479. http://dx.doi.org/10.1353/jhe.0.0057

Cowie, B. (2012). A model of formative assessment in science education. Assessment in Education: Principles, Policy & Practice, 6(1), 101-116. https://doi.org/10.1080/09695949993026

Cowie, B. (2012). Focusing on the Classroom: Assessment for Learning. In B. J. Fraser, K. G.Tobin, & C. J. McRobbie (Eds.), Second International Handbook of Science Education(pp. 679-690). London: Springer.

Crocker, L. M., & Algina, J. (1986). Introduction to classical and modern test theory. Holt, Rinehart and Winston.

Cronbach, L. & Meehl J. (1995). Construct validity after thirty years. In R. L. Linn (Ed.), Educational measurement (3rd ed., pp. 161-182). American Council on Education and Macmillan Publishing Company.

Darling-Hammond, L. (2010). Performance counts: Assessment systems that support high-quality learning. Washington, DC: Council for Chief State School Officers.

David Kamens and A. Benavot. 2011, "National, regional and international learning assessments: Trends among developing countries, 1960-2009." Globalisation, Societies and Education 9 (2) 285-300

Davie, L. E, Livingstone, D. W., & Hart, D., (Eds.). (1999). Public Attitudes towards Education in Ontario 1998: The Twelfth OISE/UT Survey. Publisher.

De Bortoli, L & Thomson, S. (2009).The achievement of Australia's Indigenous students in PISA 2000-2006.Camberwell: ACER.

De Lange, J. (2007). Large-scale assessment and mathematics education. In F. K. Lester, Jr. (Ed.), Second handbook of research on mathematics teaching and learning (pp. 1111–1142). Charlotte, NC: Information Age Publishing

DeMars, C. (2010). *Item response theory*. Oxford University Press.

Department of Basic Education. (2012). Report on the Annual National Assessments of 2011.

Department of Education School-Based Management Technical Working Group. (2009). Basic Education Sector Reform Agenda (BESRA). DepED Complex, Meralco Avenue, Pasig City 1600.

DepED Order No. 91, s. 2009, "Prescribing Guidelines and Requirements Regulating the Use of National Examination Results." Retrieved May 25, 2015 from www.deped.gov.ph

Dewey, J. (1975). Democracy and education: An introduction to the philosophy of education. New York, NY: Free Press.

Dewey, J. (1975). Democracy and education: An introduction to the philosophy of education. New York, NY: Free Press.

Dillion, S. (2010, December 7). Top test scores from Shanghai stun educators. New York Times.

Retrieved from http://www.nytimes.com/2010/12/07/education/07education.html?pagewanted¼1&_r¼2

Dissemination Strategies for the National Assessment of Educational Progress, The Widmeyer Group, p. 7. 4. NAEP 1992

Dobbin, F., Simmons, B., & Garrett, G. (2007). The global diffusion of public policies: Social construction, coercion, competition, or learning? Annual Review of Sociology, 33, 449–472.

Dobbins, M., & Christ, C. (2019). Do they matter in education politics? The influence of political parties and teacher unions on school governance reforms in Spain. Journal of Education Policy, 34(1), 61–82.

Dolton, P., & MacInnes, J. (2006). The impact of parental income and education on the schooling of their children. Scottish Journal of Political Economy, 53(5), 430-460.

Doshi, R., Kelley, J., & Simmons, B. (2004). The Power of Ranking: The Ease of Doing Business Indicator and Global Regulatory Behavior. International Organization, 1-33.

Duckworth, A. L., Yeager, D. S. (2015). Measurement matters: Assessing personal qualities other than cognitive ability for educational purposes. Educational Researcher, 44, 237–251.

Durry, C,. (2010). Building environmental assessment methods: Applications and development trends. Journal of Building Performance Evaluation, 21(4), 300-308. https://doi.org/10.1080/096132199369417

Earl, L. M. (2012). Assessment as Learning: Using Classroom Assessment to Maximize Student Learning (2nd ed.). SAGE Publications.

Earl, L., & Torrance, N. (2000). Embedding Accountability and Improvement Into Large-Scale Assessment: What Difference Does It Make? Peabody Journal of Education, 75, 114-141. https://doi.org/10.1207/S15327930PJE7504_6

Earl, L., & Torrance, N. (2000). Embedding Accountability and Improvement Into Large-Scale Assessment: What Difference Does It Make? Peabody Journal of Education, 75(4), [Page range]. https://doi.org/10.1207/S15327930PJE7504_6

Ebel, R. L. (1979). *Essentials of educational measurement* (3rd ed.). Englewood Cliffs, NJ: Prentice Hall. [Google Scholar]

Education Quality and Accountability Office (EQAO). (2009-2010). EQAO Test Results [PDF]. EQAO.

Elementary Teachers' Federation of Ontario. (2012). EQAO testing. Retrieved October 26, 2012, from http://www.etfo.ca/issuesineducation/eqaotesting/pages/default.aspx.

Embretson, S. E. (1983). Construct validity: Construct representation versus nomothetic span. Psychological Bulletin, 93(1), 179-197.

Figazzolo, L. (2009). Impact of PISA 2006 on the education policy debate. Retrieved from http://download.eiie.org/docs/IRISDocuments/ResearchWebsiteDocuments/2009-00036-01-E.p

Finn, B. (2015). Measuring Motivation in Low-Stakes Assessments. ETS Research Report SS Series

Finney, S. J. (2011). Low-stakes testing and psychological reactance: Using the Hong Psychological Reactance Scale to better understand compliant and non-compliant examinees. International Journal of Testing, 11, 248-270. https://doi.org/10.1080/15305058.2011.570884

Fisher, W. P. (2011). Invariant thinking vs. invariant instrument. Rasch Measurement Transactions, 14(1), 736-737.

Fives, H., & Buehl, M. M. (2012). Spring cleaning for the "messy" construct of teachers' beliefs: What are they? Which have been examined? What can they tell us? In K. R. Harris, S. Graham, & T. Urdan (Eds.), APA Educational Psychology Handbook: Individual Differences and Cultural and Contextual Factors (Vol. 2, pp. 471–499). Washington: American Psychological Association

Flake, J. K., Pek, J., & Hehman, E. (2017). Construct Validation in Social and Personality Research: Current Practice and Recommendations. *Social Psychological and Personality Science*, *8*(4), 370-378. https://doi.org/10.1177/1948550617693063

Foy, P., & Olson, J. F. (Eds.). (TIMSS 2007). TIMSS 2007 User Guide for the International Database. Retrieved from https://timssandpirls.bc.edu/TIMSS2007/PDF/TIMSS2007_UserGuide.pdf Top of Form

Frederiksen, J. R., & Collins, A. (1989). A systems approach to educational testing. Educational Researcher, 18(9), 27-32.

Freedman and Wilson and a last word. Interchange, 26(1), 81-96.

Froese-Germain, Bernie (2004). Educational Accountability With a Human Face. Ottawa: Canadian Teachers' Federation.

Fulcher, K. H (2014). A simple model for learning improvement: University of Illinois and Indiana University, National Institute for Learning Outcomes Assessment (NILOA).

Fullan M. The return of large-scale educational reform. *Journal of Educational Change*. 2000;1(1):5–28. [Google Scholar].

Gambell, T., & Hunter, D. (2004). Teacher scoring of large-scale assessment: Professional development or debilitation? Curriculum Studies, Curriculum and Pedagogy. https://doi.org/10.1080/0022027032000190696.

Gay, L.R. (1996). Educational research. New Jersey: Prentice Hall, Inc

Gierl, M. J., Bisanz, J., Bisanz, G. L., Boughton, K. A., & Khaliq, S. N. (2001). Illustrating the utility of differential bundle functioning analyses to identify and interpret group differences on achievement tests. Educational Measurement: Issues and Practice, 20(2), 26–36. https://doi.org/10.1111/j.1745-3992.2001.tb00060.x

Giordano, G. (2005). How testing came to dominate American schools. New York, NY: Peter Lang.

Global Partnership for Education. (2018). GPE Results Report 2018. Retrieved from https://www.globalpartnership.org/sites/default/files/2018-06-gpe-results-report-2018-web4.pdf

Goldberg, G. L., & Roswell, B. S. (2000). From perception to practice: The impact of teachers' scoring experience on performance-based instruction and classroom assessment. Educational Assessment, 6, 257–290.

Goodwin, L. D., & Leech, N. L. (2003). The meaning of validity in the new standards for educational and psychological testing: Implications for measurement courses. Pages 181-191. https://doi.org/10.1080/07481756.2003.11909741

Green, A. (2013). Exploring language assessment and testing: Language in action (1st ed.). Routledge. https://doi.org/10.4324/9781315889627

Grek, S. (2009). Governing by the numbers: The PISA "effect" in Europe. Journal of Education Policy, 24(1), 23–37.

Gronlund, N. E. (1985). *Measurement and Evaluation in Teaching* . New York, NY: Ma Mehrens, W. A. , & Lehman, I. J. (1991). *Measurement and evaluation in education and psychology* (4th ed.). Belmont, CA: Wadsworth/Thomson Learning. [Google Scholar]cmillan. [Google Scholar]

Gruber, K. H. (2006). The German "PISA-shock": Some aspects of the extraordinary impact of the OECD's PISA study on the German education system. Paper presented at the Cross-national attraction in education: Accounts from England and Germany. Oxford, England: Symposium Books

Guest, G., MacQueen, K. M., & Namey, E. E. (2012). Applied Thematic Analysis. Sage.

Gulliksen, H. (1950). The reliability of speeded tests. Psychometrika, 15(2), 111–125. https://doi.org/10.1002/j.2333-8504.1950.tb00876.x

Gutiérrez DMA, Vázquez F G (2008) Resultados nacionales: segundo estudio regional comparativo y explicativo 2006 (SERCE).Mexico, DF: Instituto Nacional para la Evaluación de la Educación (INEE).

Haertel, E. H., & Herman, J. L. (2005). Uses and misuses of tests: Three case studies. Educational Measurement: Issues and Practice, 24(1), 28-37.

Haladyna, T. M., & Downing, S. M. (Eds.). (2011). Handbook of Test Development. Routledge.

Hambleton, R. K., & Jones, R. W. (1993). An NCME instructional module on. *Educational measurement: issues and practice*, *12*(3), 38-47.

Hamilton, L. S., & Snow, R. E. (1997). Enhancing the validity and usefulness of large-scale educational assessments: IV. NELS:88 science achievement to 12th grade. *American Educational Research Journal, 34*(1), 151–173.

Hamilton, L. S., Nussbaum, E. M., & Snow, R. E. (1997). Interview procedures for validating science assessments. Applied Measurement in Education, 10(2), 181–200. https://doi.org/10.1207/s15324818ame1002_5

Harlen, W. (2006). 'Assessment for learning and assessment of learning.' In: Harlen, W. (Ed) *ASE Guide to Primary Science.* Hatfield: Association for Science Education.

Hart, W. A. (2012). What lessons can we learn? Journal Name, Volume(Number), Page range. https://doi.org/10.1111/j.1467-9752.2012.00882.x

Hatch, T. (2013). Beneath the surface of accountability: Answerability, responsibility and capacity-building in recent education reforms in Norway. Journal of Educational Change, 14(2), 113–138.

Hertzog, M.A. (2008). Considerations in determining sample size for pilot studies . Research in Nursing & Health, 31, 180-191.

Hijji, B. M. (2017). Flaws of multiple choice questions in teacher-constructed nursing examinations: A pilot descriptive study. Journal of Nursing Education, 56(8), 490-496. doi: 10.3928/01484834-20170712-08

Ho, A. D., (2012). Estimating achievement gaps from test scores reported in ordinal "proficiency" categories. *Journal of Educational and Behavioral Statistics* , *37*, 489-517.

Hopfenbeck, T. N., & Maul, A. (2013). Examining evidence for the validity of PISA learning strategy scales based on student response processes. International Journal of Testing, 11(2), 95–121. doi:10.1080/15305058.2010.529977

Hotiu, A. (2006). *The relationship between item difficulty and discrimination indices in multiple-choice tests in a Physical science course* (MSc thesis). Florida Atlantic University, Boca Raton, FL. Retrieved June 14, 2015 from www.physics.Au/eAdu/research/education/A.Hotiu_thesis.pdf [Google Scholar]

Howie, S., Van Staden, S., Tshele, M., Dowse, C., & Zimmerman, L. (2012). PIRLS 2011, South African children's reading achievement literacy achievement: Summary report. Centre for Evaluation and Assessment, University of Pretoria, Pretoria.

Hungi, N., Makuwa, D., Ross, K., Saito, M., Dolata, S., Van Cappelle, F., et al. (2010). SACMEQ project results: Pupil achievement levels in reading and mathematics (Working Document Number 1). Retrieved October 9, 2018, from http://www.wd01-sacmeq-iiiresultspupil.

Hussain Alkharusi (2008) Effects of Classroom Assessment Practices on Students' Achievement Goals, Educational Assessment, 13:4, 243-266, DOI: 10.1080/10627190802602509

Ifeanyieze, F.O. and Aneke, K. (2013) Competencies Required by Teachers in School Based Assessment of Student in Secondary School. Nigerian Journal of Educational Research and Evaluation, 12, 29-37.

Implementation of Large-Scale Education Assessments, First Edition. Edited by Petra Lietz, John C. Cresswell, Keith F. Rust and Raymond J. Adams. © 2017 John Wiley & Sons Ltd. Published 2017 by John Wiley & Sons Ltd.

Imran, M. & Earthman, G. I. (2002). School facility conditions and student academic achievement. Los Angeles, CA: UCLA's Institute for Democracy, Education, and Access (IDEA)

Jennings, B. M., Staggers, N., & Brosch, L. R. (1999). A classification scheme for outcome indicators. Journal Name, Volume(Number), Page range. https://doi.org/10.1111/j.1547-5069.1999.tb00524.x

Johanson, G. and Brooks, G. (2010) Initial Scale Development: Sample Size for Pilot Studies. Educational and Psychological Measurement, 70, 394-400. http://dx.doi.org/10.1177/0013164409355692

Johansson, S. (2016). International large-scale assessments: What uses, what consequences? *Educational Research*, *58* (2), 139-148.

Johnson, E. G. (1992). The design of the National Assessment of Educational Progress. Journal of Educational Measurement, 29(2), 95–110. https://doi.org/10.1111/j.1745-3984.1992.tb00369.x

Joseph, G. (2013). A Study on School Factors Influencing Students' Attitude Towards Learning Mathematics in the Community Secondary Schools in Tanzania: The case of Bukoba Municipal Council in Kagera Region. (Masters dissertation). Retrieved from http://repository.out.ac.tz/919/

Joseph, G. (2013). A Study on School Factors Influencing Students' Attitude Towards Learning Mathematics in the Community Secondary Schools in Tanzania: The case of Bukoba Municipal Council in Kagera Region. (Masters dissertation). Retrieved from http://repository.out.ac.tz/919

Kays, J. L., Holmes, S. D., Corales, T. A., Deise, R. K. (2015). Item response theory and health outcomes measurement in the 21st century. Medical Care, 38(9 Suppl), S11-S18.

Kellaghan, T., & Greaney, V. (2001). Using assessment to improve the quality of education. Paris: International Institute for Educational Planning.

Kellaghan, T., & Greaney, V. (2009, June 23). Using the results of a national assessment of educational achievement. Top of Form

Khan, S. H., & Saeed, M. (2010). Evaluating the quality of BEd programme: Students' views of their college experiences. Journal Name, Volume(Issue), Page numbers.

KHAN, S. (2006) An evaluation of the exercises provided in the English compulsory textbook for class X, [Unpublished MA dissertation] Faculty of English Linguistics, University of Karachi.

Khattak, S. G. (2012). Assessment in schools in Pakistan. SA-eDUC, 9(2), 129-138, 1-13.

Kirkpatrick, R., & Zang, Y. (2011). The negative influences of exam-oriented education on Chinese high school students: Backwash from classroom to child. Language testing in Asia, 1, 36.

Kirsch, I., Lennon, M., von Davier, M., Gonzalez, E., & Yamamoto, K. (2013). On the growing importance of international large-scale assessments. In von Davier, M., Gonzalez, E., Kirsch, I., Yamamoto, K. (Eds.), The Role of International Large-Scale Assessments: Perspectives from Technology, Economy, and Educational Research (pp 1-11). New York, NY: Springer.

Klinger, D. A., DeLuca, C., & Miller, T. (2008). The evolving culture of large-scale assessments in Canadian education. Canadian Journal of Educational Administration and Policy, (76), 1-34.

Klinger, D. A., Maggi, S., & D'Angiulli, A. (2011). School accountability and assessment: Should we put the roof first? The Educational Forum, 75(2), 114–128.

Klinger. D. A., & Rogers, W. T. (2011). Teachers' perceptions of large-scale assessment programs within low-stakes accountability frameworks. International Journal of Testing, 11(2), 122–143

Köhler C., Robitzsch A., Hartig J. (2020). A bias-corrected RMSD item fit statistic: An evaluation and comparison to alternatives. Journal of Educational and Behavioral Statistics, 45(3), 251-273.

Krishnan, V. (2013). *The early child development instrument (EDI): An item analysis using classical test theory (CTT) on Alberta's data*. Community-University Partnership (CUP), Faculty of Extension, University of Alberta, Edmonton, Alberta, Canada.

Labaree, D. F. (1997). Public goods, private goods: The American struggle over educational goals. American Educational Research Journal, 34, 39–81.

Lane, C. D. (1999). Equity and excellence in educational testing and assessment. Educational Assessment, 17(4). https://doi.org/10.1177/073428299901700408

Liang, L. L., Chen, S., Chen, X., Kaya, O. N., Adams, A. D., Macklin, M., & Ebenezer, J. (2006). Student Understanding of Science and Scientific Inquiry (SUSSI): Revision and Further Validation of an

Assessment Instrument. Paper presented at the 2006 Annual Conference of the National Association for Research in Science Teaching (NARST), San Francisco, CA, April 3-6, 2006.

Le Grange, L., Reddy, C. (1998). Continuous Assessment: An Introduction and Guidelines to Implementation. South Africa: Juta.

Lee J, Kang C. A litmus test of school accountability policy effects in Korea: cross-validating high-stakes test results for academic excellence and equity. *Asia Pacific Journal of Education.* 2019 doi: 10.1080/02188791.2019.1598851. [CrossRef] [Google Scholar]

Leighton, J. P., (2004). *Cognitive diagnostic assessment for education: Theory and applications.* Cambridge University Press. https://doi.org/10.1017/CBO9780511611186

Lewis, S. 2000. 'PISA "Yet To Come": Governing Schooling through Time, Difference

Linn, R. L., & Baker, E. L. (1996). Accountability systems: Implications of requirements of the No Child Left behind Act of 2001. Educational Researcher, 35(4), 3-14.

Linn. R. L. (2000). Assessment and accountability. Educational Researcher, 29(2), 4–17.

Linn. R. L. (2003, July). Accountability: Responsibility and seasonable expectations (Report No. 601). Retrieved from http://www.cse.ucla.edu/products/reports/r601.pdf

Lissitz, R. W., & Samuelsen, K. (2007). A suggested change in terminology and emphasis regarding validity and education. Educational Researcher, 36, 437-448.

Lockheed, M., and A. Verspoor. 1990. Improving Primary Education in Developing Countries. Washington, DC: World Bank

Lord, F. M., & Novick, M. R. (1968). Statistical theories of mental test scores. Addison-Wesley Publishing Company.

MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. Psychological Methods, 4(1), 84–99. https://doi.org/10.1037/1082-989X.4.1.84

Madaus, G. F., West, M. M., Harmon, M. C., Lomax, R. G., & Viator, K. A. (1992). *The influence of testing on teaching math and science in grades 4–12: Executive summary*. Chestnut Hill, MA: Boston College, Center for the Study of Testing, Evaluation, and Educational Policy.

Mahmud, J. Item response theory: A basic concept. Educational Research and Reviews. Vol. 12(5), (pp. 258-266) 10 March, 2017, http://www.academicjournals.org/journal/ERR/article-fulltext-pdf/53ED36963043, doi: 10.5897/ERR2017. 3147. (2017).

Maki, Peggy L. "Developing an Assessment Plan to Learn about Student Learning." *The Journal of Academic Librarianship* 28.1 (2002): 8–13. *ScienceDirect*. Web. The Journal of Academic Librarianship

Making. Paper presented at the America's Conference on Information Systems (AMCIS), New York, NY

Maldives. International Journal of humanities and social science, 1(15), 277-281. Retrieved from

Maligalig, D. S., & Albert, J. R. G. (2008). Measures for Assessing Basic Education in the Philippines (Discussion Paper Series No. 2008-16).

Marcussen, M. (2004). Multilateral surveillance and the OECD: Playing the idea game. In K. Armingeon & M. Beyeler (Eds.), The OECD and European welfare states (pp. 13–31). Cheltenham, UK/Northampton, MA: Edward Elgar Publishing.

Mardapi, D. (2012). Pengukuran Penilaian dan Evaluasi Pendidikan. Nuha Medika

Martins, L., & Veiga, P. (2010). Do inequalities in parents' education play an important role in PISA student mathematics achievement test score disparities? Economics of Education Review, 29(6), 1016–1033. doi:10.1016/j. econedurev.2010.05.001

Mata, M. D., Monteiro, V., & Peixoto, F. (2012). Attitudes towards mathematics: Effects of individual, motivational, and social support factors. Child development research, 2012. https://doi.org/10.1155/2012/876028.

McBride, B.J., (2001). School-based programs for young children with autism: Blending approaches to meet individual needs. Paper presented at the International Council for Exceptional Children Conference, Vancouver, BC, Canada.

Mcleod, S. (2007). Visual Perception Theory. Retrieved Oct. 07,2015 from http://www.simplypsychology.org/perception-theories.html

McNaughton, D., (2012). AAC and Communication in the Workplace. Webcast available at http://mcn.educ.psu.edu/dbm/2012/2012_aac_emp/2012_aac_emp.htm

Meaghan, D. E., & Casas, F. R. (1995). On standardized achievement testing: Response to

Meckes, L., & Carrasco, R. (2010). Two decades of SIMCE: An overview of the National Assessment System in Chile. *Assessment in Education: Principles, Policy & Practice*. doi:10.1080/09695941003696214

Mehrens, W.A. & Lehmann, I.J. (1991). *Measurement and evaluation in education and psychology*. (3rd Ed.) New York: Holt, Rinehart and Winston.

Mertler, C. (2010). Secondary teachers' assessment literacy: Does classroom experience make a difference? American Secondary Education, 33, 49–64.

Messick, S. (1989). Validity. In R. L. Linn (Ed.), Educational measurement (3rd ed., pp. 13-103). American Council on Education and Macmillan Publishing Company.

Messick, S. (1994). Meaning and values in test validation: The science and ethics of assessment. Educational Researcher, 18(2), 5-11.

Meyer, H.-D., & Benavot, A. (2013). PISA, power and policy the emergence of global educational governance. Oxford: Oxford Studies in Comparative Education, Symposium Books.

Ministry of Education. (1998). Ministry of Education position paper: Assessment (schooling sector). Wellington: Learning Media.

Ministry of Education. (2007). The New Zealand curriculum. Wellington: Learning Media

Mmereki, O. A. (2011). Quality of Primary School Inputs in Botswana. Gaborone: Ministry of Education

Mohamed, L., & Waheed, H. (2011). Secondary students' attitude towards mathematics in a selected school of Maldives. International Journal of humanities and social science, 1(15), 277-281. Retrieved from https://www.researchgate.net/profile/Hussain_Waheed/publication/266009828.

Mohamed, L., & Waheed, H. (2011). Secondary students' attitude towards mathematics in a selected school of Maldives. International Journal of humanities and social science, 1(15), 277-281. Retrieved from http://bit.ly/2VU1IVA

Mohamed, L., & Waheed, H. (2011). Secondary students' attitude towards mathematics in a selected school of

Mohan, R. (2023). Measurement, Evaluation and Assessment in Education (2nd ed.). Publisher.

Moscovich, I. (2006). The Big Book of Brain Games. New York: Workman Publishing

Mukherjee, P. & Lahiri, S. K. (2015). Analysis of multiple-choice questions (MCQs): Item

Mullis, I., Martin, (2019). Mathematics benchmarking report TIMSS, Eighth grade. Chestnut Hill, MA: Boston College.

Mullis, I., Martin, M., Beaton, A., Gonzalez, E., Kelly, D., & Smith, T. A. (1998). Mathematics achievement in Missouri and Oregon in an international context: 1997 TIMSS benchmarking. Chestnut Hill, MA: TIMSS International Study Center.

Mulyasa. (2014). Menjadi Guru Profesional: Menciptakan Pembelajaran Kreatif dan Menyenangkan. PT Remaja Rosdakarya.

Mutai, K. J. (2011). Attitudes towards learning and performance in mathematics among students in selected secondary schools in Bureti district, Kenya (Masters Dissertation). Retrieved from http://irlibrary.ku.ac.ke/bitstream/handle/123456789/609/JACKSON%20KIPRONOH.

Nagy, P. (2000). The three roles of assessment: Gatekeeping, accountability, and instructional diagnosis. Canadian Journal of Education, 25, 262-2

National Achievement Test [2016]. Retrieved from http://neas.gov.pk/PublicationDetail/ZTM5YTk5OWEtZDYxNi00NTZiLTgyYzUtZWNlNGIxZDQxZjJk

National Center for Education Statistics (NCES). (2019). NAEP 2017 Mathematics and Reading

National Education Assessment System (NEAS). (2016). Short Report of NAT 2016. http://neas.gov.pk/SiteImage/Misc/files/Short-Report-of-NAT-2016.pdf

Nelson, R. F. (2000). Personal and environmental factors that influence early childhood teachers' practice. Journal of Instructional Psychology, 27(2). Retrieved from http://findarticles.com/p/articles/mi_m0FCG/is_2_27/ai_63365162/?tag=content ;col1

Ngussa, B. M., & Mbuti, E. E. (2017). The Influence of Humour on Learners' Attitude and Mathematics Achievement: A Case of Secondary Schools in Arusha City, Tanzania. Journal of Educational Research, 2(3), 170 -181. Retrieved from https://www.researchgate.net/publication/315776039

Nichols, S. L., & Berliner, D. C. (2007). Collateral damage: How high-stakes testing corrupts America's schools. Cambridge, MA: Harvard Education Press

Nichols, S. L., Berliner, D. C. (2007). Collateral damage: How high-stakes testing corrupts America's schools. Cambridge, MA: Harvard Education Press.

No Child Left Behind Act of 2001 . (2002). Public Law No. 107–110, Congress.

Nzomo, N. & Makuwa, D.K. (2006) How Can Countries Move From Cross-national Research Results to Dissemination, and then to Policy Reform? Case Studies from Kenya and Namibia, in Ross, K.N. & Genevois, I.L. (Eds) Cross-national Studies and the Quality of Education: planning their design and managing their impact, pp 213–228. Paris: IIEP-UNESCO.

OECD (2011a), Education at a Glance 2011: OECD Indicators, OECD, Paris.

Oermann, M. H. & Gaberson K. B. (2014). Evaluation and testing in nursing education of multiple choice questions of speech sounds of English. *International Research Journal of MMC*, *2*(1), 15-28.

Onuka, O.U.A. and Akinyemi, T.F. (2012. Assessment as a Tool for Learning Improvement at the SSS Education Level. In: Abe Charles, V. and Onuka, A.O., Eds., Revolutionizing Assessment and Evaluation Procedures in Education, Stirling-Horden Publishers Ltd., Ibadan, 37-52

Organisation for Economic Co-operation and Development . (2016a). PISA 2015: Results in focus. PISA. Retrieved from https://www.oecd.org/pisa/pisa-2015-results-in-focus.pdf

Organisation for Economic Co-operation and Development. (2016a). PISA 2015: Results in focus. PISA. Retrieved from https://www.oecd.org/pisa/pisa-2015-results-in-focus.pdf

Organisation for Economic Co-operation and Development. (2016b). PISA 2015 results (Volume I): Excellence and equity in education. Paris: OECD Publishing.

Organization for Economic Co-operation and Development. (2012). PISA 2015 results (Volume I): Excellence and equity in education. Paris: OECD Publishing.

Oyebola, J.O. (2013) Assessment: A Transformation Tool for Enhancing Students' Learning. Nigerian Journal of Educational Research and Evaluation, 12, 92-100.

Popham, W. J. (2001). Big change questions: "Should large-scale assessement be used for accountability?", Journal of Educational Change, 1, 283–289.

Preacher, K. J., & MacCallum, R. C. (2002). Exploratory factor analysis in behavior genetics research: factor recovery with small sample sizes. Behavior Genetics, 32(2), 153–161. https://doi.org/10.1023/A:1015210025234

Reading Report Card for the Nation and the States, NCES, Office of Educational Research and Improvement, U.S. Department of Education, September 1993.

Reddy, V., Juan, A., van der Berg, S., & van Broekhuizen, H. (2016). South African National Assessment Program: An Analysis of Mathematics and Language Achievement in Schools.

Rehman, Abdul & Aslam, Ayesha & Hassan, Syed. (2018). Item Analysis Of Multiple Choice Questions. 38. 291-93.

Rezigalla, A. A. (2022). Item Analysis: Concept and Application. *Medical Education for the 21st Century*, 105.

Rust, C. (2002) Learning & Teaching Briefing Paper Series. Oxford Centre for Staff & Learning Development OCSLD, Oxfords Brookes University.

Sanders, W. L., & Horn, S. P. (1998). Research findings from the Tennessee Value-Added Assessment System (TVAAS) database: Implications for educational evaluation and research. *Journal of Personnel Evaluation in Education, 12*(3), 247–256. https://doi.org/10.1023/A:1008067210518

Santrock, J. W., Woloshyn, V. E., Gallagher, T. L., Di Petta, T., & Marini, Z. A. (2004). Educational Psychology (1st Canadian ed.). Toronto, ON: McGraw-Hill Ryerson

Sarmah, A., & Puri, P. (2014). Attitude towards Mathematics of the Students Studying in Diploma Engineering Institute (Polytechnic) of Sikkim. Journal of Research & Method in Education, 4(6). Retrieved from http://www.academia.edu/download/36434404/B04630610.pdf

Saylor.org (n.d). Sensation and Perception, retrieved June 14,2015 from

Schildkamp, K., & Lai, M. K. (2013). Data-based decision making : conclusions and a data use framework. In K. Schildkamp, M. K. Lai, & L. Earl (Eds.), *Data-based decision making in education: Challenges and opportunities* (pp. 177-191). (Studies in educational leadership; Vol. 17, No. 17). Springer. https://doi.org/10.1007/978-94-007-4816-3_10

Schleicher, D. J.,(2019). Evaluating the effectiveness of performance management: A 30-year integrative conceptual review. Journal of Applied Psychology, 104(7), 851–887. https://doi.org/10.1037/apl0000368.

Scott, J. (1990). Assessing documentary sources. In: A matter o f record. Documentary sources in social research (pp. 19-35). C am bridge: Polity Press. ISBN 978-0-7456-0070-

Scott, J.C., & Mead, A.D. (2011). Foundations for Measurement.

Scott, S., Scott, D. E., & Webber, C. F. (2015). Assessment in Education: Implications for Leadership. Publisher.

Serafini, F. (2001). Three paradigms of assessment: Measurement, procedure, and inquiry. The Reading Teacher, 54.4, 384–39

Sharma, L. R. (2021). Analysis of difficulty index, discrimination index and distractor efficiency

Shiel, G., Kellaghan, T., & Moran, G. (May 2010). Standardised Testing in Lower Secondary Education. Educational Research Centre, St Patrick's College, Dublin.

Shohamy, E. (2012). 10. A Critical Perspective on the Use of English as a Medium of Instruction at Universities. In A. Doiz, D. Lasagabaster & J. Sierra (Ed.), *English-Medium Instruction at Universities: Global Challenges* (pp. 196-210). Bristol, Blue Ridge Summit: Multilingual Matters. https://doi.org/10.21832/9781847698162-014

Shohamy, E. G. (2001).The power of tests: A critical perspective on the uses of language tests

Sim, S. M., & Rasiah, R. I. (2006). Relationship between item difficulty and discrimination indices in true/false-type multiple choice questions of a para-clinical multidisciplinary paper.

Singapore Ministry of Education. (2007). Summative assessment report. Publisher: Singapore Ministry of Education.*ademy of Medicine Singapore*, *35*(2), 67.

Sjøberg, S. (2015). OECD, PISA, and globalization: The influence of the international assessment regime. In C. H. Tienken & C. A. Mullen (Eds.), Education policy perils. Tackling the tough issues (pp. 102–133). New York, NY: Routledge

Skwarchuk, S.-L. (2005). Formal and informal home learning activities in relation to children's early numeracy and literacy skills: The development of a home numeracy model. Journal of Experimental Child Psychology, 121, 63-84. https://doi.org/10.1016/j.jecp.2013.11.006

Smith, M. L., & Fey, P. (2000).Validity and accountability of highstakes testing. Journal of Teacher Education, 51, 334-344. http://dx.doi.org/10.1177/0022487100051005002

Stevenson, A., & Waite, M. (Eds.). (2011). Concise Oxford English Dictionary. (20th Ed.). New York, NY: Oxford University Press.

Steyer, R., Schmitt, M., & Eid, M. (1999). Latent state–trait theory and research in personality and individual differences. *European Journal of Personality, 13*(5, Spec Issue), 389–408. https://doi.org/10.1002/(SICI)1099-0984(199909/10)13:5<389::AID-PER361>3.0.CO;2-A

Stiggins, R. (2005). From Formative Assessment to Assessment for Learning: A Path to Success in StandardsBased Schools. Phi Delta Kappan, 87, 324-328.

Stiggins, R. J. (2001). Student-involved classroom assessment (3rd ed.). Upper Saddle River, NJ: Prentice Hall.

Stiggins, R. J. (2001). The unful□lled promise of classroom assessment. Educational Measurement: Issues and Practice , 20(3), 5–15.

Stiggins, R. J. (2005). From formative assessment to assessment for learning: A path to success in standards-based schools. Phi Delta Kappan, 87(4), 324-328. DOI: 10.1177/003172170508700414.

Suen, H. K. (1990). Principles of Test Theories. New York: Routledge. https://doi.org/10.4324/9780203052211

Swan, M. & Burkhardt, H., (2012). Designing assessment of performance in mathematics. *MARS: Mathematics Assessment Resource Service, Shell Centre Team, University of Nottingham.*

Syyeda, F. (2016). Understanding Attitudes towards Mathematics (ATM) using a Multimodal modal Model: An Exploratory Case Study with Secondary School Children in England. Cambridge Open-Review Educational Research e-Journal, 3, 32-62. Retrieved from http://corerj.soc.srcf.net/?page_id=224

Syyeda, F. (2016). Understanding Attitudes Towards Mathematics (ATM) using a Multimodal modal Model: An Exploratory Case Study with Secondary School Children in England. Cambridge Open-Review Educational Research e-Journal, 3, 32-62. Retrieved from http://corerj.soc.srcf.net/?page_id=224

Taras, M. (2005). Assessment—summative and formative—some theoretical reflections. British Journal of Educational Studies, 53(4), 466-478.

teacher accountability, and quality classroom instruction. The Qualitative Report, TES (2019). What is formative assessment? A definition. The Times Educational Supplement. Tes News Online edition. Accessed on 15th February 2021 at: https://www.tes.com/news/pedagogy-focus-formative-assessment

The impact of assessment on student learning Shihab Jimaa / Procedia - Social and Behavioral Sciences 28 (2011) 718 – 721

Thorndike, R. M. , Cunningham, G. K. , Thorndike, R. L. , & Hagen, E. P. (1991). *Measurement and evaluation in psychology and education* (5th ed.). New York, NY: MacMillan. [Google Scholar]

Tomlinson, C. A., & Moon, T. R. (2013). Assessment and student success in a differentiated classroom. Alexandria, VA: Association for Supervision and Curriculum Development.

Tosuncuoglu, I. (2018). Importance of Assessment in ELT. Journal of Education and Training Studies, 6, 163-167. https://doi.org/10.11114/jets.v6i9.3443

Tshabalala, T., & Ncube, A. C. (2016). Causes of poor performance of ordinary level pupils in mathematics in rural secondary schools in Nkayi district: Learner's attributions. Nova Journal of Medical and Biological Sciences, 1(1). Retrieved from http://nova-jmbs.com/index.php/jmbs/article/view/13

Ukwuije, R.P.I. and Opkara, I.M. (2013) School Based Assessment-Implication for Educational Transformation. Nigerian Journal of Educational Research and Evaluation, 12, 1-8.

UNESCO Bangkok. (2017). What Can National Achievement Testing 2017 Data Tell Us? Student Performance in Grades 4 and 8. https://bangkok.unesco.org/content/what-can-national-achievement-testing-2017-data-tell-us-student-performance-grades-4-and-8

UNESCO International Bureau of Education. (2007). Inclusive student assessment (Document code: IBE/2007/OP/CD/03).

UNESCO. (2016). Large Scale Assessment in the Asia-Pacific: Policies, Practices, and Trends. Retrieved from https://unesdoc.unesco.org/ark:/48223/pf0000375107/PDF/375107eng.pdf.multi

Valencia, R. R., & Villarreal, B. J. (2003). Improving students' reading performance via standards-based school reform: A critique. The Reading Teacher, 56, 612-621.

Van den Heever, E. (2016). The Impact of Assessment of National Achievement on Schools in South Africa.

Van der Walt, S., Maree, J. G., Ellis, S. M., & Phatudi-Makola, M. (2018). Teacher Perspectives on the Impact of National Assessment of Educational Progress (ANA) Data: A South African Case Study

Van der Walt, S., Maree, J. G., Ellis, S. M., & Phatudi-Makola, M. (2018). Teacher Perspectives on the Impact of National Assessment of Educational Progress (ANA) Data: A South African Case Study.

Volante, L. (2007). Educational quality and accountability in Ontario: Past, present, and future. Canadian Journal of Educational Administration and Policy, 58. Retrieved from https://umanitoba.ca/publications/cjeap/articles/ volante_educational%20_quality.html

Volante, L., & Cherubini, L. (2007). Connecting educational leadership with multi-level assessment reform. International Electronic Journal for Leadership in Learning. Retrieved from Corpus ID: 151267079.

Walkington, C. A. (2013). Using adaptive learning technologies to personalize instruction to student interests: The impact of relevant contexts on performance and learning outcomes. *Journal of Educational Psychology, 105*(4), 932–945. https://doi.org/10.1037/a0031882

Webb, Norman. (2002). Assessment Literacy in a Standards-Based Education Setting.

Weir, J. (2005). Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. Journal of Strength and Conditioning Research / National Strength & Conditioning Association, 19, 231-240. https://doi.org/10.1519/15184.1

Whiston, S. C. (2016). *Principles and applications of assessment in counseling*. Cengage Learning.

Wiberg, M. (2004). CLASSICAL TEST THEORY vs. ITEM RESPONSE THEORY: An evaluation of the theory test in the Swedish driving-license test. Psychology.

Williamson, C. (2017). Teachers' role in school-based assessment as part of public examinations. US-China Education Review B, 7(6), 301-307. https://doi.org/10.17265/2161-6248/2017.06.005

Wojtczak, A. (2002). Glossary of assessment terms. The Association for Medical Education in Europe (AMEE).

Young, C., McNamara, G., Brown, M., & O'Hara, J. (2018). Adopting and adapting: school leaders in the age of data-informed decision making. *Educational Assessment, Evaluation and Accountability*, *30*, 133 - 158.

Zhao, Y. (2018a). Personalizable education for greatness. Kappa Delta Pi Record, 54, 109–115.

Zhao, Y. (2018c). Shifting the education paradigm: Why international borrowing is no longer sufficient for improving education in China. ECNU Review of Education, 1, 76–106

Zhao, Y. (2018d). What works may hurt: Side effects in education. New York, NY: Teachers College Press

Annex-I

NEAS School ID | | | | | | **Roll No** | | |

Achievement Test for Grade 8
Subject: Mathematics

Name of Student ______________________ Boy ☐ Girl ☐

Father's Name ______________________

Name and Address of School ______________________

Date ______________

This test has been prepared according to the prescribed National Curriculum 2006 for Mathematics of Grade 8. It contains topics such as sets, real numbers, number system, square root, algebra, geometry, measurement and information handling. Some instructions, sample and exercise questions have been given in the beginning to enable the students solve the test more conveniently.

INSTRUCTIONS FOR STUDENTS

(i) Read the questions carefully and answer them as per instructions.

(ii) Do not spend too much time on any question. Solve difficult questions at the end.

(iii) If you mark a wrong option, erase it with rubber, then tick the answer you think is correct.

(iv) For some questions the option are not given, Solve these questions yourself. One of the examples is given in the exercise question No.3. Use compass, protractor and scale if required while solving these questions.

(v) Do the rough work in the blank space given infront of every question in the test booklet.

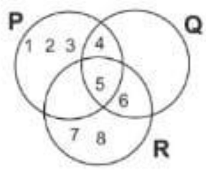
(vi) Mark only one option as a correct answer for a question.

(vii) Calculators, mobile phones or any other textual material is not allowed during test.

Mathematics
Sample Question

INSTRUCTIONS: The following question has four options. Only one of these options is correct answer. The correct answer of this question is {4,5,6,7,8} therefore, box $\boxed{C}$ is ticked "✓". In case of next questions tick the correct boxes.

Question No 1. If



For $(P \cap Q) \cup R$, the correct answer is:

$\boxed{A}$ { 4,5 }

$\boxed{B}$ { 4,5,6 }

$\boxed{C}$ ✓ { 4,5,6,7,8 }

$\boxed{D}$ { 1,2,3,4,5 }

Exercise Question No. 1

Question. The cube root of 4574.296 is

$\boxed{A}$ 68.5

$\boxed{B}$ 67.6

$\boxed{C}$ 21.4

$\boxed{D}$ 16.6

Exercise Question No. 2

Question According to $\triangle ABC$

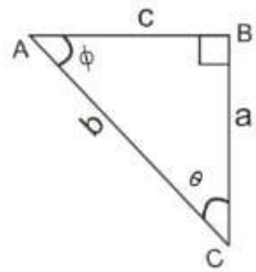
$\cot\theta =$

[A] $\frac{b}{c}$

[B] $\frac{a}{b}$

[C] $\frac{a}{c}$

[D] $\frac{c}{b}$



Exercise Question No. 3

Question. The shade of a certain tree is 30m away from a place. The angle of the peak of the tree from that place is 45°. Find the height of the tree.
(Where as $\tan 45^\circ = 1$, $\sin 45^\circ = \frac{1}{\sqrt{2}}$, $\cos 45^\circ = \frac{1}{\sqrt{2}}$)

Solve:

$$\tan 45^\circ = \frac{Prep}{base}$$

$$1 = \frac{Prep}{30}$$

$$1 \times 30 = Prep$$

Therefore Height of tree = 30m

Ans. 30m

Answer: Ex Q. No. 1 [D] Ex Q. No. 2 [C] Ex Q. No. 3 **Solved**

Note: You have only "2" hours to solve this test

Mathematics Test

INSTRUCTIONS: Four options have been given for each question. Among these, only one option is correct. Read the questions carefully and tick "✓" the correct option A, B, C and D

Q No 1. Which of the following is correct for Base 5 system?

A 0,1,2,3,4,5,6,7,8,9

B 0,1,2,3,4,5,6,7

C 0,1,2,3,4

D 0,1

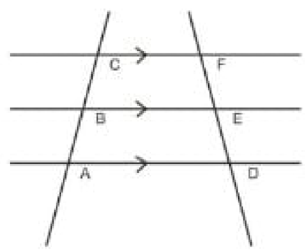
Q No 2. If AB = BC then the true statement about the given figure is

A $DE \neq EF$

B $DE = EF$

C $DE > EF$

D $DE < EF$



Q No 3. Factorize:

$$18x^2-15y^2+24z^2 =$$

A $3(6x^2-5y^2+8z^2)$

B $6(3x^2-15y^2+4z^2)$

C $3(6x^2-15y^2+8z^2)$

D $3(6x^2-5y^2+24z^2)$

Q No 4. The square of 19 in the following is:

A 361

B 19

C 10

D 38

Q No 5. 9.63215... is which decimal fraction?

A Recurring

B Terminating

C Rational

D Non-terminating and non recurring

Q No 6. If the length of one side of regular hexagon is given after drawing the line of the given length the second step will be:

A With the help of ruler, two parallel lines are drawn of the given line.

B With the help of protector, an angle of 120^0 is drawn at one end of the given line.

C Double length of the given line and parallel to the given line is drawn

D With the help of protector, an angle of 140^0 is drawn at the both end of the given line.

Q No 7. In 3,5,9,2,3,5,2,8,7,3,1 the mode is:

A 1

B 3

C 5

D 8

Q No 8. A square shape garden has areas 0.64 square kilometers. What will be the length of each side?

A 0.16 km

B 0.08 km

C 0.8 km

D 8.0 km

Q No 9. Which of the following is the algebraic expression that has two variables?

A $x^6 - 1$

B $x^4 - 4x^2 + 4$

C $x^3 - 3xyz + y^2$

D $x^2 - 2xy + y^2$

Q No 10. $3x+5y-9=0$, $6y=7x-2$

Which of the following is the correct answer for the above equations?

A Linear equations having two variables

B Quadratic equations having two variables

C Cubic equations having two variables

D Degree six equations having two variables

Q No 11. If $A = \{3,7,9,10\}$, $B = \{2,3,4,6\}$, $C = \{1,2,3,5\}$

and $A\cup(B\cap C) = (A\cup B)\cap(A\cup C)$

then the correct answer for the above statement is

A $\{2,3\}$

B $\{2,3,7,9,10\}$

C $\{2,3,4,6,7,9,10\}$

D $\{1,2,3,5,7,9,10\}$

Q No 12. If $a = 7cm$, $b = 8cm$ and $c = 7cm$ are the sides of a triangle ABC, then the area of traingle ABC will be

A $92\ cm^2$

B $22.97\ cm^2$

C $11\ cm^2$

D $6.9\ cm^2$

Q No 13. The square root of 3136 is

A 36

B 46

C 56

D 66

Q No 14. If $mn - x = 0$ and $x - 2pq = 0$

Which of the following is true if x is eliminated from the above equations?

A $mn = 0$

B $2pq = 0$

C $mn + 2pq = 0$

D $mn - 2pq = 0$

Q No 15. How many angles and sides are there in a regular octagon?

A 8 sides and 4 angles.

B 6 sides and 8 angles.

C 6 sides and 6 angles.

D 8 sides and 8 angles.

Q No 16. How is $(1226)_{10}$ written in base 5 system?

- A $(12211)_5$
- B $(14401)_5$
- C $(10441)_5$
- D $(10414)_5$

Q No 17. A line that passes through two points of a circle is called:

- A Perpendicular line
- B Adjacent line
- C Intersecting line
- D Parallel line

Q No 18. In terms of tax, which of the following statements is correct for exemption?

- A Tax payer should not be taxed at all.
- B Impossing tax on tax payer is not right.
- C Tax should not be deducted from tax payer's income upto a specific limit.
- D Tax payer should fix a limit and he/she should not pay tax upto that limit.

Q No 19. Solve:

$$(2x - 6)(x + 11) =$$

A $x^2 - 66x + 16$

B $x^2 - 16x + 11$

C $2x^2 + 16x - 66$

D $2x^2 + 11x - 66$

Q No 20. Hamid sold out Khalid,s car to Javaid for Rs.150,000. Hamid got 5% commission from both of them. The total amount of rupees as commission he got is:

A 157,500

B 150,000

C 15,000

D 7,500

Q No 21. The sum of $(3224)_5$ and $(1302)_2$ is

A $(10031)_5$

B $(4531)_5$

C $(4526)_5$

D $(4031)_5$

Q No 22. Which of the following statements is true?

A $(A \cup B)' = A' \cup B'$

B $(A \cap B)' = A' \cap B'$

C $(A \cap B)' = A' \cup B'$

D $(B' \cap A) = A' \cup B'$

Q No 23. θ and $\emptyset$ are two different acute angles in a right angled triangle. which statement is true for?

$\sin\theta =$

A $\cot\theta$

B $\cot\emptyset$

C $\cos\theta$

D $\cos\emptyset$

Q No 24. In demonstrative geometery the geometrical problems are proved with

A reasons

B capmass

C scale and compass

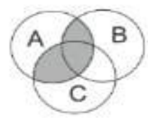
D scale, protractor and compass

[redacted]

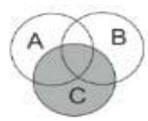
Q No 25. From the following Venn Diagrams, which of the shaded areas represents

$$A\cap(B\cup C) = (A\cap B)\cup(A\cap C)$$

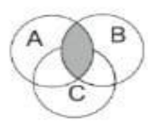
A



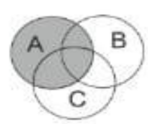
B



C



D



Q No 26(a). Look at this figure and then select the correct answer for:

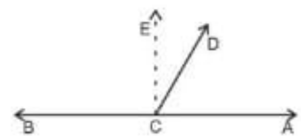
$m\angle ACD =$

A $m\angle BCE + m\angle ECD$

B $m\angle DCB + m\angle BCE$

C $m\angle ACE - m\angle DCE$

D $m\angle ECD + m\angle DCA$



Q No 26(b). $m\angle ACD + m\angle DCB =$

A $m\angle BCE + m\angle ECA$

B $m\angle BCE + m\angle ECD$

C $m\angle ACE + m\angle ACD$

D $m\angle BCE + m\angle ACD$

Q No 26(c). $m\angle DCB =$

A $m\angle ACD + m\angle DCE$

B $m\angle BCE + m\angle ECD$

C $m\angle BCD - m\angle DCE$

D $m\angle BCD - m\angle ACD$

Achievement Test for Grade 4

General Science

Name of Student ____________________________ Boy ☐ Girl ☐

Father's Name ____

Name and Address of School ___________________________________

____

Date: ___________________

This test is set according to the prescribed National Curriculum. It contains on Geology, Biology, Astrology, and Physics. Clear instructions are given before each question. Besides this, general directions are also provided. One model and three exercise questions are done for you to facilitate you solve the test in a proper way.

INSTRUCTIONS FOR THE STUDENTS

1. Read each question carefully and tick ✓ the right answer out of four options. In other case, write relevant answer in the space given below each question. Examples are present on next page.
2. If you are not sure about the answer to a question, tick the box next to the answer you think is best.
3. Don't spend much time on any specific question solve the easy one earlier and then use the remaining time to solve complex questions.
4. If you mark a wrong option, erase it with rubber, then tick the answer you think is correct.

Dictionary, text book, cell phones are not allowed at the Centre.

General Science

Example Question

INSTRUCTION: The following question has four options. One of these options is correct answer. The correct answer of this question is "Flower", therefore, box B is "✓" ticked. In case of the next questions, tick the correct box. For some questions, you will write your answer, in the space below the question with the help of words and pictures.

Q: Which part of a plant produces seed?

- A Leaf
- B ✓ Flower
- C Stem
- D Root

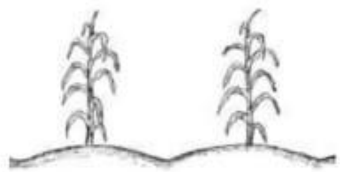
Practice Question No: 1

Q: Energy from the sun can be changed into electricity for people to use. Write one other source of energy which can be changed into electricity.

Answer: ______________________________

Practice Question No: 2

Q: Haris has a small vegetable garden?



Which natural resources does Haris use to grow plants?

- A water and wind
- B water and soil
- C soil and wind
- D oil and wind

Practice Question No: 3

Q: Zeeshan stirs a pot of boiling soup and leaves her spoon in the pot.

Later, the spoon is too hot to pick up.

What material is the spoon most likely made from ?

- A wood
- B rubber
- C plastic
- D metal

DO NOT TURN OVER THE PAGE AND LISTEN TO NEXT INSTRUCTIONS CAREFULLY.

Answers: Q.1 Water, dung etc Q.2 $\boxed{B}$ Q.3 $\boxed{D}$

YOU HAVE TWO HOURS TO COMPLETE THIS TEST.

General Science Test

INSTRUCTIONS: Four options have been given for each question. Among these, only one option is correct. Read the questions carefully and tick '✓' the correct option $\boxed{A}$, $\boxed{B}$, $\boxed{C}$ and $\boxed{D}$.

Q No. 1 Which, among the following objects is the heaviest?

$\boxed{A}$ car

$\boxed{B}$ bus

$\boxed{C}$ pen

$\boxed{D}$ ball

Q No. 2 What is required to animals for their survival?

$\boxed{A}$ soil

$\boxed{B}$ smoke

$\boxed{C}$ food

$\boxed{D}$ stone

Q No.3 Yousaf covers 50 meter distance in 10 seconds, and Maryam covers the same distance in 20 seconds. Who is fast?

- A Yousaf
- B Maryam
- C both
- D none of the above

Q No. 4 A balanced diet contains:

- A meat, pulses, vegetables, fruit
- B pulses, meat
- C pulses, vegetables
- D vegetables, meat

Q No. 5 Which of the following is dissolve in water?

- A salt
- B iron fillings
- C rubber
- D wood particals

Q No. 6 Smoke of vehicles pollutes:

- A air
- B food
- C water
- D soil

Q No. 7 Which of the following seasons the days are short and nights are longest?

- A autumn
- B spring
- C winter
- D summer

Q No.8 The total number of bones in human body is

- A 210
- B 206
- C 200
- D 220

Q No.9 A parrot is not visible in green trees because its:

- [A] beak is red
- [B] feathers are green
- [C] sits on tall trees
- [D] eats green chili fondly

Q No.10 If a magnet is cut into half, its number of poles of each part will be:

- [A] one
- [B] two
- [C] three
- [D] four

Q No.11 Excess use of salt in food can cause the disease of:

- [A] diabatese
- [B] high blood presure
- [C] tumer
- [D] skin diseases

Q No.12 A box contains a mixture of iron and sand. Tell the method by which iron can be separated from sand:

- [A] sand
- [B] salt
- [C] wood
- [D] stone

Q No.13 Which of the following is a noise?

- [A] piano
- [B] flute
- [C] horn
- [D] guitar

Q No.14 Which of the following groups of food contains carbohydrates?

- [A] wheat-rice
- [B] water-sweets
- [C] pulses-fruit
- [D] carrot-apple

Q No.15 Which of the following minerals present in milk?

A calcium

B sulpher

C carbon

D nitrogen

Q No.16 Which food eaten by herbivores animal?

A meat

B bones

C leaves and grass

D fish

Q No.17 At what temperature the plants and animals die?

A 300°C

B 400°C

C extreme temperature

D 100°F

Q No.18 Different organisms are shown in the following scene which organism makes its own food?

A duck

B frog

C man

D trees

Q No.19 The sunlight is necessary for plants because in its presence plants:

A prepare chlorophyll

B do not grow

C do not provide fruit

D do not prepare food

Q No.20 Open circuit is that circuit where the current:

A stops

B flows inereares

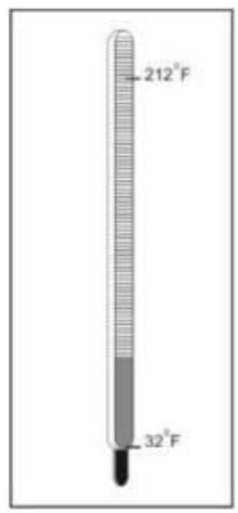
C becomes smooth

D slows down

Q No.21 The earth takes 24 hours to complete its rotation. This shows:

- A the time of the sun rise
- B distance between the earth and the sun
- C charge of seasons
- D duration of day and night

Q No.22 What is the name of the temperature scale given in the diagram?



- A celcius
- B fahrenheit
- C kelvin
- D none of the above

Q No.23 How do the earth, the moon and the sun make movements?

A the moon revolves around the earth and the earth around the sun

B the moon revolves around the sun and the earth around the moon

C the sun and the moon revolve around the earth

D the sun revolves around the moon and the moon around the earth

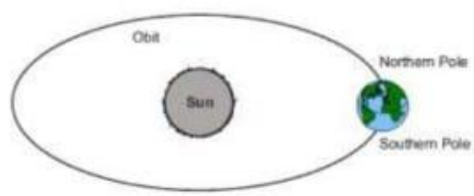
Q No.24 In figure earth is revolving around the sun in its own axis. The reason on South Pole is:

A winter

B summer

C spring

D autumn



Q No.25 What is produced if drum is beaten with stick?

- A light
- B sound
- C electrical energy
- D magnetic force

Q No.26 Wahaj put a thermometer in a beaker filled with water. He notices that the level of mercury in thermometer rises. It is because mercury:

- A pulls the pressure on water up
- B increases due to gravity
- C expands due to heat of water
- D expands due to room temperature

Q No.27 If the Earth takes the place of Uranus then the duration of a year will:

- A increase
- B decrease
- C remain same
- D change with season

Q No.28 Which of the following statements is true about stomata?

- A small pores in leaf
- B omit gases
- C provide oxygen only to plant
- D provide oxygen only to animals

Q No.29 If a person covers 48 kilometers in 6 seconds, his speed will be:

- A 16 m/sec
- B 12 m/sec
- C 10 m/sec
- D 8 m/sec

Q No.30 A body pressed air-filled balloon under his hands. It will come in its original form by.

- A continuous pull
- B removing
- C tilting the ballon
- D bursting of balloon

Q No.31 Which of the following characteristics has Sana inherited from her parents?

- A favourite food of mother
- B favourite pink colour of mother
- C favourite blue colour of father
- D father's skin colour

Q No.32 The Colour of mercury in thermometer is

- A green
- B silver
- C purple
- D blue

Q No.33 Which one of the following comes back to its original form after removing the applied force?

- A clay
- B sponge
- C block of wood
- D glass

Q No.34 Sound **cannot** pass through:

A glass

B oil

C oxygen

D vacuum

Q No.35 In extreme hot, the tiers of vehicles burst, because air in the tube

A changes into solid

B changes into liquid

C increases in volume due to heat

D changes into vapours

Q No.36 Complete the blank boxes with proper answer in the given table.

Animal	Habital Air	Movement	Respiration
Fish			
Bird			

Q No.37 The stages of life cycle of a plant and a frog are given below. Read them carefully and then associate them with the life cycle of plant and frog.

flower, seed, growth, sewing seed, tadpole, plant, egg, frog.

Plant: ______________________________________

Frog: ______________________________________

Q. No.38 What is required to plants for their growth?

Ans.

____

____

____

____

Q No.39 Name any two parts of human body which are jointed.

1. ______________________ 2. ____________________

Q No.40 Write the names of parts of human head.

Ans. ____

____

Annex -II

TEACHER QUESTIONNAIRE

The "**LARGE SCALE ASSESSMENT IN PAKISTAN: A CRITICAL ANALYSIS OF DEVELOPMENT AND APPLICATION OF NATIONAL ACHIEVEMENT Test** is the research study you are requested to share your experience/views through participation in this study. Your responses would be kept anonymous and used only for research purposes. Your assistance in the study would be highly appreciated.

1. Name ....

2. Please indicate your gender

Male ...... Female ...

2. Institution ....

3. Area/Province

AJK ......ICT......

4. Subject: _______________

5. Qualification: Masters in Islamic studies _______________

6. Experience of development of assessment ....

7. Experience of application of assessment activities ....

INSTRUCTIONS FOR ANSWERING SURVEY ITEMS:

Please state your level of agreement for the following statements regarding "National Achievement Test (NAT)". You will use a five-point scale to indicate the degree to which you feel is appropriate. The scale ranges from: 1 = strongly agree to 5 = strongly disagree

For example, using the five-point scale, how appropriate do you feel the following purpose is for the tests and/or examinations that you are going to talk about?

1. *Improve the achievement of all students.*

Strongly Agree	*Agree*	*Neutral*	*Disagree*	*Strongly Disagree*
1	*2*	*3*	*4*	*5*

Statement	**1 Strongly agree**	**2 Agree**	**3 Neutral**	**4 Disagree**	**5 Strongly disagree**
Do you think National Achievement Test					
improves the high academic standards					
improves the students achievement					
improves the teacher's effectiveness					
perform as an accountability tool					
administration is too early to get accurate achievement data					

Do you think that NAT monitor the progress of student by:					
Statement	**1 Strongly agree**	**2 Agree**	**3 Neutral**	**4 Disagree**	**5 Strongly disagree**
measuring student achievement					
aligned with classroom assessment instruments that reflect item format and content in NAT.					

determining how well the students are learning the intended curriculum					
providing a "snap-shot" of what can do.					
helping students to concentrate on their effort in learning					

Do you think that national achievement test impact the instructional methodology by					
Statement	**1 Strongly agree**	**2 Agree**	**3 Neutral**	**4 Disagree**	**5 Strongly disagree**
Increasing teachers' assessment, knowledge and skills					
Improving and enhancing teaching					
providing common measures so that teachers can link their own assessments to national standards					
evaluating the quality of students					
evaluating the quality of teachers					
evaluating the quality of schools					

Being a teacher, do you think that National Achievement Test					
Statement	**1 Strongly agree**	**2 Agree**	**3 Neutral**	**4 Disagree**	**5 Strongly disagree**

Opt the national assessment framework for assessment					
Provide data for data-based decision-making at student level					
Provide data for data-based decision-making at class level					
Provide data for data-based decision-making at school level					

Being a teacher, do you think that you/your					
Statement	**1 Strongly agree**	**2 Agree**	**3 Neutral**	**4 Disagree**	**5 Strongly disagree**
know how to interpret assessment results.					
know how to use the test results					
principals know how to interpret the test results					
public know how to interpret the test results					

Instruction: Please read the questions and tick your answer.

Q1- Do you actually use the results of NAT being a teacher? On the basis of your experience

Yes No

Q2. Do you feel that National Achievement test have value?

- In evaluating the quality of students
- In evaluating the quality of teachers

- In evaluating the quality of schools

Q3. One of the purposes of NAT is to measure the functionality of the education system. Do you think this is a good indicator of the education system? Why or why not?

....
....
....
....

Q4. Do you think the standards in NAT align with the standards used in classroom instructive assessments?

....
....
....
....

Q5 What are the factors influencing poor student achievement in NAT?

....
....
....

Q6 Are you satisfied with NAT

Questionnaire Design

- Yes) No

Content Coverage

(Yes) No

Assessment Time

- (Yes) No

Q7 Is there any mechanism for data dissemination of NAT result at your school level?

- Yes (No)

Q8. What do you think are the challenges that you/teachers are facing in application of NAT in an institute?

....
....

....
....
....

Q9- Please suggests .Best ways to effectively applying NAT results as a Teacher?

....
....
....
....
....

Q10. What are the resources you think that teachers and school administers need most to help improve student performance in NAT?

....
....
....
....
....

ANNEX- III
CHECKLIST FOR REPORTING AND DISSEMINATION OF NAT

Checklist Item	Target Audience	Description	Scale
Report Development			
Develop group-level performance reports for teachers	Teachers	Create detailed reports that provide insights into student performance at a group level.	Not Started / Ongoing / Implemented
Simplify overall performance reports for parents	Parents	Prepare simplified reports that give an overview of their child's performance.	Not Started / Ongoing / Implemented
Aggregate data for community stakeholders	Community Organizations/NGOs	Provide aggregated data to identify educational gaps and prioritize interventions.	Not Started / Ongoing / Implemented
Communication Channels			
Develop user-friendly online platforms	All Stakeholders	Create accessible online platforms for stakeholders to easily access and interpret data.	Not Started / Ongoing / Implemented
Organize workshops for data usage	All Stakeholders	Conduct workshops to teach stakeholders how to effectively use and interpret the data provided.	Not Started / Ongoing / Implemented
Training and Support			
Hold training sessions to enhance data interpretation skills	Teachers and Community Stakeholders	Provide training sessions to help stakeholders understand and utilize data effectively.	Not Started / Ongoing / Implemented
Create data summaries and infographics for wider dissemination	All Stakeholders	Develop and distribute data summaries and infographics to make data more accessible and understandable to a wider audience.	Not Started / Ongoing / Implemented

ANNEX- IV

F.No 5(1)/NCC/2021
Government of Pakistan
Ministry of Federal Education & Professional Training
National Curriculum Council, Secretariat (NCC)

D.No: 5047
Date: 10-11-2021

Subject: **REQUESTING DATA ACCESS FOR THE PHD RESEARCH STUDY**

The undersigned is pursuing PhD in Education from National University of Modern Languages, Islamabad. After successful completion of course work, the phase of research work has started. The thesis titled as "Large Scale Assessment in Pakistan: a Critical Analysis of Development and Application of National Achievement Test".

2. The objectives of this study will be:

i. To provide a comprehensive critically analysis psychometric properties of NAT 2016 by using IRT framework"
ii. To generate objective and accurate information about its quality through teacher perceptions those participated in NAT 2016
iii. To propose a model for proper dissemination of data of NAT results to different stakeholder to improve assessment system in Pakistan.

3. In order to work on this topic, I need the following data from National Education Assesemnt System (NEAS):

- Soft data of the questionnaire
- Soft data of student response
- Details of teachers and schools participated in NAT 2016

4. It is requested to forward my request to Additional Secretary (F/A) through a proper channel to have this information from the National Education Assessment System (NEAS)

SAIMA ABBAS
Education officer
9/11/21

AEA (Assessment)

9/11/21

Director (NCC)

5. Kindly request NEAS to provide data.

AS/JS(PT) 6. Please do the needful.

Mohyuddin Ahmad Wani
Additional Secretary
10/11/21.

7. Urgent Pl.
DS(FE)
S.O NCC (AB)
Abbas SO 11.11

Dy.No. 1988 AS(FE&PT)
Dated 11-11-21

Dy.No. 1310 JS(PT)
Dated 11/11/2021

Dy.No. 2043/2021 SO(NCC)
12-11-2021

-1702-

No.F.4(5)/2015-NCC
Government of Pakistan
Ministry of Federal Education
& Professional Training
****

Islamabad: the 15th November 2021.

SUBJECT: <u>REQUESTING DATA ACCESS FOR THE PH.D RESEARCH STUDY.</u>

The undersigned is directed to refer to the above subject and to state that Ms.Saima Abbas, Education Officer, National Curriculum Council (NCC) is currently doing Ph-D in Education from National University of Modern Languages (NUML), Islamabad. After successful completion of course work, the phase of her research work has started. The title of her thesis is "Large Scale Assessment in Pakistan; a Critical Analysis of Development and Application of National Achievement Test". The objectives of this dtudy will be as under:-

i. To provide a comprehensive critically analysis psychometric properties of NAT 2016 by using IRT framework"
ii. To generate objective and accurate information about its quality through teacher preceptions those participated in NAT 2016
iii. To propose a model for proper dissemination of data of NAT results to different stakeholders to improve assessment system in Pakistan.

2. In order to work on this topic, she needs the following data from National Education Assesement System (NEAS):-

- Soft data of the questionnaire
- Soft data of student response
- Details of teachers and schools participanted in NAT 2016

3. It is requested that the above-stated information / documents may please be furnished urgently.

4. This issues with approval of the Additional Secretary, M/o FE&PT.

(Farooque Aizam Abro)
Assistant Educational Adviser
Ph: 051-9205852

National Coordinator,
National Education Assessment System (NEAS),
M/o FE&PT, Islamabad.

Copy to APS to Ms. Saima Abbas. Education Officer (NCC), Islamabad.

ANNEX-V

Seeking permission to use Research Tool Inbox ×

saima Abbas <saima.abbas832@gmail.com> Mar 18, 2022, 4:22 AM
to dklinger, don.klinger, bcc: Aisha

Dear Dr. Don A. Klinger!
Greetings from Pakistan!
I am Saima Abbas (PhD Scholar) from National University for Modern Language (NUML), Islamabad, Pakistan. My research titled, as LARGE **SCALE ASSESSMENT IN PAKISTAN: A CRITICAL ANALYSIS OF DEVELOPMENT AND APPLICATION OF NATIONAL ACHIEVEMENT TEST**.

I am writing this to inform you that recently, I found a questionnaire used in your study exploring Large Scale Assessments in Canada. I want to use your Questionnaire in my research and I need your permission in this regard. Furthermore, I am in search of other literature relevant to my topic and also need some more information about your tool.

Hereby, I would like to use and print your tool under the following conditions:
· I will use this tool only for my research study and will not sell or use it with any compensated or curriculum development activities.
· I will include the copyright statement on all copies of the tool.
· I will send a copy of my completed research study to your attention upon completion of the study. If these are acceptable terms and conditions, please indicate so by replying to me through e-mail: saima.abbas832@gmail.com Contact No-03345213859.
I already emailed you regarding it and I need your immediate response to facilitate my research study.
Thanking in anticipation
Best Regards,

AW01 +0.07% 7:43 PM 5/27/2024

Don Klinger <don.klinger@queensu.ca> Mar 20, 2022, 12:53 PM
to me

Saima,

Yes, I have already sent you a previous email allowing you to use this survey. I am sending this response using a different email address in case the other went into a trash folder.

Nāku noa, nā

Professor Don Albert Klinger
Pro Vice-Chancellor
Te Wānanga Toi Tangata - Division of Education
Ph: +64 7 838 4281 | Mobile: +64 022 013 1957 | https://www.linkedin.com/in/don-klinger-99bb4535/
University of Waikato | Private Bag 3105
Hamilton 3240 | New Zealand
don.klinger@waikato.ac.nz
University of Waikato

The content of this email is confidential and intended for the recipient specified in message only, unless otherwise stated.

saima Abbas <saima.abbas832@gmail.com> Mar 22, 2022, 5:23 AM
to Aisha

AW01 +0.07% 7:42 PM 5/27/2024

ANNEX-VI

LIST OF SCHOOL PARTICIPATED IN NATIONAL ACHIEVEMENT TEST 2016

NEAS_ID	Province	District	Tehsil	Name of Schools	Number of Students
A171	Azad Jammu & Kashmir	Hattian	Hattian	MSQ BANDI LANGLA	7
A172	Azad Jammu & Kashmir	Hattian	Hattian	GGPS MAKNAIT	4
A173	Azad Jammu & Kashmir	Haveli	Mumtaz abad	GBPS SANGAL BALA	9
A174	Azad Jammu & Kashmir	Kotli	Dulliya jattan	GBPS AUMB SAYYADAN	7
A175	Azad Jammu & Kashmir	Mirpur	Mirpur	GGPS SECTOR G-1 MIRPUR	10
A176	Azad Jammu & Kashmir	Muzaffarabad	Muzaffarabad	GBPS KANGRAN	6
A177	Azad Jammu & Kashmir	Muzaffarabad	Pattika	GBMS BATANGI NOORI SERI	7
A178	Azad Jammu & Kashmir	Muzaffarabad	Muzaffarabad	GBPS KAMHAR BANDI GHARBI	8
A179	Azad Jammu & Kashmir	Muzaffarabad	Muzaffarabad	GGPS NARULL SATRHA	19
A180	Azad Jammu & Kashmir	Neelam	Sharda	GGPS JANAWAI	12
A181	Azad Jammu & Kashmir	Neelam	Athmuqam	GBPS GUJAN	6
A182	Azad Jammu & Kashmir	Sudhnoti	Baloch	GGPS KAIRY MINHAS	7
A183	Azad Jammu & Kashmir	Sudhnoti	Pallandari	GGPS KOHAS SIDDIAN	3
A184	Azad Jammu & Kashmir	Sudhnoti	Mong	GGPS KHOUKHRAL	5

ICT193	Islamabad	IMS (I-V) Sawan Garden	40
ICT194	Islamabad	IMSG (I-V), HERDOGHER	70
ICT195	Islamabad	IMSJ (I-V), G-6/2	70
ICT196	Islamabad	IMSJ (I-V), I-8/1	70
ICT197	Islamabad	IMSJ (I-V), F-6-3,	56
ICT198	Islamabad	IMSJ (I-V), F-7/2	70
ICT199	Islamabad	IMSJ (I-V), NO 1, G-6/1-4,	70
ICT200	Islamabad	IMSJ (I-V) NO 1, E-8	70

LIST OF SCHOOL PARTICIPATED IN NATIONAL ACHIEVEMENT TEST 2019

1000101	**1**	**10001**	**A0001**	**1**	**1**	**GGMS BRING BUN**	**2**	**1**
1000102	1	10001	A0001	2	1	GGMS BRING BUN	2	1
1000103	1	10001	A0001	3	1	GGMS BRING BUN	2	1
1000104	1	10001	A0001	4	1	GGMS BRING BUN	2	1
1000105	1	10001	A0001	5	1	GGMS BRING BUN	2	1
1000106	1	10001	A0001	6	1	GGMS BRING BUN	2	1
1000107	1	10001	A0001	7	1	GGMS BRING BUN	2	1
1000108	1	10001	A0001	8	1	GGMS BRING BUN	2	1
1000109	1	10001	A0001	9	1	GGMS BRING BUN	2	1
1000110	1	10001	A0001	10	1	GGMS BRING BUN	2	1
1000111	1	10001	A0001	11	1	GGMS BRING BUN	2	1
1000112	1	10001	A0001	12	1	GGMS BRING BUN	2	1
1000113	1	10001	A0001	13	1	GGMS BRING BUN	2	1
1000114	1	10001	A0001	14	1	GGMS BRING BUN	2	1
1000115	1	10001	A0001	15	1	GGMS BRING BUN	2	1
1000116	1	10001	A0001	16	1	GGMS BRING BUN	2	1

1000117	1	10001	A0001	17	1	GGMS BRING BUN	2	1
1000118	1	10001	A0001	18	1	GGMS BRING BUN	2	1
1000121	1	10001	A0001	21	1	GGMS BRING BUN	2	1
1000122	1	10001	A0001	22	1	GGMS BRING BUN	2	1
1000201	1	10002	A0002	1	1	GGHS KHORI CHANNA	2	1
1000202	1	10002	A0002	2	1	GGHS KHORI CHANNA	2	1
1000203	1	10002	A0002	3	1	GGHS KHORI CHANNA	2	1
1000204	1	10002	A0002	4	1	GGHS KHORI CHANNA	2	1
1000205	1	10002	A0002	5	1	GGHS KHORI CHANNA	2	1
1000206	1	10002	A0002	6	1	GGHS KHORI CHANNA	2	1
1000207	1	10002	A0002	7	1	GGHS KHORI CHANNA	2	1
1000208	1	10002	A0002	8	1	GGHS KHORI CHANNA	2	1
1000209	1	10002	A0002	9	1	GGHS KHORI CHANNA	2	1
1000210	1	10002	A0002	10	1	GGHS KHORI CHANNA	2	1
1000211	1	10002	A0002	11	1	GGHS KHORI CHANNA	2	1
1000221	1	10002	A0002	21	1	GGHS KHORI CHANNA	2	1
1000222	1	10002	A0002	22	1	GGHS KHORI CHANNA	2	1
1000301	1	10003	A0003	1	1	GBHS RATNOI	1	1
1000302	1	10003	A0003	2	1	GBHS RATNOI	1	1
1000303	1	10003	A0003	3	1	GBHS RATNOI	1	1
1000304	1	10003	A0003	4	1	GBHS RATNOI	1	1
1000305	1	10003	A0003	5	1	GBHS RATNOI	1	1

1000306	1	10003	A0003	6	1	GBHS RATNOI	1	1
1000307	1	10003	A0003	7	1	GBHS RATNOI	1	1
1000308	1	10003	A0003	8	1	GBHS RATNOI	1	1
1000309	1	10003	A0003	9	1	GBHS RATNOI	1	1
1000310	1	10003	A0003	10	1	GBHS RATNOI	1	1
1000311	1	10003	A0003	11	1	GBHS RATNOI	1	1
1000312	1	10003	A0003	12	1	GBHS RATNOI	1	1
1000313	1	10003	A0003	13	1	GBHS RATNOI	1	1
1000314	1	10003	A0003	14	1	GBHS RATNOI	1	1
1000315	1	10003	A0003	15	1	GBHS RATNOI	1	1
1000316	1	10003	A0003	16	1	GBHS RATNOI	1	1
1000317	1	10003	A0003	17	1	GBHS RATNOI	1	1
1000318	1	10003	A0003	18	1	GBHS RATNOI	1	1
1000319	1	10003	A0003	19	1	GBHS RATNOI	1	1
1000320	1	10003	A0003	20	1	GBHS RATNOI	1	1
1000321	1	10003	A0003	21	1	GBHS RATNOI	1	1
1000322	1	10003	A0003	22	1	GBHS RATNOI	1	1
1000401	1	10004	A0004	1	1	GBHS BANI MAHALDARA	1	1
1000402	1	10004	A0004	2	1	GBHS BANI MAHALDARA	1	1
1000403	1	10004	A0004	3	1	GBHS BANI MAHALDARA	1	1
1000404	1	10004	A0004	4	1	GBHS BANI MAHALDARA	1	1
1000405	1	10004	A0004	5	1	GBHS BANI MAHALDARA	1	1

1000406	1	10004	A0004	6	1	GBHS BANI MAHALDARA	1	1
1000407	1	10004	A0004	7	1	GBHS BANI MAHALDARA	1	1
1000408	1	10004	A0004	8	1	GBHS BANI MAHALDARA	1	1
1000409	1	10004	A0004	9	1	GBHS BANI MAHALDARA	1	1
1000410	1	10004	A0004	10	1	GBHS BANI MAHALDARA	1	1
1000411	1	10004	A0004	11	1	GBHS BANI MAHALDARA	1	1
1000412	1	10004	A0004	12	1	GBHS BANI MAHALDARA	1	1
1000413	1	10004	A0004	13	1	GBHS BANI MAHALDARA	1	1
1000414	1	10004	A0004	14	1	GBHS BANI MAHALDARA	1	1
1000415	1	10004	A0004	15	1	GBHS BANI MAHALDARA	1	1
1000416	1	10004	A0004	16	1	GBHS BANI MAHALDARA	1	1
1000417	1	10004	A0004	17	1	GBHS BANI MAHALDARA	1	1
1000418	1	10004	A0004	18	1	GBHS BANI MAHALDARA	1	1
1000419	1	10004	A0004	19	1	GBHS BANI MAHALDARA	1	1
1000420	1	10004	A0004	20	1	GBHS BANI MAHALDARA	1	1
1000501	1	10005	A0005	1	1	GGPS KEEMER	2	1
1000502	1	10005	A0005	2	1	GGPS KEEMER	2	1
1000503	1	10005	A0005	3	1	GGPS KEEMER	2	1
1000504	1	10005	A0005	4	1	GGPS KEEMER	2	1
1000505	1	10005	A0005	5	1	GGPS KEEMER	2	1
1000506	1	10005	A0005	6	1	GGPS KEEMER	2	1
1000507	1	10005	A0005	7	1	GGPS KEEMER	2	1

1000508	1	10005	A0005	8	1	GGPS KEEMER	2	1
1000509	1	10005	A0005	9	1	GGPS KEEMER	2	1
1000521	1	10005	A0005	21	1	GGPS KEEMER	2	1
1000522	1	10005	A0005	22	1	GGPS KEEMER	2	1
1000601	1	10006	A0006	1	1	GBMS NALIAN	1	1
1000602	1	10006	A0006	2	1	GBMS NALIAN	1	1
1000603	1	10006	A0006	3	1	GBMS NALIAN	1	1
1000604	1	10006	A0006	4	1	GBMS NALIAN	1	1
1000605	1	10006	A0006	5	1	GBMS NALIAN	1	1
1000606	1	10006	A0006	6	1	GBMS NALIAN	1	1
1000607	1	10006	A0006	7	1	GBMS NALIAN	1	1
1000608	1	10006	A0006	8	1	GBMS NALIAN	1	1
1000609	1	10006	A0006	9	1	GBMS NALIAN	1	1
1000610	1	10006	A0006	10	1	GBMS NALIAN	1	1
1000611	1	10006	A0006	11	1	GBMS NALIAN	1	1
1000612	1	10006	A0006	12	1	GBMS NALIAN	1	1
1000613	1	10006	A0006	13	1	GBMS NALIAN	1	1
1000614	1	10006	A0006	14	1	GBMS NALIAN	1	1
1000615	1	10006	A0006	15	1	GBMS NALIAN	1	1
1000616	1	10006	A0006	16	1	GBMS NALIAN	1	1
1000617	1	10006	A0006	17	1	GBMS NALIAN	1	1
1000618	1	10006	A0006	18	1	GBMS NALIAN	1	1

1000619	1	10006	A0006	19	1	GBMS NALIAN	1	1
1000620	1	10006	A0006	20	1	GBMS NALIAN	1	1
1000701	1	10007	A0007	1	1	GGHS MALOTE	2	1
1000702	1	10007	A0007	2	1	GGHS MALOTE	2	1
1000703	1	10007	A0007	3	1	GGHS MALOTE	2	1
1000704	1	10007	A0007	4	1	GGHS MALOTE	2	1
1000705	1	10007	A0007	5	1	GGHS MALOTE	2	1
1000706	1	10007	A0007	6	1	GGHS MALOTE	2	1
1000707	1	10007	A0007	7	1	GGHS MALOTE	2	1
1000708	1	10007	A0007	8	1	GGHS MALOTE	2	1
1000709	1	10007	A0007	9	1	GGHS MALOTE	2	1
1000710	1	10007	A0007	10	1	GGHS MALOTE	2	1
1000711	1	10007	A0007	11	1	GGHS MALOTE	2	1
1000712	1	10007	A0007	12	1	GGHS MALOTE	2	1
1000713	1	10007	A0007	13	1	GGHS MALOTE	2	1
1000714	1	10007	A0007	14	1	GGHS MALOTE	2	1
1000715	1	10007	A0007	15	1	GGHS MALOTE	2	1
1000716	1	10007	A0007	16	1	GGHS MALOTE	2	1
1000717	1	10007	A0007	17	1	GGHS MALOTE	2	1
1000718	1	10007	A0007	18	1	GGHS MALOTE	2	1
1000721	1	10007	A0007	21	1	GGHS MALOTE	2	1
1000722	1	10007	A0007	22	1	GGHS MALOTE	2	1

1000801	1	10008	A0008	1	1	GBMS KANGRA SHARQI	1	1
1000802	1	10008	A0008	2	1	GBMS KANGRA SHARQI	1	1
1000803	1	10008	A0008	3	1	GBMS KANGRA SHARQI	1	1
1000804	1	10008	A0008	4	1	GBMS KANGRA SHARQI	1	1
1000805	1	10008	A0008	5	1	GBMS KANGRA SHARQI	1	1
1000806	1	10008	A0008	6	1	GBMS KANGRA SHARQI	1	1
1000807	1	10008	A0008	7	1	GBMS KANGRA SHARQI	1	1
1000808	1	10008	A0008	8	1	GBMS KANGRA SHARQI	1	1
1000809	1	10008	A0008	9	1	GBMS KANGRA SHARQI	1	1
1000810	1	10008	A0008	10	1	GBMS KANGRA SHARQI	1	1
1000811	1	10008	A0008	11	1	GBMS KANGRA SHARQI	1	1
1000812	1	10008	A0008	12	1	GBMS KANGRA SHARQI	1	1
1000813	1	10008	A0008	13	1	GBMS KANGRA SHARQI	1	1
1000814	1	10008	A0008	14	1	GBMS KANGRA SHARQI	1	1
1000815	1	10008	A0008	15	1	GBMS KANGRA SHARQI	1	1
1000816	1	10008	A0008	16	1	GBMS KANGRA SHARQI	1	1
1000817	1	10008	A0008	17	1	GBMS KANGRA SHARQI	1	1
1000818	1	10008	A0008	18	1	GBMS KANGRA SHARQI	1	1
1000819	1	10008	A0008	19	1	GBMS KANGRA SHARQI	1	1
1000901	1	10009	A0009	1	1	GBMS DOBAY	1	1
1000902	1	10009	A0009	2	1	GBMS DOBAY	1	1
1000903	1	10009	A0009	3	1	GBMS DOBAY	1	1

1000904	1	10009	A0009	4	1	GBMS DOBAY	1	1
1000905	1	10009	A0009	5	1	GBMS DOBAY	1	1
1000906	1	10009	A0009	6	1	GBMS DOBAY	1	1
1000907	1	10009	A0009	7	1	GBMS DOBAY	1	1
1000908	1	10009	A0009	8	1	GBMS DOBAY	1	1
1000909	1	10009	A0009	9	1	GBMS DOBAY	1	1
1001001	1	10010	A0010	1	1	GGMS ABOWALA	2	1
1001002	1	10010	A0010	2	1	GGMS ABOWALA	2	1
1001003	1	10010	A0010	3	1	GGMS ABOWALA	2	1
1001004	1	10010	A0010	4	1	GGMS ABOWALA	2	1
1001005	1	10010	A0010	5	1	GGMS ABOWALA	2	1
1001006	1	10010	A0010	6	1	GGMS ABOWALA	2	1
1001007	1	10010	A0010	7	1	GGMS ABOWALA	2	1
1001008	1	10010	A0010	8	1	GGMS ABOWALA	2	1
1001009	1	10010	A0010	9	1	GGMS ABOWALA	2	1
1001010	1	10010	A0010	10	1	GGMS ABOWALA	2	1
1001011	1	10010	A0010	11	1	GGMS ABOWALA	2	1
1001012	1	10010	A0010	12	1	GGMS ABOWALA	2	1
1001013	1	10010	A0010	13	1	GGMS ABOWALA	2	1
1001014	1	10010	A0010	14	1	GGMS ABOWALA	2	1
1001015	1	10010	A0010	15	1	GGMS ABOWALA	2	1
1001016	1	10010	A0010	16	1	GGMS ABOWALA	2	1

1001017	1	10010	A0010	17	1	GGMS ABOWALA	2	1
1001018	1	10010	A0010	18	1	GGMS ABOWALA	2	1
1001019	1	10010	A0010	19	1	GGMS ABOWALA	2	1
1001020	1	10010	A0010	20	1	GGMS ABOWALA	2	1
1001021	1	10010	A0010	21	1	GGMS ABOWALA	2	1
1001022	1	10010	A0010	22	1	GGMS ABOWALA	2	1
1001101	1	10011	A0011	1	1	GBHS MALOTE	1	1
1001102	1	10011	A0011	2	1	GBHS MALOTE	1	1
1001103	1	10011	A0011	3	1	GBHS MALOTE	1	1
1001104	1	10011	A0011	4	1	GBHS MALOTE	1	1
1001105	1	10011	A0011	5	1	GBHS MALOTE	1	1
1001106	1	10011	A0011	6	1	GBHS MALOTE	1	1
1001107	1	10011	A0011	7	1	GBHS MALOTE	1	1
1001108	1	10011	A0011	8	1	GBHS MALOTE	1	1
1001109	1	10011	A0011	9	1	GBHS MALOTE	1	1
1001110	1	10011	A0011	10	1	GBHS MALOTE	1	1
1001111	1	10011	A0011	11	1	GBHS MALOTE	1	1
1001112	1	10011	A0011	12	1	GBHS MALOTE	1	1
1001113	1	10011	A0011	13	1	GBHS MALOTE	1	1
1001114	1	10011	A0011	14	1	GBHS MALOTE	1	1
1001115	1	10011	A0011	15	1	GBHS MALOTE	1	1
1001116	1	10011	A0011	16	1	GBHS MALOTE	1	1

1001117	1	10011	A0011	17	1	GBHS MALOTE	1	1
1001118	1	10011	A0011	18	1	GBHS MALOTE	1	1
1001119	1	10011	A0011	19	1	GBHS MALOTE	1	1
1001120	1	10011	A0011	20	1	GBHS MALOTE	1	1
1001201	1	10012	A0012	1	1	MSQ BINAKHA (Winter)	2	1
1001202	1	10012	A0012	2	1	MSQ BINAKHA (Winter)	2	1
1001203	1	10012	A0012	3	1	MSQ BINAKHA (Winter)	2	1
1001204	1	10012	A0012	4	1	MSQ BINAKHA (Winter)	2	1
1001205	1	10012	A0012	5	1	MSQ BINAKHA (Winter)	2	1
1001206	1	10012	A0012	6	1	MSQ BINAKHA (Winter)	2	1
1001207	1	10012	A0012	7	1	MSQ BINAKHA (Winter)	2	1
1001208	1	10012	A0012	8	1	MSQ BINAKHA (Winter)	2	1
1001209	1	10012	A0012	9	1	MSQ BINAKHA (Winter)	2	1
1001210	1	10012	A0012	10	1	MSQ BINAKHA (Winter)	2	1
1001211	1	10012	A0012	11	1	MSQ BINAKHA (Winter)	2	1
1001212	1	10012	A0012	12	1	MSQ BINAKHA (Winter)	2	1
1001221	1	10012	A0012	21	1	MSQ BINAKHA (Winter)	2	1
1001222	1	10012	A0012	22	1	MSQ BINAKHA (Winter)	2	1
1001301	1	10013	A0013	1	1	GGHS KALRI	2	1
1001302	1	10013	A0013	2	1	GGHS KALRI	2	1
1001303	1	10013	A0013	3	1	GGHS KALRI	2	1
1001304	1	10013	A0013	4	1	GGHS KALRI	2	1

1001305	1	10013	A0013	5	1	GGHS KALRI	2	1
1001306	1	10013	A0013	6	1	GGHS KALRI	2	1
1001307	1	10013	A0013	7	1	GGHS KALRI	2	1
1001308	1	10013	A0013	8	1	GGHS KALRI	2	1
1001309	1	10013	A0013	9	1	GGHS KALRI	2	1
1001310	1	10013	A0013	10	1	GGHS KALRI	2	1
1001311	1	10013	A0013	11	1	GGHS KALRI	2	1
1001312	1	10013	A0013	12	1	GGHS KALRI	2	1
1001313	1	10013	A0013	13	1	GGHS KALRI	2	1
1001314	1	10013	A0013	14	1	GGHS KALRI	2	1
1001315	1	10013	A0013	15	1	GGHS KALRI	2	1
1001316	1	10013	A0013	16	1	GGHS KALRI	2	1
1001317	1	10013	A0013	17	1	GGHS KALRI	2	1
1001318	1	10013	A0013	18	1	GGHS KALRI	2	1
1001319	1	10013	A0013	19	1	GGHS KALRI	2	1
1001320	1	10013	A0013	20	1	GGHS KALRI	2	1
1001321	1	10013	A0013	21	1	GGHS KALRI	2	1
1001322	1	10013	A0013	22	1	GGHS KALRI	2	1
1001401	1	10014	A0014	1	2	GGHS CHHATRO	2	1
1001402	1	10014	A0014	2	2	GGHS CHHATRO	2	1
1001403	1	10014	A0014	3	2	GGHS CHHATRO	2	1
1001404	1	10014	A0014	4	2	GGHS CHHATRO	2	1

1001405	1	10014	A0014	5	2	GGHS CHHATRO	2	1
1001406	1	10014	A0014	6	2	GGHS CHHATRO	2	1
1001407	1	10014	A0014	7	2	GGHS CHHATRO	2	1
1001408	1	10014	A0014	8	2	GGHS CHHATRO	2	1
1001409	1	10014	A0014	9	2	GGHS CHHATRO	2	1
1001410	1	10014	A0014	10	2	GGHS CHHATRO	2	1
1001411	1	10014	A0014	11	2	GGHS CHHATRO	2	1
1001412	1	10014	A0014	12	2	GGHS CHHATRO	2	1
1001413	1	10014	A0014	13	2	GGHS CHHATRO	2	1
1001414	1	10014	A0014	14	2	GGHS CHHATRO	2	1
1001415	1	10014	A0014	15	2	GGHS CHHATRO	2	1
1001416	1	10014	A0014	16	2	GGHS CHHATRO	2	1
1001417	1	10014	A0014	17	2	GGHS CHHATRO	2	1
1001418	1	10014	A0014	18	2	GGHS CHHATRO	2	1
1001419	1	10014	A0014	19	2	GGHS CHHATRO	2	1
1001420	1	10014	A0014	20	2	GGHS CHHATRO	2	1
1001421	1	10014	A0014	21	2	GGHS CHHATRO	2	1
1001422	1	10014	A0014	22	2	GGHS CHHATRO	2	1
1001501	1	10015	A0015	1	1	GGMS THARA	2	1
1001502	1	10015	A0015	2	1	GGMS THARA	2	1
1001503	1	10015	A0015	3	1	GGMS THARA	2	1
1001504	1	10015	A0015	4	1	GGMS THARA	2	1

1001505	1	10015	A0015	5	1	GGMS THARA	2	1
1001506	1	10015	A0015	6	1	GGMS THARA	2	1
1001507	1	10015	A0015	7	1	GGMS THARA	2	1
1001508	1	10015	A0015	8	1	GGMS THARA	2	1
1001509	1	10015	A0015	9	1	GGMS THARA	2	1
1001510	1	10015	A0015	10	1	GGMS THARA	2	1
1001511	1	10015	A0015	11	1	GGMS THARA	2	1
1001512	1	10015	A0015	12	1	GGMS THARA	2	1
1001513	1	10015	A0015	13	1	GGMS THARA	2	1
1001514	1	10015	A0015	14	1	GGMS THARA	2	1
1001515	1	10015	A0015	15	1	GGMS THARA	2	1
1001516	1	10015	A0015	16	1	GGMS THARA	2	1
1001517	1	10015	A0015	17	1	GGMS THARA	2	1
1001518	1	10015	A0015	18	1	GGMS THARA	2	1
1001601	1	10016	A0016	1	1	GBMS KHARANG	1	1
1001602	1	10016	A0016	2	1	GBMS KHARANG	1	1
1001603	1	10016	A0016	3	1	GBMS KHARANG	1	1
1001604	1	10016	A0016	4	1	GBMS KHARANG	1	1
1001605	1	10016	A0016	5	1	GBMS KHARANG	1	1
1001606	1	10016	A0016	6	1	GBMS KHARANG	1	1
1001607	1	10016	A0016	7	1	GBMS KHARANG	1	1
1001608	1	10016	A0016	8	1	GBMS KHARANG	1	1

1001609	1	10016	A0016	9	1	GBMS KHARANG	1	1
1001610	1	10016	A0016	10	1	GBMS KHARANG	1	1
1001611	1	10016	A0016	11	1	GBMS KHARANG	1	1
1001612	1	10016	A0016	12	1	GBMS KHARANG	1	1
1001613	1	10016	A0016	13	1	GBMS KHARANG	1	1
1001614	1	10016	A0016	14	1	GBMS KHARANG	1	1
1001615	1	10016	A0016	15	1	GBMS KHARANG	1	1
1001616	1	10016	A0016	16	1	GBMS KHARANG	1	1
1001617	1	10016	A0016	17	1	GBMS KHARANG	1	1
1001618	1	10016	A0016	18	1	GBMS KHARANG	1	1
1001619	1	10016	A0016	19	1	GBMS KHARANG	1	1
1001620	1	10016	A0016	20	1	GBMS KHARANG	1	1
1001621	1	10016	A0016	21	1	GBMS KHARANG	1	1
1001622	1	10016	A0016	22	1	GBMS KHARANG	1	1
1001701	1	10017	A0017	1	1	GBHS AWERA	3	1
1001702	1	10017	A0017	2	1	GBHS AWERA	3	1
1001703	1	10017	A0017	3	1	GBHS AWERA	3	1
1001704	1	10017	A0017	4	1	GBHS AWERA	3	1
1001705	1	10017	A0017	5	1	GBHS AWERA	3	1
1001706	1	10017	A0017	6	1	GBHS AWERA	3	1
1001707	1	10017	A0017	7	1	GBHS AWERA	3	1
1001708	1	10017	A0017	8	1	GBHS AWERA	3	1

1001709	1	10017	A0017	9	1	GBHS AWERA	3	1
1001710	1	10017	A0017	10	1	GBHS AWERA	3	1
1001711	1	10017	A0017	11	1	GBHS AWERA	3	1
1001712	1	10017	A0017	12	1	GBHS AWERA	3	1
1001713	1	10017	A0017	13	1	GBHS AWERA	3	1
1001714	1	10017	A0017	14	1	GBHS AWERA	3	1
1001715	1	10017	A0017	15	1	GBHS AWERA	3	1
1001716	1	10017	A0017	16	1	GBHS AWERA	3	1
1001717	1	10017	A0017	17	1	GBHS AWERA	3	1
1001718	1	10017	A0017	18	1	GBHS AWERA	3	1
1001719	1	10017	A0017	19	1	GBHS AWERA	3	1
1001720	1	10017	A0017	20	1	GBHS AWERA	3	1
1001721	1	10017	A0017	21	1	GBHS AWERA	3	1
1001722	1	10017	A0017	22	1	GBHS AWERA	3	1
1001801	1	10018	A0018	1	1	GBMS BANDI CHAKKAN	3	1
1001802	1	10018	A0018	2	1	GBMS BANDI CHAKKAN	3	1
1001803	1	10018	A0018	3	1	GBMS BANDI CHAKKAN	3	1
1001804	1	10018	A0018	4	1	GBMS BANDI CHAKKAN	3	1
1001805	1	10018	A0018	5	1	GBMS BANDI CHAKKAN	3	1
1001806	1	10018	A0018	6	1	GBMS BANDI CHAKKAN	3	1
1001807	1	10018	A0018	7	1	GBMS BANDI CHAKKAN	3	1
1001808	1	10018	A0018	8	1	GBMS BANDI CHAKKAN	3	1

1001809	1	10018	A0018	9	1	GBMS BANDI CHAKKAN	3	1
1001810	1	10018	A0018	10	1	GBMS BANDI CHAKKAN	3	1
1001811	1	10018	A0018	11	1	GBMS BANDI CHAKKAN	3	1
1001812	1	10018	A0018	12	1	GBMS BANDI CHAKKAN	3	1
1001813	1	10018	A0018	13	1	GBMS BANDI CHAKKAN	3	1
1001821	1	10018	A0018	21	1	GBMS BANDI CHAKKAN	3	1
1001822	1	10018	A0018	22	1	GBMS BANDI CHAKKAN	3	1
1001901	1	10019	A0019	1	2	GGHS SARRAN	2	1
1001902	1	10019	A0019	2	2	GGHS SARRAN	2	1
1001903	1	10019	A0019	3	2	GGHS SARRAN	2	1
1001904	1	10019	A0019	4	2	GGHS SARRAN	2	1
1001905	1	10019	A0019	5	2	GGHS SARRAN	2	1
1001906	1	10019	A0019	6	2	GGHS SARRAN	2	1
1001907	1	10019	A0019	7	2	GGHS SARRAN	2	1
1001908	1	10019	A0019	8	2	GGHS SARRAN	2	1
1001909	1	10019	A0019	9	2	GGHS SARRAN	2	1
1001910	1	10019	A0019	10	2	GGHS SARRAN	2	1
1001911	1	10019	A0019	11	2	GGHS SARRAN	2	1
1001912	1	10019	A0019	12	2	GGHS SARRAN	2	1
1001913	1	10019	A0019	13	2	GGHS SARRAN	2	1
1001914	1	10019	A0019	14	2	GGHS SARRAN	2	1
1001915	1	10019	A0019	15	2	GGHS SARRAN	2	1

1001916	1	10019	A0019	16	2	GGHS SARRAN	2	1
1001917	1	10019	A0019	17	2	GGHS SARRAN	2	1
1001918	1	10019	A0019	18	2	GGHS SARRAN	2	1
1001919	1	10019	A0019	19	2	GGHS SARRAN	2	1
1001921	1	10019	A0019	21	2	GGHS SARRAN	2	1
1001922	1	10019	A0019	22	2	GGHS SARRAN	2	1
1002001	1	10020	A0020	1	1	GBMS HATTIAN BALA	1	1
1002002	1	10020	A0020	2	1	GBMS HATTIAN BALA	1	1
1002003	1	10020	A0020	3	1	GBMS HATTIAN BALA	1	1
1002004	1	10020	A0020	4	1	GBMS HATTIAN BALA	1	1
1002005	1	10020	A0020	5	1	GBMS HATTIAN BALA	1	1
1002006	1	10020	A0020	6	1	GBMS HATTIAN BALA	1	1
1002007	1	10020	A0020	7	1	GBMS HATTIAN BALA	1	1
1002008	1	10020	A0020	8	1	GBMS HATTIAN BALA	1	1
1002009	1	10020	A0020	9	1	GBMS HATTIAN BALA	1	1
1002010	1	10020	A0020	10	1	GBMS HATTIAN BALA	1	1
1002011	1	10020	A0020	11	1	GBMS HATTIAN BALA	1	1
1002012	1	10020	A0020	12	1	GBMS HATTIAN BALA	1	1
1002013	1	10020	A0020	13	1	GBMS HATTIAN BALA	1	1
1002014	1	10020	A0020	14	1	GBMS HATTIAN BALA	1	1
1002015	1	10020	A0020	15	1	GBMS HATTIAN BALA	1	1
1002016	1	10020	A0020	16	1	GBMS HATTIAN BALA	1	1

1002017	1	10020	A0020	17	1	GBMS HATTIAN BALA	1	1
1002018	1	10020	A0020	18	1	GBMS HATTIAN BALA	1	1
1002019	1	10020	A0020	19	1	GBMS HATTIAN BALA	1	1
1002021	1	10020	A0020	21	1	GBMS HATTIAN BALA	1	1
1002022	1	10020	A0020	22	1	GBMS HATTIAN BALA	1	1
1002101	1	10021	A0021	1	1	GGPS THILL	2	1
1002102	1	10021	A0021	2	1	GGPS THILL	2	1
1002103	1	10021	A0021	3	1	GGPS THILL	2	1
1002104	1	10021	A0021	4	1	GGPS THILL	2	1
1002105	1	10021	A0021	5	1	GGPS THILL	2	1
1002106	1	10021	A0021	6	1	GGPS THILL	2	1
1002201	1	10022	A0022	1	1	GBHS GIAEEN	3	1
1002202	1	10022	A0022	2	1	GBHS GIAEEN	3	1
1002203	1	10022	A0022	3	1	GBHS GIAEEN	3	1
1002204	1	10022	A0022	4	1	GBHS GIAEEN	3	1
1002205	1	10022	A0022	5	1	GBHS GIAEEN	3	1
1002206	1	10022	A0022	6	1	GBHS GIAEEN	3	1
1002207	1	10022	A0022	7	1	GBHS GIAEEN	3	1
1002208	1	10022	A0022	8	1	GBHS GIAEEN	3	1
1002209	1	10022	A0022	9	1	GBHS GIAEEN	3	1
1002210	1	10022	A0022	10	1	GBHS GIAEEN	3	1
1002211	1	10022	A0022	11	1	GBHS GIAEEN	3	1

1002301	1	10023	A0023	1	1	GBPS BERHI BALA	3	1
1002302	1	10023	A0023	2	1	GBPS BERHI BALA	3	1
1002303	1	10023	A0023	3	1	GBPS BERHI BALA	3	1
1002304	1	10023	A0023	4	1	GBPS BERHI BALA	3	1
1002305	1	10023	A0023	5	1	GBPS BERHI BALA	3	1
1002306	1	10023	A0023	6	1	GBPS BERHI BALA	3	1
1002307	1	10023	A0023	7	1	GBPS BERHI BALA	3	1
1002308	1	10023	A0023	8	1	GBPS BERHI BALA	3	1
1002309	1	10023	A0023	9	1	GBPS BERHI BALA	3	1
1002310	1	10023	A0023	10	1	GBPS BERHI BALA	3	1
1002311	1	10023	A0023	11	1	GBPS BERHI BALA	3	1
1002312	1	10023	A0023	12	1	GBPS BERHI BALA	3	1
1002313	1	10023	A0023	13	1	GBPS BERHI BALA	3	1
1002314	1	10023	A0023	14	1	GBPS BERHI BALA	3	1
1002315	1	10023	A0023	15	1	GBPS BERHI BALA	3	1
1002316	1	10023	A0023	16	1	GBPS BERHI BALA	3	1
1002317	1	10023	A0023	17	1	GBPS BERHI BALA	3	1
1002318	1	10023	A0023	18	1	GBPS BERHI BALA	3	1
1002319	1	10023	A0023	19	1	GBPS BERHI BALA	3	1
1002320	1	10023	A0023	20	1	GBPS BERHI BALA	3	1
1002401	1	10024	A0024	1	1	GGMS SAMLAR NO. 1	2	1
1002402	1	10024	A0024	2	1	GGMS SAMLAR NO. 1	2	1

1002403	1	10024	A0024	3	1	GGMS SAMLAR NO. 1	2	1
1002404	1	10024	A0024	4	1	GGMS SAMLAR NO. 1	2	1
1002405	1	10024	A0024	5	1	GGMS SAMLAR NO. 1	2	1
1002406	1	10024	A0024	6	1	GGMS SAMLAR NO. 1	2	1
1002407	1	10024	A0024	7	1	GGMS SAMLAR NO. 1	2	1
1002408	1	10024	A0024	8	1	GGMS SAMLAR NO. 1	2	1
1002409	1	10024	A0024	9	1	GGMS SAMLAR NO. 1	2	1
1002410	1	10024	A0024	10	1	GGMS SAMLAR NO. 1	2	1
1002411	1	10024	A0024	11	1	GGMS SAMLAR NO. 1	2	1
1002412	1	10024	A0024	12	1	GGMS SAMLAR NO. 1	2	1
1002413	1	10024	A0024	13	1	GGMS SAMLAR NO. 1	2	1
1002414	1	10024	A0024	14	1	GGMS SAMLAR NO. 1	2	1
1002415	1	10024	A0024	15	1	GGMS SAMLAR NO. 1	2	1
1002416	1	10024	A0024	16	1	GGMS SAMLAR NO. 1	2	1
1002417	1	10024	A0024	17	1	GGMS SAMLAR NO. 1	2	1
1002418	1	10024	A0024	18	1	GGMS SAMLAR NO. 1	2	1
1002419	1	10024	A0024	19	1	GGMS SAMLAR NO. 1	2	1
1002420	1	10024	A0024	20	1	GGMS SAMLAR NO. 1	2	1
1002421	1	10024	A0024	21	1	GGMS SAMLAR NO. 1	2	1
1002422	1	10024	A0024	22	1	GGMS SAMLAR NO. 1	2	1
1002501	1	10025	A0025	1	1	GBHS SERI	1	1
1002502	1	10025	A0025	2	1	GBHS SERI	1	1

1002503	1	10025	A0025	3	1	GBHS SERI	1	1
1002504	1	10025	A0025	4	1	GBHS SERI	1	1
1002505	1	10025	A0025	5	1	GBHS SERI	1	1
1002506	1	10025	A0025	6	1	GBHS SERI	1	1
1002507	1	10025	A0025	7	1	GBHS SERI	1	1
1002508	1	10025	A0025	8	1	GBHS SERI	1	1
1002509	1	10025	A0025	9	1	GBHS SERI	1	1
1002601	1	10026	A0026	1	1	GBPS CHANGUN	3	1
1002602	1	10026	A0026	2	1	GBPS CHANGUN	3	1
1002603	1	10026	A0026	3	1	GBPS CHANGUN	3	1
1002604	1	10026	A0026	4	1	GBPS CHANGUN	3	1
1002605	1	10026	A0026	5	1	GBPS CHANGUN	3	1
1002606	1	10026	A0026	6	1	GBPS CHANGUN	3	1
1002607	1	10026	A0026	7	1	GBPS CHANGUN	3	1
1002608	1	10026	A0026	8	1	GBPS CHANGUN	3	1
1002609	1	10026	A0026	9	1	GBPS CHANGUN	3	1
1002610	1	10026	A0026	10	1	GBPS CHANGUN	3	1
1002611	1	10026	A0026	11	1	GBPS CHANGUN	3	1
1002612	1	10026	A0026	12	1	GBPS CHANGUN	3	1
1002621	1	10026	A0026	21	1	GBPS CHANGUN	3	1
1002622	1	10026	A0026	22	1	GBPS CHANGUN	3	1
1002701	1	10027	A0027	1	1	GBPS CHARNARI	3	1

1002702	1	10027	A0027	2	1	GBPS CHARNARI	3	1
1002703	1	10027	A0027	3	1	GBPS CHARNARI	3	1
1002704	1	10027	A0027	4	1	GBPS CHARNARI	3	1
1002705	1	10027	A0027	5	1	GBPS CHARNARI	3	1
1002706	1	10027	A0027	6	1	GBPS CHARNARI	3	1
1002707	1	10027	A0027	7	1	GBPS CHARNARI	3	1
1002708	1	10027	A0027	8	1	GBPS CHARNARI	3	1
1002709	1	10027	A0027	9	1	GBPS CHARNARI	3	1
1002710	1	10027	A0027	10	1	GBPS CHARNARI	3	1
1002711	1	10027	A0027	11	1	GBPS CHARNARI	3	1
1002712	1	10027	A0027	12	1	GBPS CHARNARI	3	1
1002713	1	10027	A0027	13	1	GBPS CHARNARI	3	1
1002714	1	10027	A0027	14	1	GBPS CHARNARI	3	1
1002715	1	10027	A0027	15	1	GBPS CHARNARI	3	1
1002716	1	10027	A0027	16	1	GBPS CHARNARI	3	1
1002717	1	10027	A0027	17	1	GBPS CHARNARI	3	1
1002801	1	10028	A0028	1	1	GBPS KUND KALAR	1	1
1002802	1	10028	A0028	2	1	GBPS KUND KALAR	1	1
1002803	1	10028	A0028	3	1	GBPS KUND KALAR	1	1
1002804	1	10028	A0028	4	1	GBPS KUND KALAR	1	1
1002805	1	10028	A0028	5	1	GBPS KUND KALAR	1	1
1002806	1	10028	A0028	6	1	GBPS KUND KALAR	1	1

1002807	1	10028	A0028	7	1	GBPS KUND KALAR	1	1
1002808	1	10028	A0028	8	1	GBPS KUND KALAR	1	1
1002809	1	10028	A0028	9	1	GBPS KUND KALAR	1	1
1002810	1	10028	A0028	10	1	GBPS KUND KALAR	1	1
1002811	1	10028	A0028	11	1	GBPS KUND KALAR	1	1
1002812	1	10028	A0028	12	1	GBPS KUND KALAR	1	1
1002813	1	10028	A0028	13	1	GBPS KUND KALAR	1	1
1002814	1	10028	A0028	14	1	GBPS KUND KALAR	1	1
1002815	1	10028	A0028	15	1	GBPS KUND KALAR	1	1
1002816	1	10028	A0028	16	1	GBPS KUND KALAR	1	1
1002817	1	10028	A0028	17	1	GBPS KUND KALAR	1	1
1002818	1	10028	A0028	18	1	GBPS KUND KALAR	1	1
1002819	1	10028	A0028	19	1	GBPS KUND KALAR	1	1
1002820	1	10028	A0028	20	1	GBPS KUND KALAR	1	1
1002901	1	10029	A0029	1	1	GBHS KURTI	1	1
1002902	1	10029	A0029	2	1	GBHS KURTI	1	1
1002903	1	10029	A0029	3	1	GBHS KURTI	1	1
1002904	1	10029	A0029	4	1	GBHS KURTI	1	1
1002905	1	10029	A0029	5	1	GBHS KURTI	1	1
1002906	1	10029	A0029	6	1	GBHS KURTI	1	1
1002907	1	10029	A0029	7	1	GBHS KURTI	1	1
1002908	1	10029	A0029	8	1	GBHS KURTI	1	1

1002909	1	10029	A0029	9	1	GBHS KURTI	1	1
1002910	1	10029	A0029	10	1	GBHS KURTI	1	1
1002911	1	10029	A0029	11	1	GBHS KURTI	1	1
1002912	1	10029	A0029	12	1	GBHS KURTI	1	1
1002913	1	10029	A0029	13	1	GBHS KURTI	1	1
1002914	1	10029	A0029	14	1	GBHS KURTI	1	1
1002915	1	10029	A0029	15	1	GBHS KURTI	1	1
1002916	1	10029	A0029	16	1	GBHS KURTI	1	1
1002917	1	10029	A0029	17	1	GBHS KURTI	1	1
1002918	1	10029	A0029	18	1	GBHS KURTI	1	1
1002919	1	10029	A0029	19	1	GBHS KURTI	1	1
1002920	1	10029	A0029	20	1	GBHS KURTI	1	1
1002921	1	10029	A0029	21	1	GBHS KURTI	1	1
1002922	1	10029	A0029	22	1	GBHS KURTI	1	1
1003001	1	10030	A0030	1	2	GGHS KANALI	2	1
1003002	1	10030	A0030	2	2	GGHS KANALI	2	1
1003003	1	10030	A0030	3	2	GGHS KANALI	2	1
1003004	1	10030	A0030	4	2	GGHS KANALI	2	1
1003005	1	10030	A0030	5	2	GGHS KANALI	2	1
1003006	1	10030	A0030	6	2	GGHS KANALI	2	1
1003007	1	10030	A0030	7	2	GGHS KANALI	2	1
1003008	1	10030	A0030	8	2	GGHS KANALI	2	1

1003009	1	10030	A0030	9	2	GGHS KANALI	2	1
1003010	1	10030	A0030	10	2	GGHS KANALI	2	1
1003011	1	10030	A0030	11	2	GGHS KANALI	2	1
1003012	1	10030	A0030	12	2	GGHS KANALI	2	1
1003013	1	10030	A0030	13	2	GGHS KANALI	2	1
1003014	1	10030	A0030	14	2	GGHS KANALI	2	1
1003015	1	10030	A0030	15	2	GGHS KANALI	2	1
1003016	1	10030	A0030	16	2	GGHS KANALI	2	1
1003017	1	10030	A0030	17	2	GGHS KANALI	2	1
1003101	1	10031	A0031	1	2	GGHS SECTOR F-1 MIRPUR	2	1
1003102	1	10031	A0031	2	2	GGHS SECTOR F-1 MIRPUR	2	1
1003103	1	10031	A0031	3	2	GGHS SECTOR F-1 MIRPUR	2	1
1003104	1	10031	A0031	4	2	GGHS SECTOR F-1 MIRPUR	2	1
1003105	1	10031	A0031	5	2	GGHS SECTOR F-1 MIRPUR	2	1
1003106	1	10031	A0031	6	2	GGHS SECTOR F-1 MIRPUR	2	1
1003107	1	10031	A0031	7	2	GGHS SECTOR F-1 MIRPUR	2	1
1003108	1	10031	A0031	8	2	GGHS SECTOR F-1 MIRPUR	2	1
1003109	1	10031	A0031	9	2	GGHS SECTOR F-1 MIRPUR	2	1
1003110	1	10031	A0031	10	2	GGHS SECTOR F-1 MIRPUR	2	1
1003111	1	10031	A0031	11	2	GGHS SECTOR F-1 MIRPUR	2	1
1003112	1	10031	A0031	12	2	GGHS SECTOR F-1 MIRPUR	2	1
1003113	1	10031	A0031	13	2	GGHS SECTOR F-1 MIRPUR	2	1

1003114	1	10031	A0031	14	2	GGHS SECTOR F-1 MIRPUR	2	1
1003115	1	10031	A0031	15	2	GGHS SECTOR F-1 MIRPUR	2	1
1003116	1	10031	A0031	16	2	GGHS SECTOR F-1 MIRPUR	2	1
1003117	1	10031	A0031	17	2	GGHS SECTOR F-1 MIRPUR	2	1
1003118	1	10031	A0031	18	2	GGHS SECTOR F-1 MIRPUR	2	1
1003119	1	10031	A0031	19	2	GGHS SECTOR F-1 MIRPUR	2	1
1003120	1	10031	A0031	20	2	GGHS SECTOR F-1 MIRPUR	2	1
1003121	1	10031	A0031	21	2	GGHS SECTOR F-1 MIRPUR	2	1
1003122	1	10031	A0031	22	2	GGHS SECTOR F-1 MIRPUR	2	1
1003201	1	10032	A0032	1	2	GGMS THOTHAL	2	1
1003202	1	10032	A0032	2	2	GGMS THOTHAL	2	1
1003203	1	10032	A0032	3	2	GGMS THOTHAL	2	1
1003204	1	10032	A0032	4	2	GGMS THOTHAL	2	1
1003205	1	10032	A0032	5	2	GGMS THOTHAL	2	1
1003206	1	10032	A0032	6	2	GGMS THOTHAL	2	1
1003207	1	10032	A0032	7	2	GGMS THOTHAL	2	1
1003208	1	10032	A0032	8	2	GGMS THOTHAL	2	1
1003209	1	10032	A0032	9	2	GGMS THOTHAL	2	1
1003210	1	10032	A0032	10	2	GGMS THOTHAL	2	1
1003211	1	10032	A0032	11	2	GGMS THOTHAL	2	1
1003212	1	10032	A0032	12	2	GGMS THOTHAL	2	1
1003213	1	10032	A0032	13	2	GGMS THOTHAL	2	1

1003214	1	10032	A0032	14	2	GGMS THOTHAL	2	1
1003222	1	10032	A0032	22	2	GGMS THOTHAL	2	1
1003301	1	10033	A0033	1	1	GGHS NEW CITY KALYAL (Shift-2)	2	1
1003302	1	10033	A0033	2	1	GGHS NEW CITY KALYAL (Shift-2)	2	1
1003303	1	10033	A0033	3	1	GGHS NEW CITY KALYAL (Shift-2)	2	1
1003304	1	10033	A0033	4	1	GGHS NEW CITY KALYAL (Shift-2)	2	1
1003305	1	10033	A0033	5	1	GGHS NEW CITY KALYAL (Shift-2)	2	1
1003306	1	10033	A0033	6	1	GGHS NEW CITY KALYAL (Shift-2)	2	1
1003307	1	10033	A0033	7	1	GGHS NEW CITY KALYAL (Shift-2)	2	1
1003308	1	10033	A0033	8	1	GGHS NEW CITY KALYAL (Shift-2)	2	1
1003309	1	10033	A0033	9	1	GGHS NEW CITY KALYAL (Shift-2)	2	1
1003310	1	10033	A0033	10	1	GGHS NEW CITY KALYAL (Shift-2)	2	1
1003311	1	10033	A0033	11	1	GGHS NEW CITY KALYAL (Shift-2)	2	1
1003312	1	10033	A0033	12	1	GGHS NEW CITY KALYAL (Shift-2)	2	1
1003313	1	10033	A0033	13	1	GGHS NEW CITY KALYAL (Shift-2)	2	1

1003314	1	10033	A0033	14	1	GGHS NEW CITY KALYAL (Shift-2)	2	1
1003315	1	10033	A0033	15	1	GGHS NEW CITY KALYAL (Shift-2)	2	1
1003316	1	10033	A0033	16	1	GGHS NEW CITY KALYAL (Shift-2)	2	1
1003317	1	10033	A0033	17	1	GGHS NEW CITY KALYAL (Shift-2)	2	1
1003318	1	10033	A0033	18	1	GGHS NEW CITY KALYAL (Shift-2)	2	1
1003319	1	10033	A0033	19	1	GGHS NEW CITY KALYAL (Shift-2)	2	1
1003320	1	10033	A0033	20	1	GGHS NEW CITY KALYAL (Shift-2)	2	1
1003321	1	10033	A0033	21	1	GGHS NEW CITY KALYAL (Shift-2)	2	1
1003322	1	10033	A0033	22	1	GGHS NEW CITY KALYAL (Shift-2)	2	1
1003401	1	10034	A0034	1	1	GGMS PRAME KOT	2	1
1003402	1	10034	A0034	2	1	GGMS PRAME KOT	2	1
1003403	1	10034	A0034	3	1	GGMS PRAME KOT	2	1
1003404	1	10034	A0034	4	1	GGMS PRAME KOT	2	1
1003405	1	10034	A0034	5	1	GGMS PRAME KOT	2	1
1003406	1	10034	A0034	6	1	GGMS PRAME KOT	2	1
1003407	1	10034	A0034	7	1	GGMS PRAME KOT	2	1
1003408	1	10034	A0034	8	1	GGMS PRAME KOT	2	1

1003409	1	10034	A0034	9	1	GGMS PRAME KOT	2	1
1003410	1	10034	A0034	10	1	GGMS PRAME KOT	2	1
1003411	1	10034	A0034	11	1	GGMS PRAME KOT	2	1
1003412	1	10034	A0034	12	1	GGMS PRAME KOT	2	1
1003413	1	10034	A0034	13	1	GGMS PRAME KOT	2	1
1003421	1	10034	A0034	21	1	GGMS PRAME KOT	2	1
1003422	1	10034	A0034	22	1	GGMS PRAME KOT	2	1
1003501	1	10035	A0035	1	2	GBMS SALEH GALLI	3	1
1003502	1	10035	A0035	2	2	GBMS SALEH GALLI	3	1
1003503	1	10035	A0035	3	2	GBMS SALEH GALLI	3	1
1003504	1	10035	A0035	4	2	GBMS SALEH GALLI	3	1
1003505	1	10035	A0035	5	2	GBMS SALEH GALLI	3	1
1003506	1	10035	A0035	6	2	GBMS SALEH GALLI	3	1
1003507	1	10035	A0035	7	2	GBMS SALEH GALLI	3	1
1003508	1	10035	A0035	8	2	GBMS SALEH GALLI	3	1
1003509	1	10035	A0035	9	2	GBMS SALEH GALLI	3	1
1003510	1	10035	A0035	10	2	GBMS SALEH GALLI	3	1
1003511	1	10035	A0035	11	2	GBMS SALEH GALLI	3	1
1003512	1	10035	A0035	12	2	GBMS SALEH GALLI	3	1
1003513	1	10035	A0035	13	2	GBMS SALEH GALLI	3	1
1003514	1	10035	A0035	14	2	GBMS SALEH GALLI	3	1
1003515	1	10035	A0035	15	2	GBMS SALEH GALLI	3	1

1003516	1	10035	A0035	16	2	GBMS SALEH GALLI	3	1
1003517	1	10035	A0035	17	2	GBMS SALEH GALLI	3	1
1003518	1	10035	A0035	18	2	GBMS SALEH GALLI	3	1
1003521	1	10035	A0035	21	2	GBMS SALEH GALLI	3	1
1003601	1	10036	A0036	1	1	GBHS RANJATA	1	1
1003602	1	10036	A0036	2	1	GBHS RANJATA	1	1
1003603	1	10036	A0036	3	1	GBHS RANJATA	1	1
1003604	1	10036	A0036	4	1	GBHS RANJATA	1	1
1003605	1	10036	A0036	5	1	GBHS RANJATA	1	1
1003606	1	10036	A0036	6	1	GBHS RANJATA	1	1
1003607	1	10036	A0036	7	1	GBHS RANJATA	1	1
1003608	1	10036	A0036	8	1	GBHS RANJATA	1	1
1003609	1	10036	A0036	9	1	GBHS RANJATA	1	1
1003610	1	10036	A0036	10	1	GBHS RANJATA	1	1
1003611	1	10036	A0036	11	1	GBHS RANJATA	1	1
1003612	1	10036	A0036	12	1	GBHS RANJATA	1	1
1003613	1	10036	A0036	13	1	GBHS RANJATA	1	1
1003614	1	10036	A0036	14	1	GBHS RANJATA	1	1
1003615	1	10036	A0036	15	1	GBHS RANJATA	1	1
1003616	1	10036	A0036	16	1	GBHS RANJATA	1	1
1003617	1	10036	A0036	17	1	GBHS RANJATA	1	1
1003618	1	10036	A0036	18	1	GBHS RANJATA	1	1

1003619	1	10036	A0036	19	1	GBHS RANJATA	1	1
1003620	1	10036	A0036	20	1	GBHS RANJATA	1	1
1003621	1	10036	A0036	21	1	GBHS RANJATA	1	1
1003622	1	10036	A0036	22	1	GBHS RANJATA	1	1
1003701	1	10037	A0037	1	2	GGHS JUGLARI	2	1
1003702	1	10037	A0037	2	2	GGHS JUGLARI	2	1
1003703	1	10037	A0037	3	2	GGHS JUGLARI	2	1
1003704	1	10037	A0037	4	2	GGHS JUGLARI	2	1
1003705	1	10037	A0037	5	2	GGHS JUGLARI	2	1
1003706	1	10037	A0037	6	2	GGHS JUGLARI	2	1
1003707	1	10037	A0037	7	2	GGHS JUGLARI	2	1
1003708	1	10037	A0037	8	2	GGHS JUGLARI	2	1
1003709	1	10037	A0037	9	2	GGHS JUGLARI	2	1
1003710	1	10037	A0037	10	2	GGHS JUGLARI	2	1
1003711	1	10037	A0037	11	2	GGHS JUGLARI	2	1
1003712	1	10037	A0037	12	2	GGHS JUGLARI	2	1
1003713	1	10037	A0037	13	2	GGHS JUGLARI	2	1
1003714	1	10037	A0037	14	2	GGHS JUGLARI	2	1
1003715	1	10037	A0037	15	2	GGHS JUGLARI	2	1
1003716	1	10037	A0037	16	2	GGHS JUGLARI	2	1
1003717	1	10037	A0037	17	2	GGHS JUGLARI	2	1
1003718	1	10037	A0037	18	2	GGHS JUGLARI	2	1

1003719	1	10037	A0037	19	2	GGHS JUGLARI	2	1
1003720	1	10037	A0037	20	2	GGHS JUGLARI	2	1
1003721	1	10037	A0037	21	2	GGHS JUGLARI	2	1
1003722	1	10037	A0037	22	2	GGHS JUGLARI	2	1
1003801	1	10038	A0038	1	1	GBHS PALLANDRI	1	1
1003802	1	10038	A0038	2	1	GBHS PALLANDRI	1	1
1003803	1	10038	A0038	3	1	GBHS PALLANDRI	1	1
1003804	1	10038	A0038	4	1	GBHS PALLANDRI	1	1
1003805	1	10038	A0038	5	1	GBHS PALLANDRI	1	1
1003806	1	10038	A0038	6	1	GBHS PALLANDRI	1	1
1003807	1	10038	A0038	7	1	GBHS PALLANDRI	1	1
1003808	1	10038	A0038	8	1	GBHS PALLANDRI	1	1
1003809	1	10038	A0038	9	1	GBHS PALLANDRI	1	1
1003810	1	10038	A0038	10	1	GBHS PALLANDRI	1	1
1003811	1	10038	A0038	11	1	GBHS PALLANDRI	1	1
1003812	1	10038	A0038	12	1	GBHS PALLANDRI	1	1
1003813	1	10038	A0038	13	1	GBHS PALLANDRI	1	1
1003814	1	10038	A0038	14	1	GBHS PALLANDRI	1	1
1003815	1	10038	A0038	15	1	GBHS PALLANDRI	1	1
1003816	1	10038	A0038	16	1	GBHS PALLANDRI	1	1
1003817	1	10038	A0038	17	1	GBHS PALLANDRI	1	1
1003818	1	10038	A0038	18	1	GBHS PALLANDRI	1	1

1003819	1	10038	A0038	19	1	GBHS PALLANDRI	1	1
1003820	1	10038	A0038	20	1	GBHS PALLANDRI	1	1
1003821	1	10038	A0038	21	1	GBHS PALLANDRI	1	1
1003822	1	10038	A0038	22	1	GBHS PALLANDRI	1	1
1003901	1	10039	A0039	1	1	GBPS CHAMIAN KHAITER	3	1
1003902	1	10039	A0039	2	1	GBPS CHAMIAN KHAITER	3	1
1003903	1	10039	A0039	3	1	GBPS CHAMIAN KHAITER	3	1
1003904	1	10039	A0039	4	1	GBPS CHAMIAN KHAITER	3	1
1003905	1	10039	A0039	5	1	GBPS CHAMIAN KHAITER	3	1
1003906	1	10039	A0039	6	1	GBPS CHAMIAN KHAITER	3	1
1003907	1	10039	A0039	7	1	GBPS CHAMIAN KHAITER	3	1
1003908	1	10039	A0039	8	1	GBPS CHAMIAN KHAITER	3	1
1003909	1	10039	A0039	9	1	GBPS CHAMIAN KHAITER	3	1
1003910	1	10039	A0039	10	1	GBPS CHAMIAN KHAITER	3	1
1003911	1	10039	A0039	11	1	GBPS CHAMIAN KHAITER	3	1
1003912	1	10039	A0039	12	1	GBPS CHAMIAN KHAITER	3	1
1003913	1	10039	A0039	13	1	GBPS CHAMIAN KHAITER	3	1
1003914	1	10039	A0039	14	1	GBPS CHAMIAN KHAITER	3	1
1003915	1	10039	A0039	15	1	GBPS CHAMIAN KHAITER	3	1
1003916	1	10039	A0039	16	1	GBPS CHAMIAN KHAITER	3	1
1003917	1	10039	A0039	17	1	GBPS CHAMIAN KHAITER	3	1
1003921	1	10039	A0039	21	1	GBPS CHAMIAN KHAITER	3	1

1003922	1	10039	A0039	22	1	GBPS CHAMIAN KHAITER	3	1
1004001	1	10040	A0040	1	1	GGPS LALA	3	1
1004002	1	10040	A0040	2	1	GGPS LALA	3	1
1004003	1	10040	A0040	3	1	GGPS LALA	3	1
1004004	1	10040	A0040	4	1	GGPS LALA	3	1
1004005	1	10040	A0040	5	1	GGPS LALA	3	1
1004006	1	10040	A0040	6	1	GGPS LALA	3	1
1004007	1	10040	A0040	7	1	GGPS LALA	3	1
1004008	1	10040	A0040	8	1	GGPS LALA	3	1
1004009	1	10040	A0040	9	1	GGPS LALA	3	1
1004010	1	10040	A0040	10	1	GGPS LALA	3	1
1004011	1	10040	A0040	11	1	GGPS LALA	3	1
1004012	1	10040	A0040	12	1	GGPS LALA	3	1
1004013	1	10040	A0040	13	1	GGPS LALA	3	1
1004014	1	10040	A0040	14	1	GGPS LALA	3	1
1004015	1	10040	A0040	15	1	GGPS LALA	3	1
1004021	1	10040	A0040	21	1	GGPS LALA	3	1
1004022	1	10040	A0040	22	1	GGPS LALA	3	1
1004101	1	10041	A0041	1	1	GBHS DAB	1	1
1004102	1	10041	A0041	2	1	GBHS DAB	1	1
1004103	1	10041	A0041	3	1	GBHS DAB	1	1
1004104	1	10041	A0041	4	1	GBHS DAB	1	1

1004105	1	10041	A0041	5	1	GBHS DAB	1	1
1004106	1	10041	A0041	6	1	GBHS DAB	1	1
1004107	1	10041	A0041	7	1	GBHS DAB	1	1
1004108	1	10041	A0041	8	1	GBHS DAB	1	1
1004109	1	10041	A0041	9	1	GBHS DAB	1	1
1004110	1	10041	A0041	10	1	GBHS DAB	1	1
1004111	1	10041	A0041	11	1	GBHS DAB	1	1
1004112	1	10041	A0041	12	1	GBHS DAB	1	1
1004113	1	10041	A0041	13	1	GBHS DAB	1	1
1004114	1	10041	A0041	14	1	GBHS DAB	1	1
1004115	1	10041	A0041	15	1	GBHS DAB	1	1
1004116	1	10041	A0041	16	1	GBHS DAB	1	1
1004117	1	10041	A0041	17	1	GBHS DAB	1	1
1004118	1	10041	A0041	18	1	GBHS DAB	1	1
1004119	1	10041	A0041	19	1	GBHS DAB	1	1
1004120	1	10041	A0041	20	1	GBHS DAB	1	1
1004121	1	10041	A0041	21	1	GBHS DAB	1	1
1004122	1	10041	A0041	22	1	GBHS DAB	1	1
1004201	1	10042	A0042	1	1	MSQ BUTTANGI	3	1
1004202	1	10042	A0042	2	1	MSQ BUTTANGI	3	1
1004203	1	10042	A0042	3	1	MSQ BUTTANGI	3	1
1004204	1	10042	A0042	4	1	MSQ BUTTANGI	3	1

1004205	1	10042	A0042	5	1	MSQ BUTTANGI	3	1
1004206	1	10042	A0042	6	1	MSQ BUTTANGI	3	1
1004207	1	10042	A0042	7	1	MSQ BUTTANGI	3	1
1004208	1	10042	A0042	8	1	MSQ BUTTANGI	3	1
1004209	1	10042	A0042	9	1	MSQ BUTTANGI	3	1
1004210	1	10042	A0042	10	1	MSQ BUTTANGI	3	1
1004211	1	10042	A0042	11	1	MSQ BUTTANGI	3	1
1004212	1	10042	A0042	12	1	MSQ BUTTANGI	3	1
1004213	1	10042	A0042	13	1	MSQ BUTTANGI	3	1
1004214	1	10042	A0042	14	1	MSQ BUTTANGI	3	1
1004215	1	10042	A0042	15	1	MSQ BUTTANGI	3	1
1004216	1	10042	A0042	16	1	MSQ BUTTANGI	3	1
1004217	1	10042	A0042	17	1	MSQ BUTTANGI	3	1
1004218	1	10042	A0042	18	1	MSQ BUTTANGI	3	1
1004219	1	10042	A0042	19	1	MSQ BUTTANGI	3	1
1004220	1	10042	A0042	20	1	MSQ BUTTANGI	3	1
1004221	1	10042	A0042	21	1	MSQ BUTTANGI	3	1
1004222	1	10042	A0042	22	1	MSQ BUTTANGI	3	1
1004301	1	10043	A0043	1	1	GBMS KEL	1	1
1004302	1	10043	A0043	2	1	GBMS KEL	1	1
1004303	1	10043	A0043	3	1	GBMS KEL	1	1
1004304	1	10043	A0043	4	1	GBMS KEL	1	1

1004305	1	10043	A0043	5	1	GBMS KEL	1	1
1004306	1	10043	A0043	6	1	GBMS KEL	1	1
1004307	1	10043	A0043	7	1	GBMS KEL	1	1
1004308	1	10043	A0043	8	1	GBMS KEL	1	1
1004309	1	10043	A0043	9	1	GBMS KEL	1	1
1004310	1	10043	A0043	10	1	GBMS KEL	1	1
1004311	1	10043	A0043	11	1	GBMS KEL	1	1
1004312	1	10043	A0043	12	1	GBMS KEL	1	1
1004313	1	10043	A0043	13	1	GBMS KEL	1	1
1004314	1	10043	A0043	14	1	GBMS KEL	1	1
1004315	1	10043	A0043	15	1	GBMS KEL	1	1
1004316	1	10043	A0043	16	1	GBMS KEL	1	1
1004317	1	10043	A0043	17	1	GBMS KEL	1	1
1004318	1	10043	A0043	18	1	GBMS KEL	1	1
1004319	1	10043	A0043	19	1	GBMS KEL	1	1
1004320	1	10043	A0043	20	1	GBMS KEL	1	1
1004321	1	10043	A0043	21	1	GBMS KEL	1	1
1004322	1	10043	A0043	22	1	GBMS KEL	1	1
1004401	1	10044	A0044	1	1	GBHS BHARING	1	1
1004402	1	10044	A0044	2	1	GBHS BHARING	1	1
1004403	1	10044	A0044	3	1	GBHS BHARING	1	1
1004404	1	10044	A0044	4	1	GBHS BHARING	1	1

1004405	1	10044	A0044	5	1	GBHS BHARING	1	1
1004406	1	10044	A0044	6	1	GBHS BHARING	1	1
1004407	1	10044	A0044	7	1	GBHS BHARING	1	1
1004408	1	10044	A0044	8	1	GBHS BHARING	1	1
1004409	1	10044	A0044	9	1	GBHS BHARING	1	1
1004410	1	10044	A0044	10	1	GBHS BHARING	1	1
1004411	1	10044	A0044	11	1	GBHS BHARING	1	1
1065101	1	10651	A0651	1	2	Green Wood Public School Gojra	3	2
1065102	1	10651	A0651	2	2	Green Wood Public School Gojra	3	2
1065103	1	10651	A0651	3	2	Green Wood Public School Gojra	3	2
1065104	1	10651	A0651	4	2	Green Wood Public School Gojra	3	2
1065105	1	10651	A0651	5	2	Green Wood Public School Gojra	3	2
1065106	1	10651	A0651	6	2	Green Wood Public School Gojra	3	2
1065107	1	10651	A0651	7	2	Green Wood Public School Gojra	3	2
1065108	1	10651	A0651	8	2	Green Wood Public School Gojra	3	2
1065109	1	10651	A0651	9	2	Green Wood Public School Gojra	3	2
1065110	1	10651	A0651	10	2	Green Wood Public School Gojra	3	2

1065111	1	10651	A0651	11	2	Green Wood Public School Gojra	3	2
1065112	1	10651	A0651	12	2	Green Wood Public School Gojra	3	2
1065113	1	10651	A0651	13	2	Green Wood Public School Gojra	3	2
1065114	1	10651	A0651	14	2	Green Wood Public School Gojra	3	2
1065115	1	10651	A0651	15	2	Green Wood Public School Gojra	3	2
1065116	1	10651	A0651	16	2	Green Wood Public School Gojra	3	2
1065117	1	10651	A0651	17	2	Green Wood Public School Gojra	3	2
1065118	1	10651	A0651	18	2	Green Wood Public School Gojra	3	2
1065119	1	10651	A0651	19	2	Green Wood Public School Gojra	3	2
1065121	1	10651	A0651	21	2	Green Wood Public School Gojra	3	2
1065122	1	10651	A0651	22	2	Green Wood Public School Gojra	3	2
1065201	1	10652	A0652	1	2	Shaheen Model School Chatter Domel	3	2
1065202	1	10652	A0652	2	2	Shaheen Model School Chatter Domel	3	2
1065203	1	10652	A0652	3	2	Shaheen Model School Chatter Domel	3	2

1065204	1	10652	A0652	4	2	Shaheen Model School Chatter Domel	3	2
1065205	1	10652	A0652	5	2	Shaheen Model School Chatter Domel	3	2
1065206	1	10652	A0652	6	2	Shaheen Model School Chatter Domel	3	2
1065207	1	10652	A0652	7	2	Shaheen Model School Chatter Domel	3	2
1065208	1	10652	A0652	8	2	Shaheen Model School Chatter Domel	3	2
1065209	1	10652	A0652	9	2	Shaheen Model School Chatter Domel	3	2
1065210	1	10652	A0652	10	2	Shaheen Model School Chatter Domel	3	2
1065211	1	10652	A0652	11	2	Shaheen Model School Chatter Domel	3	2
1065212	1	10652	A0652	12	2	Shaheen Model School Chatter Domel	3	2
1065213	1	10652	A0652	13	2	Shaheen Model School Chatter Domel	3	2
1065214	1	10652	A0652	14	2	Shaheen Model School Chatter Domel	3	2
1065215	1	10652	A0652	15	2	Shaheen Model School Chatter Domel	3	2
1065216	1	10652	A0652	16	2	Shaheen Model School Chatter Domel	3	2
1065217	1	10652	A0652	17	2	Shaheen Model School Chatter Domel	3	2

1065218	1	10652	A0652	18	2	Shaheen Model School Chatter Domel	3	2
1065219	1	10652	A0652	19	2	Shaheen Model School Chatter Domel	3	2
1065221	1	10652	A0652	21	2	Shaheen Model School Chatter Domel	3	2
1065222	1	10652	A0652	22	2	Shaheen Model School Chatter Domel	3	2
1065301	1	10653	A0653	1	1	BECS School Rajgran Komi Kot	3	3
1065302	1	10653	A0653	2	1	BECS School Rajgran Komi Kot	3	3
1065303	1	10653	A0653	3	1	BECS School Rajgran Komi Kot	3	3
1065304	1	10653	A0653	4	1	BECS School Rajgran Komi Kot	3	3
1065305	1	10653	A0653	5	1	BECS School Rajgran Komi Kot	3	3
1065306	1	10653	A0653	6	1	BECS School Rajgran Komi Kot	3	3
1065307	1	10653	A0653	7	1	BECS School Rajgran Komi Kot	3	3
1065308	1	10653	A0653	8	1	BECS School Rajgran Komi Kot	3	3
1065309	1	10653	A0653	9	1	BECS School Rajgran Komi Kot	3	3
1065310	1	10653	A0653	10	1	BECS School Rajgran Komi Kot	3	3

1065311	1	10653	A0653	11	1	BECS School Rajgran Komi Kot	3	3
1065312	1	10653	A0653	12	1	BECS School Rajgran Komi Kot	3	3
1065313	1	10653	A0653	13	1	BECS School Rajgran Komi Kot	3	3
1065314	1	10653	A0653	14	1	BECS School Rajgran Komi Kot	3	3
1065321	1	10653	A0653	21	1	BECS School Rajgran Komi Kot	3	3
1065322	1	10653	A0653	22	1	BECS School Rajgran Komi Kot	3	3
1065401	1	10654	A0654	1	1	BECS School Riat Bala Rahim Kot	3	3
1065402	1	10654	A0654	2	1	BECS School Riat Bala Rahim Kot	3	3
1065403	1	10654	A0654	3	1	BECS School Riat Bala Rahim Kot	3	3
1065404	1	10654	A0654	4	1	BECS School Riat Bala Rahim Kot	3	3
1065405	1	10654	A0654	5	1	BECS School Riat Bala Rahim Kot	3	3
1065406	1	10654	A0654	6	1	BECS School Riat Bala Rahim Kot	3	3
1065407	1	10654	A0654	7	1	BECS School Riat Bala Rahim Kot	3	3
1065421	1	10654	A0654	21	1	BECS School Riat Bala Rahim Kot	3	3

1065422	1	10654	A0654	22	1	BECS School Riat Bala Rahim Kot	3	3
1065501	1	10655	A0655	1	2	Islamia Public School Muzaffarbad	3	2
1065502	1	10655	A0655	2	2	Islamia Public School Muzaffarbad	3	2
1065503	1	10655	A0655	3	2	Islamia Public School Muzaffarbad	3	2
1065504	1	10655	A0655	4	2	Islamia Public School Muzaffarbad	3	2
1065505	1	10655	A0655	5	2	Islamia Public School Muzaffarbad	3	2
1065506	1	10655	A0655	6	2	Islamia Public School Muzaffarbad	3	2
1065507	1	10655	A0655	7	2	Islamia Public School Muzaffarbad	3	2
1065508	1	10655	A0655	8	2	Islamia Public School Muzaffarbad	3	2
1065509	1	10655	A0655	9	2	Islamia Public School Muzaffarbad	3	2
1065510	1	10655	A0655	10	2	Islamia Public School Muzaffarbad	3	2
1065511	1	10655	A0655	11	2	Islamia Public School Muzaffarbad	3	2
1065512	1	10655	A0655	12	2	Islamia Public School Muzaffarbad	3	2
1065513	1	10655	A0655	13	2	Islamia Public School Muzaffarbad	3	2

1065514	1	10655	A0655	14	2	Islamia Public School Muzaffarbad	3	2
1065515	1	10655	A0655	15	2	Islamia Public School Muzaffarbad	3	2
1065516	1	10655	A0655	16	2	Islamia Public School Muzaffarbad	3	2
1065517	1	10655	A0655	17	2	Islamia Public School Muzaffarbad	3	2
1065518	1	10655	A0655	18	2	Islamia Public School Muzaffarbad	3	2
1065519	1	10655	A0655	19	2	Islamia Public School Muzaffarbad	3	2
1065520	1	10655	A0655	20	2	Islamia Public School Muzaffarbad	3	2
1065521	1	10655	A0655	21	2	Islamia Public School Muzaffarbad	3	2
1065522	1	10655	A0655	22	2	Islamia Public School Muzaffarbad	3	2
5011201	5	50112	I 0112	1	1	IMSG (I-VIII), BOBRI	2	1
5011202	5	50112	I 0112	2	1	IMSG (I-VIII), BOBRI	2	1
5011203	5	50112	I 0112	3	1	IMSG (I-VIII), BOBRI	2	1
5011204	5	50112	I 0112	4	1	IMSG (I-VIII), BOBRI	2	1
5011205	5	50112	I 0112	5	1	IMSG (I-VIII), BOBRI	2	1
5011206	5	50112	I 0112	6	1	IMSG (I-VIII), BOBRI	2	1
5011207	5	50112	I 0112	7	1	IMSG (I-VIII), BOBRI	2	1
5011208	5	50112	I 0112	8	1	IMSG (I-VIII), BOBRI	2	1

5011209	5	50112	I 0112	9	1	IMSG (I-VIII), BOBRI	2	1
5011210	5	50112	I 0112	10	1	IMSG (I-VIII), BOBRI	2	1
5011211	5	50112	I 0112	11	1	IMSG (I-VIII), BOBRI	2	1
5011212	5	50112	I 0112	12	1	IMSG (I-VIII), BOBRI	2	1
5011213	5	50112	I 0112	13	1	IMSG (I-VIII), BOBRI	2	1
5011214	5	50112	I 0112	14	1	IMSG (I-VIII), BOBRI	2	1
5011215	5	50112	I 0112	15	1	IMSG (I-VIII), BOBRI	2	1
5011216	5	50112	I 0112	16	1	IMSG (I-VIII), BOBRI	2	1
5011217	5	50112	I 0112	17	1	IMSG (I-VIII), BOBRI	2	1
5011218	5	50112	I 0112	18	1	IMSG (I-VIII), BOBRI	2	1
5011219	5	50112	I 0112	19	1	IMSG (I-VIII), BOBRI	2	1
5011220	5	50112	I 0112	20	1	IMSG (I-VIII), BOBRI	2	1
5011221	5	50112	I 0112	21	1	IMSG (I-VIII), BOBRI	2	1
5011222	5	50112	I 0112	22	1	IMSG (I-VIII), BOBRI	2	1
5011301	5	50113	I 0113	1	1	IMSB (I-V), PIND PARIAN ,	1	1
5011302	5	50113	I 0113	2	1	IMSB (I-V), PIND PARIAN ,	1	1
5011303	5	50113	I 0113	3	1	IMSB (I-V), PIND PARIAN ,	1	1
5011304	5	50113	I 0113	4	1	IMSB (I-V), PIND PARIAN ,	1	1
5011305	5	50113	I 0113	5	1	IMSB (I-V), PIND PARIAN ,	1	1
5011306	5	50113	I 0113	6	1	IMSB (I-V), PIND PARIAN ,	1	1
5011307	5	50113	I 0113	7	1	IMSB (I-V), PIND PARIAN ,	1	1
5011308	5	50113	I 0113	8	1	IMSB (I-V), PIND PARIAN ,	1	1

5011309	5	50113	I 0113	9	1	IMSB (I-V), PIND PARIAN ,	1	1
5011310	5	50113	I 0113	10	1	IMSB (I-V), PIND PARIAN ,	1	1
5011311	5	50113	I 0113	11	1	IMSB (I-V), PIND PARIAN ,	1	1
5011312	5	50113	I 0113	12	1	IMSB (I-V), PIND PARIAN ,	1	1
5011313	5	50113	I 0113	13	1	IMSB (I-V), PIND PARIAN ,	1	1
5011314	5	50113	I 0113	14	1	IMSB (I-V), PIND PARIAN ,	1	1
5011315	5	50113	I 0113	15	1	IMSB (I-V), PIND PARIAN ,	1	1
5011316	5	50113	I 0113	16	1	IMSB (I-V), PIND PARIAN ,	1	1
5011317	5	50113	I 0113	17	1	IMSB (I-V), PIND PARIAN ,	1	1
5011318	5	50113	I 0113	18	1	IMSB (I-V), PIND PARIAN ,	1	1
5011319	5	50113	I 0113	19	1	IMSB (I-V), PIND PARIAN ,	1	1
5011320	5	50113	I 0113	20	1	IMSB (I-V), PIND PARIAN ,	1	1
5011321	5	50113	I 0113	21	1	IMSB (I-V), PIND PARIAN ,	1	1
5011322	5	50113	I 0113	22	1	IMSB (I-V), PIND PARIAN ,	1	1
5011401	5	50114	I 0114	1	1	IMCG (I-XII) MAIRA BEGWAL	2	1
5011402	5	50114	I 0114	2	1	IMCG (I-XII) MAIRA BEGWAL	2	1
5011403	5	50114	I 0114	3	1	IMCG (I-XII) MAIRA BEGWAL	2	1
5011404	5	50114	I 0114	4	1	IMCG (I-XII) MAIRA BEGWAL	2	1
5011405	5	50114	I 0114	5	1	IMCG (I-XII) MAIRA BEGWAL	2	1
5011406	5	50114	I 0114	6	1	IMCG (I-XII) MAIRA BEGWAL	2	1
5011407	5	50114	I 0114	7	1	IMCG (I-XII) MAIRA BEGWAL	2	1
5011408	5	50114	I 0114	8	1	IMCG (I-XII) MAIRA BEGWAL	2	1

5011409	5	50114	I 0114	9	1	IMCG (I-XII) MAIRA BEGWAL	2	1
5011410	5	50114	I 0114	10	1	IMCG (I-XII) MAIRA BEGWAL	2	1
5011411	5	50114	I 0114	11	1	IMCG (I-XII) MAIRA BEGWAL	2	1
5011412	5	50114	I 0114	12	1	IMCG (I-XII) MAIRA BEGWAL	2	1
5011413	5	50114	I 0114	13	1	IMCG (I-XII) MAIRA BEGWAL	2	1
5011414	5	50114	I 0114	14	1	IMCG (I-XII) MAIRA BEGWAL	2	1
5011415	5	50114	I 0114	15	1	IMCG (I-XII) MAIRA BEGWAL	2	1
5011416	5	50114	I 0114	16	1	IMCG (I-XII) MAIRA BEGWAL	2	1
5011417	5	50114	I 0114	17	1	IMCG (I-XII) MAIRA BEGWAL	2	1
5011418	5	50114	I 0114	18	1	IMCG (I-XII) MAIRA BEGWAL	2	1
5011419	5	50114	I 0114	19	1	IMCG (I-XII) MAIRA BEGWAL	2	1
5011420	5	50114	I 0114	20	1	IMCG (I-XII) MAIRA BEGWAL	2	1
5011421	5	50114	I 0114	21	1	IMCG (I-XII) MAIRA BEGWAL	2	1
5011422	5	50114	I 0114	22	1	IMCG (I-XII) MAIRA BEGWAL	2	1
5011501	5	50115	I 0115	1	1	IMSB (I-V), TALHAR	1	1
5011502	5	50115	I 0115	2	1	IMSB (I-V), TALHAR	1	1
5011503	5	50115	I 0115	3	1	IMSB (I-V), TALHAR	1	1
5011504	5	50115	I 0115	4	1	IMSB (I-V), TALHAR	1	1
5011505	5	50115	I 0115	5	1	IMSB (I-V), TALHAR	1	1
5011506	5	50115	I 0115	6	1	IMSB (I-V), TALHAR	1	1
5011507	5	50115	I 0115	7	1	IMSB (I-V), TALHAR	1	1
5011508	5	50115	I 0115	8	1	IMSB (I-V), TALHAR	1	1

5011509	5	50115	I 0115	9	1	IMSB (I-V), TALHAR	1	1
5011510	5	50115	I 0115	10	1	IMSB (I-V), TALHAR	1	1
5011511	5	50115	I 0115	11	1	IMSB (I-V), TALHAR	1	1
5011512	5	50115	I 0115	12	1	IMSB (I-V), TALHAR	1	1
5011513	5	50115	I 0115	13	1	IMSB (I-V), TALHAR	1	1
5011514	5	50115	I 0115	14	1	IMSB (I-V), TALHAR	1	1
5011515	5	50115	I 0115	15	1	IMSB (I-V), TALHAR	1	1
5011516	5	50115	I 0115	16	1	IMSB (I-V), TALHAR	1	1
5011517	5	50115	I 0115	17	1	IMSB (I-V), TALHAR	1	1
5011518	5	50115	I 0115	18	1	IMSB (I-V), TALHAR	1	1
5011519	5	50115	I 0115	19	1	IMSB (I-V), TALHAR	1	1
5011520	5	50115	I 0115	20	1	IMSB (I-V), TALHAR	1	1
5011521	5	50115	I 0115	21	1	IMSB (I-V), TALHAR	1	1
5011601	5	50116	I 0116	1	1	IMSB (I-X) MAIRA AKKU	1	1
5011602	5	50116	I 0116	2	1	IMSB (I-X) MAIRA AKKU	1	1
5011603	5	50116	I 0116	3	1	IMSB (I-X) MAIRA AKKU	1	1
5011604	5	50116	I 0116	4	1	IMSB (I-X) MAIRA AKKU	1	1
5011605	5	50116	I 0116	5	1	IMSB (I-X) MAIRA AKKU	1	1
5011606	5	50116	I 0116	6	1	IMSB (I-X) MAIRA AKKU	1	1
5011607	5	50116	I 0116	7	1	IMSB (I-X) MAIRA AKKU	1	1
5011608	5	50116	I 0116	8	1	IMSB (I-X) MAIRA AKKU	1	1
5011609	5	50116	I 0116	9	1	IMSB (I-X) MAIRA AKKU	1	1

5011610	5	50116	I 0116	10	1	IMSB (I-X) MAIRA AKKU	1	1
5011611	5	50116	I 0116	11	1	IMSB (I-X) MAIRA AKKU	1	1
5011612	5	50116	I 0116	12	1	IMSB (I-X) MAIRA AKKU	1	1
5011613	5	50116	I 0116	13	1	IMSB (I-X) MAIRA AKKU	1	1
5011614	5	50116	I 0116	14	1	IMSB (I-X) MAIRA AKKU	1	1
5011615	5	50116	I 0116	15	1	IMSB (I-X) MAIRA AKKU	1	1
5011616	5	50116	I 0116	16	1	IMSB (I-X) MAIRA AKKU	1	1
5011617	5	50116	I 0116	17	1	IMSB (I-X) MAIRA AKKU	1	1
5011618	5	50116	I 0116	18	1	IMSB (I-X) MAIRA AKKU	1	1
5011619	5	50116	I 0116	19	1	IMSB (I-X) MAIRA AKKU	1	1
5011620	5	50116	I 0116	20	1	IMSB (I-X) MAIRA AKKU	1	1
5011621	5	50116	I 0116	21	1	IMSB (I-X) MAIRA AKKU	1	1
5011701	5	50117	I 0117	1	1	IMSB (I-VIII), KOT HATHIAL (N/A)	1	1
5011702	5	50117	I 0117	2	1	IMSB (I-VIII), KOT HATHIAL (N/A)	1	1
5011703	5	50117	I 0117	3	1	IMSB (I-VIII), KOT HATHIAL (N/A)	1	1
5011704	5	50117	I 0117	4	1	IMSB (I-VIII), KOT HATHIAL (N/A)	1	1
5011705	5	50117	I 0117	5	1	IMSB (I-VIII), KOT HATHIAL (N/A)	1	1
5011706	5	50117	I 0117	6	1	IMSB (I-VIII), KOT HATHIAL (N/A)	1	1

5011707	5	50117	I 0117	7	1	IMSB (I-VIII), KOT HATHIAL (N/A)	1	1
5011708	5	50117	I 0117	8	1	IMSB (I-VIII), KOT HATHIAL (N/A)	1	1
5011709	5	50117	I 0117	9	1	IMSB (I-VIII), KOT HATHIAL (N/A)	1	1
5011710	5	50117	I 0117	10	1	IMSB (I-VIII), KOT HATHIAL (N/A)	1	1
5011711	5	50117	I 0117	11	1	IMSB (I-VIII), KOT HATHIAL (N/A)	1	1
5011712	5	50117	I 0117	12	1	IMSB (I-VIII), KOT HATHIAL (N/A)	1	1
5011713	5	50117	I 0117	13	1	IMSB (I-VIII), KOT HATHIAL (N/A)	1	1
5011714	5	50117	I 0117	14	1	IMSB (I-VIII), KOT HATHIAL (N/A)	1	1
5011715	5	50117	I 0117	15	1	IMSB (I-VIII), KOT HATHIAL (N/A)	1	1
5011716	5	50117	I 0117	16	1	IMSB (I-VIII), KOT HATHIAL (N/A)	1	1
5011717	5	50117	I 0117	17	1	IMSB (I-VIII), KOT HATHIAL (N/A)	1	1
5011718	5	50117	I 0117	18	1	IMSB (I-VIII), KOT HATHIAL (N/A)	1	1
5011719	5	50117	I 0117	19	1	IMSB (I-VIII), KOT HATHIAL (N/A)	1	1
5011720	5	50117	I 0117	20	1	IMSB (I-VIII), KOT HATHIAL (N/A)	1	1

5011721	5	50117	I 0117	21	1	IMSB (I-VIII), KOT HATHIAL (N/A)	1	1
5011801	5	50118	I 0118	1	1	IMSG (I-X) BHADANA KALAN ,	2	1
5011802	5	50118	I 0118	2	1	IMSG (I-X) BHADANA KALAN ,	2	1
5011803	5	50118	I 0118	3	1	IMSG (I-X) BHADANA KALAN ,	2	1
5011804	5	50118	I 0118	4	1	IMSG (I-X) BHADANA KALAN ,	2	1
5011805	5	50118	I 0118	5	1	IMSG (I-X) BHADANA KALAN ,	2	1
5011806	5	50118	I 0118	6	1	IMSG (I-X) BHADANA KALAN ,	2	1
5011807	5	50118	I 0118	7	1	IMSG (I-X) BHADANA KALAN ,	2	1
5011808	5	50118	I 0118	8	1	IMSG (I-X) BHADANA KALAN ,	2	1
5011809	5	50118	I 0118	9	1	IMSG (I-X) BHADANA KALAN ,	2	1
5011810	5	50118	I 0118	10	1	IMSG (I-X) BHADANA KALAN ,	2	1
5011811	5	50118	I 0118	11	1	IMSG (I-X) BHADANA KALAN ,	2	1
5011812	5	50118	I 0118	12	1	IMSG (I-X) BHADANA KALAN ,	2	1
5011813	5	50118	I 0118	13	1	IMSG (I-X) BHADANA KALAN ,	2	1

5011814	5	50118	I 0118	14	1	IMSG (I-X) BHADANA KALAN ,	2	1
5011815	5	50118	I 0118	15	1	IMSG (I-X) BHADANA KALAN ,	2	1
5011816	5	50118	I 0118	16	1	IMSG (I-X) BHADANA KALAN ,	2	1
5011817	5	50118	I 0118	17	1	IMSG (I-X) BHADANA KALAN ,	2	1
5011818	5	50118	I 0118	18	1	IMSG (I-X) BHADANA KALAN ,	2	1
5011819	5	50118	I 0118	19	1	IMSG (I-X) BHADANA KALAN ,	2	1
5011820	5	50118	I 0118	20	1	IMSG (I-X) BHADANA KALAN ,	2	1
5011901	5	50119	I 0119	1	1	IMSB (I-X), KHANNA NAI ABADI	1	1
5011902	5	50119	I 0119	2	1	IMSB (I-X), KHANNA NAI ABADI	1	1
5011903	5	50119	I 0119	3	1	IMSB (I-X), KHANNA NAI ABADI	1	1
5011904	5	50119	I 0119	4	1	IMSB (I-X), KHANNA NAI ABADI	1	1
5011905	5	50119	I 0119	5	1	IMSB (I-X), KHANNA NAI ABADI	1	1
5011906	5	50119	I 0119	6	1	IMSB (I-X), KHANNA NAI ABADI	1	1
5011907	5	50119	I 0119	7	1	IMSB (I-X), KHANNA NAI ABADI	1	1

5011908	5	50119	I 0119	8	1	IMSB (I-X), KHANNA NAI ABADI	1	1
5011909	5	50119	I 0119	9	1	IMSB (I-X), KHANNA NAI ABADI	1	1
5011910	5	50119	I 0119	10	1	IMSB (I-X), KHANNA NAI ABADI	1	1
5011911	5	50119	I 0119	11	1	IMSB (I-X), KHANNA NAI ABADI	1	1
5011912	5	50119	I 0119	12	1	IMSB (I-X), KHANNA NAI ABADI	1	1
5011913	5	50119	I 0119	13	1	IMSB (I-X), KHANNA NAI ABADI	1	1
5011914	5	50119	I 0119	14	1	IMSB (I-X), KHANNA NAI ABADI	1	1
5011915	5	50119	I 0119	15	1	IMSB (I-X), KHANNA NAI ABADI	1	1
5011916	5	50119	I 0119	16	1	IMSB (I-X), KHANNA NAI ABADI	1	1
5011917	5	50119	I 0119	17	1	IMSB (I-X), KHANNA NAI ABADI	1	1
5011918	5	50119	I 0119	18	1	IMSB (I-X), KHANNA NAI ABADI	1	1
5011919	5	50119	I 0119	19	1	IMSB (I-X), KHANNA NAI ABADI	1	1
5011920	5	50119	I 0119	20	1	IMSB (I-X), KHANNA NAI ABADI	1	1
5011921	5	50119	I 0119	21	1	IMSB (I-X), KHANNA NAI ABADI	1	1
5012001	5	50120	I 0120	1	1	IMSG (I-X), DHOKE GANGAL	2	1

5012002	5	50120	I 0120	2	1	IMSG (I-X), DHOKE GANGAL	2	1
5012003	5	50120	I 0120	3	1	IMSG (I-X), DHOKE GANGAL	2	1
5012004	5	50120	I 0120	4	1	IMSG (I-X), DHOKE GANGAL	2	1
5012005	5	50120	I 0120	5	1	IMSG (I-X), DHOKE GANGAL	2	1
5012006	5	50120	I 0120	6	1	IMSG (I-X), DHOKE GANGAL	2	1
5012007	5	50120	I 0120	7	1	IMSG (I-X), DHOKE GANGAL	2	1
5012008	5	50120	I 0120	8	1	IMSG (I-X), DHOKE GANGAL	2	1
5012009	5	50120	I 0120	9	1	IMSG (I-X), DHOKE GANGAL	2	1
5012010	5	50120	I 0120	10	1	IMSG (I-X), DHOKE GANGAL	2	1
5012011	5	50120	I 0120	11	1	IMSG (I-X), DHOKE GANGAL	2	1
5012012	5	50120	I 0120	12	1	IMSG (I-X), DHOKE GANGAL	2	1
5012013	5	50120	I 0120	13	1	IMSG (I-X), DHOKE GANGAL	2	1
5012014	5	50120	I 0120	14	1	IMSG (I-X), DHOKE GANGAL	2	1
5012015	5	50120	I 0120	15	1	IMSG (I-X), DHOKE GANGAL	2	1
5012016	5	50120	I 0120	16	1	IMSG (I-X), DHOKE GANGAL	2	1
5012017	5	50120	I 0120	17	1	IMSG (I-X), DHOKE GANGAL	2	1
5012018	5	50120	I 0120	18	1	IMSG (I-X), DHOKE GANGAL	2	1
5012019	5	50120	I 0120	19	1	IMSG (I-X), DHOKE GANGAL	2	1
5012020	5	50120	I 0120	20	1	IMSG (I-X), DHOKE GANGAL	2	1
5012021	5	50120	I 0120	21	1	IMSG (I-X), DHOKE GANGAL	2	1
5012022	5	50120	I 0120	22	1	IMSG (I-X), DHOKE GANGAL	2	1
5012101	5	50121	I 0121	1	2	IMSJ (I-V), G-9/3 (No-1)	3	1

5012102	5	50121	I 0121	2	2	IMSJ (I-V), G-9/3 (No-1)	3	1
5012103	5	50121	I 0121	3	2	IMSJ (I-V), G-9/3 (No-1)	3	1
5012104	5	50121	I 0121	4	2	IMSJ (I-V), G-9/3 (No-1)	3	1
5012105	5	50121	I 0121	5	2	IMSJ (I-V), G-9/3 (No-1)	3	1
5012106	5	50121	I 0121	6	2	IMSJ (I-V), G-9/3 (No-1)	3	1
5012107	5	50121	I 0121	7	2	IMSJ (I-V), G-9/3 (No-1)	3	1
5012108	5	50121	I 0121	8	2	IMSJ (I-V), G-9/3 (No-1)	3	1
5012109	5	50121	I 0121	9	2	IMSJ (I-V), G-9/3 (No-1)	3	1
5012110	5	50121	I 0121	10	2	IMSJ (I-V), G-9/3 (No-1)	3	1
5012111	5	50121	I 0121	11	2	IMSJ (I-V), G-9/3 (No-1)	3	1
5012112	5	50121	I 0121	12	2	IMSJ (I-V), G-9/3 (No-1)	3	1
5012113	5	50121	I 0121	13	2	IMSJ (I-V), G-9/3 (No-1)	3	1
5012114	5	50121	I 0121	14	2	IMSJ (I-V), G-9/3 (No-1)	3	1
5012115	5	50121	I 0121	15	2	IMSJ (I-V), G-9/3 (No-1)	3	1
5012116	5	50121	I 0121	16	2	IMSJ (I-V), G-9/3 (No-1)	3	1
5012117	5	50121	I 0121	17	2	IMSJ (I-V), G-9/3 (No-1)	3	1
5012118	5	50121	I 0121	18	2	IMSJ (I-V), G-9/3 (No-1)	3	1
5012119	5	50121	I 0121	19	2	IMSJ (I-V), G-9/3 (No-1)	3	1
5012120	5	50121	I 0121	20	2	IMSJ (I-V), G-9/3 (No-1)	3	1
5012121	5	50121	I 0121	21	2	IMSJ (I-V), G-9/3 (No-1)	3	1
5012201	5	50122	I 0122	1	2	IMSJ (I-V), G-10/2	3	1
5012202	5	50122	I 0122	2	2	IMSJ (I-V), G-10/2	3	1

5012203	5	50122	I 0122	3	2	IMSJ (I-V), G-10/2	3	1
5012204	5	50122	I 0122	4	2	IMSJ (I-V), G-10/2	3	1
5012205	5	50122	I 0122	5	2	IMSJ (I-V), G-10/2	3	1
5012206	5	50122	I 0122	6	2	IMSJ (I-V), G-10/2	3	1
5012207	5	50122	I 0122	7	2	IMSJ (I-V), G-10/2	3	1
5012208	5	50122	I 0122	8	2	IMSJ (I-V), G-10/2	3	1
5012209	5	50122	I 0122	9	2	IMSJ (I-V), G-10/2	3	1
5012210	5	50122	I 0122	10	2	IMSJ (I-V), G-10/2	3	1
5012211	5	50122	I 0122	11	2	IMSJ (I-V), G-10/2	3	1
5012212	5	50122	I 0122	12	2	IMSJ (I-V), G-10/2	3	1
5012213	5	50122	I 0122	13	2	IMSJ (I-V), G-10/2	3	1
5012214	5	50122	I 0122	14	2	IMSJ (I-V), G-10/2	3	1
5012215	5	50122	I 0122	15	2	IMSJ (I-V), G-10/2	3	1
5012216	5	50122	I 0122	16	2	IMSJ (I-V), G-10/2	3	1
5012217	5	50122	I 0122	17	2	IMSJ (I-V), G-10/2	3	1
5012218	5	50122	I 0122	18	2	IMSJ (I-V), G-10/2	3	1
5012219	5	50122	I 0122	19	2	IMSJ (I-V), G-10/2	3	1
5012220	5	50122	I 0122	20	2	IMSJ (I-V), G-10/2	3	1
5012221	5	50122	I 0122	21	2	IMSJ (I-V), G-10/2	3	1
5012222	5	50122	I 0122	22	2	IMSJ (I-V), G-10/2	3	1
5012301	5	50123	I 0123	1	2	IMSJ (I-V), G-8/1 (No-3)	3	1
5012302	5	50123	I 0123	2	2	IMSJ (I-V), G-8/1 (No-3)	3	1

5012303	5	50123	I 0123	3	2	IMSJ (I-V), G-8/1 (No-3)	3	1
5012304	5	50123	I 0123	4	2	IMSJ (I-V), G-8/1 (No-3)	3	1
5012305	5	50123	I 0123	5	2	IMSJ (I-V), G-8/1 (No-3)	3	1
5012306	5	50123	I 0123	6	2	IMSJ (I-V), G-8/1 (No-3)	3	1
5012307	5	50123	I 0123	7	2	IMSJ (I-V), G-8/1 (No-3)	3	1
5012308	5	50123	I 0123	8	2	IMSJ (I-V), G-8/1 (No-3)	3	1
5012309	5	50123	I 0123	9	2	IMSJ (I-V), G-8/1 (No-3)	3	1
5012310	5	50123	I 0123	10	2	IMSJ (I-V), G-8/1 (No-3)	3	1
5012311	5	50123	I 0123	11	2	IMSJ (I-V), G-8/1 (No-3)	3	1
5012312	5	50123	I 0123	12	2	IMSJ (I-V), G-8/1 (No-3)	3	1
5012313	5	50123	I 0123	13	2	IMSJ (I-V), G-8/1 (No-3)	3	1
5012314	5	50123	I 0123	14	2	IMSJ (I-V), G-8/1 (No-3)	3	1
5012315	5	50123	I 0123	15	2	IMSJ (I-V), G-8/1 (No-3)	3	1
5012316	5	50123	I 0123	16	2	IMSJ (I-V), G-8/1 (No-3)	3	1
5012317	5	50123	I 0123	17	2	IMSJ (I-V), G-8/1 (No-3)	3	1
5012318	5	50123	I 0123	18	2	IMSJ (I-V), G-8/1 (No-3)	3	1
5012319	5	50123	I 0123	19	2	IMSJ (I-V), G-8/1 (No-3)	3	1
5012320	5	50123	I 0123	20	2	IMSJ (I-V), G-8/1 (No-3)	3	1
5012321	5	50123	I 0123	21	2	IMSJ (I-V), G-8/1 (No-3)	3	1
5012322	5	50123	I 0123	22	2	IMSJ (I-V), G-8/1 (No-3)	3	1
5012401	5	50124	I 0124	1	2	IMSJ (I-V), F-6/4,	3	1
5012402	5	50124	I 0124	2	2	IMSJ (I-V), F-6/4,	3	1

5012403	5	50124	I 0124	3	2	IMSJ (I-V), F-6/4,	3	1
5012404	5	50124	I 0124	4	2	IMSJ (I-V), F-6/4,	3	1
5012405	5	50124	I 0124	5	2	IMSJ (I-V), F-6/4,	3	1
5012406	5	50124	I 0124	6	2	IMSJ (I-V), F-6/4,	3	1
5012407	5	50124	I 0124	7	2	IMSJ (I-V), F-6/4,	3	1
5012408	5	50124	I 0124	8	2	IMSJ (I-V), F-6/4,	3	1
5012409	5	50124	I 0124	9	2	IMSJ (I-V), F-6/4,	3	1
5012410	5	50124	I 0124	10	2	IMSJ (I-V), F-6/4,	3	1
5012411	5	50124	I 0124	11	2	IMSJ (I-V), F-6/4,	3	1
5012412	5	50124	I 0124	12	2	IMSJ (I-V), F-6/4,	3	1
5012413	5	50124	I 0124	13	2	IMSJ (I-V), F-6/4,	3	1
5012414	5	50124	I 0124	14	2	IMSJ (I-V), F-6/4,	3	1
5012415	5	50124	I 0124	15	2	IMSJ (I-V), F-6/4,	3	1
5012416	5	50124	I 0124	16	2	IMSJ (I-V), F-6/4,	3	1
5012417	5	50124	I 0124	17	2	IMSJ (I-V), F-6/4,	3	1
5012418	5	50124	I 0124	18	2	IMSJ (I-V), F-6/4,	3	1
5012419	5	50124	I 0124	19	2	IMSJ (I-V), F-6/4,	3	1
5012420	5	50124	I 0124	20	2	IMSJ (I-V), F-6/4,	3	1
5012421	5	50124	I 0124	21	2	IMSJ (I-V), F-6/4,	3	1
5012501	5	50125	I 0125	1	2	IMSJ (I-V), G-9/2 (No-4)	3	1
5012502	5	50125	I 0125	2	2	IMSJ (I-V), G-9/2 (No-4)	3	1
5012503	5	50125	I 0125	3	2	IMSJ (I-V), G-9/2 (No-4)	3	1

5012504	5	50125	I 0125	4	2	IMSJ (I-V), G-9/2 (No-4)	3	1
5012505	5	50125	I 0125	5	2	IMSJ (I-V), G-9/2 (No-4)	3	1
5012506	5	50125	I 0125	6	2	IMSJ (I-V), G-9/2 (No-4)	3	1
5012507	5	50125	I 0125	7	2	IMSJ (I-V), G-9/2 (No-4)	3	1
5012508	5	50125	I 0125	8	2	IMSJ (I-V), G-9/2 (No-4)	3	1
5012509	5	50125	I 0125	9	2	IMSJ (I-V), G-9/2 (No-4)	3	1
5012510	5	50125	I 0125	10	2	IMSJ (I-V), G-9/2 (No-4)	3	1
5012511	5	50125	I 0125	11	2	IMSJ (I-V), G-9/2 (No-4)	3	1
5012512	5	50125	I 0125	12	2	IMSJ (I-V), G-9/2 (No-4)	3	1
5012513	5	50125	I 0125	13	2	IMSJ (I-V), G-9/2 (No-4)	3	1
5012514	5	50125	I 0125	14	2	IMSJ (I-V), G-9/2 (No-4)	3	1
5012515	5	50125	I 0125	15	2	IMSJ (I-V), G-9/2 (No-4)	3	1
5012516	5	50125	I 0125	16	2	IMSJ (I-V), G-9/2 (No-4)	3	1
5012517	5	50125	I 0125	17	2	IMSJ (I-V), G-9/2 (No-4)	3	1
5012518	5	50125	I 0125	18	2	IMSJ (I-V), G-9/2 (No-4)	3	1
5012519	5	50125	I 0125	19	2	IMSJ (I-V), G-9/2 (No-4)	3	1
5012520	5	50125	I 0125	20	2	IMSJ (I-V), G-9/2 (No-4)	3	1
5012521	5	50125	I 0125	21	2	IMSJ (I-V), G-9/2 (No-4)	3	1
5012522	5	50125	I 0125	22	2	IMSJ (I-V), G-9/2 (No-4)	3	1
5012601	5	50126	I 0126	1	2	IMSJ (I-V), G-8/4	3	1
5012602	5	50126	I 0126	2	2	IMSJ (I-V), G-8/4	3	1
5012603	5	50126	I 0126	3	2	IMSJ (I-V), G-8/4	3	1

5012604	5	50126	I 0126	4	2	IMSJ (I-V), G-8/4	3	1
5012605	5	50126	I 0126	5	2	IMSJ (I-V), G-8/4	3	1
5012606	5	50126	I 0126	6	2	IMSJ (I-V), G-8/4	3	1
5012607	5	50126	I 0126	7	2	IMSJ (I-V), G-8/4	3	1
5012608	5	50126	I 0126	8	2	IMSJ (I-V), G-8/4	3	1
5012609	5	50126	I 0126	9	2	IMSJ (I-V), G-8/4	3	1
5012610	5	50126	I 0126	10	2	IMSJ (I-V), G-8/4	3	1
5012611	5	50126	I 0126	11	2	IMSJ (I-V), G-8/4	3	1
5012612	5	50126	I 0126	12	2	IMSJ (I-V), G-8/4	3	1
5012613	5	50126	I 0126	13	2	IMSJ (I-V), G-8/4	3	1
5012614	5	50126	I 0126	14	2	IMSJ (I-V), G-8/4	3	1
5012615	5	50126	I 0126	15	2	IMSJ (I-V), G-8/4	3	1
5012616	5	50126	I 0126	16	2	IMSJ (I-V), G-8/4	3	1
5012617	5	50126	I 0126	17	2	IMSJ (I-V), G-8/4	3	1
5012618	5	50126	I 0126	18	2	IMSJ (I-V), G-8/4	3	1
5012619	5	50126	I 0126	19	2	IMSJ (I-V), G-8/4	3	1
5012620	5	50126	I 0126	20	2	IMSJ (I-V), G-8/4	3	1
5012621	5	50126	I 0126	21	2	IMSJ (I-V), G-8/4	3	1
5012622	5	50126	I 0126	22	2	IMSJ (I-V), G-8/4	3	1
5012701	5	50127	I 0127	1	2	IMSJ (I-V), G-6/4,	3	1
5012702	5	50127	I 0127	2	2	IMSJ (I-V), G-6/4,	3	1
5012703	5	50127	I 0127	3	2	IMSJ (I-V), G-6/4,	3	1

5012704	5	50127	I 0127	4	2	IMSJ (I-V), G-6/4,	3	1
5012705	5	50127	I 0127	5	2	IMSJ (I-V), G-6/4,	3	1
5012706	5	50127	I 0127	6	2	IMSJ (I-V), G-6/4,	3	1
5012707	5	50127	I 0127	7	2	IMSJ (I-V), G-6/4,	3	1
5012708	5	50127	I 0127	8	2	IMSJ (I-V), G-6/4,	3	1
5012709	5	50127	I 0127	9	2	IMSJ (I-V), G-6/4,	3	1
5012710	5	50127	I 0127	10	2	IMSJ (I-V), G-6/4,	3	1
5012711	5	50127	I 0127	11	2	IMSJ (I-V), G-6/4,	3	1
5012712	5	50127	I 0127	12	2	IMSJ (I-V), G-6/4,	3	1
5012713	5	50127	I 0127	13	2	IMSJ (I-V), G-6/4,	3	1
5012714	5	50127	I 0127	14	2	IMSJ (I-V), G-6/4,	3	1
5012715	5	50127	I 0127	15	2	IMSJ (I-V), G-6/4,	3	1
5012716	5	50127	I 0127	16	2	IMSJ (I-V), G-6/4,	3	1
5012717	5	50127	I 0127	17	2	IMSJ (I-V), G-6/4,	3	1
5012718	5	50127	I 0127	18	2	IMSJ (I-V), G-6/4,	3	1
5012719	5	50127	I 0127	19	2	IMSJ (I-V), G-6/4,	3	1
5012720	5	50127	I 0127	20	2	IMSJ (I-V), G-6/4,	3	1
5012721	5	50127	I 0127	21	2	IMSJ (I-V), G-6/4,	3	1
5012722	5	50127	I 0127	22	2	IMSJ (I-V), G-6/4,	3	1
5012801	5	50128	I 0128	1	2	IMSG (I-VIII), I-8/1	3	1
5012802	5	50128	I 0128	2	2	IMSG (I-VIII), I-8/1	3	1
5012803	5	50128	I 0128	3	2	IMSG (I-VIII), I-8/1	3	1

5012804	5	50128	I 0128	4	2	IMSG (I-VIII), I-8/1	3	1
5012805	5	50128	I 0128	5	2	IMSG (I-VIII), I-8/1	3	1
5012806	5	50128	I 0128	6	2	IMSG (I-VIII), I-8/1	3	1
5012807	5	50128	I 0128	7	2	IMSG (I-VIII), I-8/1	3	1
5012808	5	50128	I 0128	8	2	IMSG (I-VIII), I-8/1	3	1
5012809	5	50128	I 0128	9	2	IMSG (I-VIII), I-8/1	3	1
5012810	5	50128	I 0128	10	2	IMSG (I-VIII), I-8/1	3	1
5012811	5	50128	I 0128	11	2	IMSG (I-VIII), I-8/1	3	1
5012812	5	50128	I 0128	12	2	IMSG (I-VIII), I-8/1	3	1
5012813	5	50128	I 0128	13	2	IMSG (I-VIII), I-8/1	3	1
5012814	5	50128	I 0128	14	2	IMSG (I-VIII), I-8/1	3	1
5012815	5	50128	I 0128	15	2	IMSG (I-VIII), I-8/1	3	1
5012816	5	50128	I 0128	16	2	IMSG (I-VIII), I-8/1	3	1
5012817	5	50128	I 0128	17	2	IMSG (I-VIII), I-8/1	3	1
5012818	5	50128	I 0128	18	2	IMSG (I-VIII), I-8/1	3	1
5012819	5	50128	I 0128	19	2	IMSG (I-VIII), I-8/1	3	1
5012820	5	50128	I 0128	20	2	IMSG (I-VIII), I-8/1	3	1
5012821	5	50128	I 0128	21	2	IMSG (I-VIII), I-8/1	3	1
5012822	5	50128	I 0128	22	2	IMSG (I-VIII), I-8/1	3	1
5012901	5	50129	I 0129	1	2	IMSG (I-X), G-11/2	3	1
5012902	5	50129	I 0129	2	2	IMSG (I-X), G-11/2	3	1
5012903	5	50129	I 0129	3	2	IMSG (I-X), G-11/2	3	1

5012904	5	50129	I 0129	4	2	IMSG (I-X), G-11/2	3	1
5012905	5	50129	I 0129	5	2	IMSG (I-X), G-11/2	3	1
5012906	5	50129	I 0129	6	2	IMSG (I-X), G-11/2	3	1
5012907	5	50129	I 0129	7	2	IMSG (I-X), G-11/2	3	1
5012908	5	50129	I 0129	8	2	IMSG (I-X), G-11/2	3	1
5012909	5	50129	I 0129	9	2	IMSG (I-X), G-11/2	3	1
5012910	5	50129	I 0129	10	2	IMSG (I-X), G-11/2	3	1
5012911	5	50129	I 0129	11	2	IMSG (I-X), G-11/2	3	1
5012912	5	50129	I 0129	12	2	IMSG (I-X), G-11/2	3	1
5012913	5	50129	I 0129	13	2	IMSG (I-X), G-11/2	3	1
5012914	5	50129	I 0129	14	2	IMSG (I-X), G-11/2	3	1
5012915	5	50129	I 0129	15	2	IMSG (I-X), G-11/2	3	1
5012916	5	50129	I 0129	16	2	IMSG (I-X), G-11/2	3	1
5012917	5	50129	I 0129	17	2	IMSG (I-X), G-11/2	3	1
5012918	5	50129	I 0129	18	2	IMSG (I-X), G-11/2	3	1
5012919	5	50129	I 0129	19	2	IMSG (I-X), G-11/2	3	1
5012920	5	50129	I 0129	20	2	IMSG (I-X), G-11/2	3	1
5012921	5	50129	I 0129	21	2	IMSG (I-X), G-11/2	3	1
5012922	5	50129	I 0129	22	2	IMSG (I-X), G-11/2	3	1
5013001	5	50130	I 0130	1	2	IMSJ (I-V), G-6/2	3	1
5013002	5	50130	I 0130	2	2	IMSJ (I-V), G-6/2	3	1
5013003	5	50130	I 0130	3	2	IMSJ (I-V), G-6/2	3	1

5013004	5	50130	I 0130	4	2	IMSJ (I-V), G-6/2	3	1
5013005	5	50130	I 0130	5	2	IMSJ (I-V), G-6/2	3	1
5013006	5	50130	I 0130	6	2	IMSJ (I-V), G-6/2	3	1
5013007	5	50130	I 0130	7	2	IMSJ (I-V), G-6/2	3	1
5013008	5	50130	I 0130	8	2	IMSJ (I-V), G-6/2	3	1
5013009	5	50130	I 0130	9	2	IMSJ (I-V), G-6/2	3	1
5013010	5	50130	I 0130	10	2	IMSJ (I-V), G-6/2	3	1
5013011	5	50130	I 0130	11	2	IMSJ (I-V), G-6/2	3	1
5013012	5	50130	I 0130	12	2	IMSJ (I-V), G-6/2	3	1
5013013	5	50130	I 0130	13	2	IMSJ (I-V), G-6/2	3	1
5013014	5	50130	I 0130	14	2	IMSJ (I-V), G-6/2	3	1
5013015	5	50130	I 0130	15	2	IMSJ (I-V), G-6/2	3	1
5013016	5	50130	I 0130	16	2	IMSJ (I-V), G-6/2	3	1
5013017	5	50130	I 0130	17	2	IMSJ (I-V), G-6/2	3	1
5013018	5	50130	I 0130	18	2	IMSJ (I-V), G-6/2	3	1
5013019	5	50130	I 0130	19	2	IMSJ (I-V), G-6/2	3	1
5013020	5	50130	I 0130	20	2	IMSJ (I-V), G-6/2	3	1
5013021	5	50130	I 0130	21	2	IMSJ (I-V), G-6/2	3	1
5013022	5	50130	I 0130	22	2	IMSJ (I-V), G-6/2	3	1
5013101	5	50131	I 0131	1	2	IMSJ (I-V), G-7/4,	3	1
5013102	5	50131	I 0131	2	2	IMSJ (I-V), G-7/4,	3	1
5013103	5	50131	I 0131	3	2	IMSJ (I-V), G-7/4,	3	1

5013104	5	50131	I 0131	4	2	IMSJ (I-V), G-7/4,	3	1
5013105	5	50131	I 0131	5	2	IMSJ (I-V), G-7/4,	3	1
5013106	5	50131	I 0131	6	2	IMSJ (I-V), G-7/4,	3	1
5013107	5	50131	I 0131	7	2	IMSJ (I-V), G-7/4,	3	1
5013108	5	50131	I 0131	8	2	IMSJ (I-V), G-7/4,	3	1
5013109	5	50131	I 0131	9	2	IMSJ (I-V), G-7/4,	3	1
5013110	5	50131	I 0131	10	2	IMSJ (I-V), G-7/4,	3	1
5013111	5	50131	I 0131	11	2	IMSJ (I-V), G-7/4,	3	1
5013112	5	50131	I 0131	12	2	IMSJ (I-V), G-7/4,	3	1
5013113	5	50131	I 0131	13	2	IMSJ (I-V), G-7/4,	3	1
5013114	5	50131	I 0131	14	2	IMSJ (I-V), G-7/4,	3	1
5013115	5	50131	I 0131	15	2	IMSJ (I-V), G-7/4,	3	1
5013116	5	50131	I 0131	16	2	IMSJ (I-V), G-7/4,	3	1
5013117	5	50131	I 0131	17	2	IMSJ (I-V), G-7/4,	3	1
5013118	5	50131	I 0131	18	2	IMSJ (I-V), G-7/4,	3	1
5013119	5	50131	I 0131	19	2	IMSJ (I-V), G-7/4,	3	1
5013120	5	50131	I 0131	20	2	IMSJ (I-V), G-7/4,	3	1
5013121	5	50131	I 0131	21	2	IMSJ (I-V), G-7/4,	3	1
5013122	5	50131	I 0131	22	2	IMSJ (I-V), G-7/4,	3	1
5013201	5	50132	I 0132	1	2	IMCB, I-10/1	1	1
5013202	5	50132	I 0132	2	2	IMCB, I-10/1	1	1
5013203	5	50132	I 0132	3	2	IMCB, I-10/1	1	1

5013204	5	50132	I 0132	4	2	IMCB, I-10/1	1	1
5013205	5	50132	I 0132	5	2	IMCB, I-10/1	1	1
5013206	5	50132	I 0132	6	2	IMCB, I-10/1	1	1
5013207	5	50132	I 0132	7	2	IMCB, I-10/1	1	1
5013208	5	50132	I 0132	8	2	IMCB, I-10/1	1	1
5013209	5	50132	I 0132	9	2	IMCB, I-10/1	1	1
5013210	5	50132	I 0132	10	2	IMCB, I-10/1	1	1
5013211	5	50132	I 0132	11	2	IMCB, I-10/1	1	1
5013212	5	50132	I 0132	12	2	IMCB, I-10/1	1	1
5013213	5	50132	I 0132	13	2	IMCB, I-10/1	1	1
5013214	5	50132	I 0132	14	2	IMCB, I-10/1	1	1
5013215	5	50132	I 0132	15	2	IMCB, I-10/1	1	1
5013216	5	50132	I 0132	16	2	IMCB, I-10/1	1	1
5013217	5	50132	I 0132	17	2	IMCB, I-10/1	1	1
5013218	5	50132	I 0132	18	2	IMCB, I-10/1	1	1
5013219	5	50132	I 0132	19	2	IMCB, I-10/1	1	1
5013220	5	50132	I 0132	20	2	IMCB, I-10/1	1	1
5013221	5	50132	I 0132	21	2	IMCB, I-10/1	1	1
5013222	5	50132	I 0132	22	2	IMCB, I-10/1	1	1
5013301	5	50133	I 0133	1	2	IMCG, I-8/4	2	1
5013302	5	50133	I 0133	2	2	IMCG, I-8/4	2	1
5013303	5	50133	I 0133	3	2	IMCG, I-8/4	2	1

5013304	5	50133	I 0133	4	2	IMCG, I-8/4	2	1
5013305	5	50133	I 0133	5	2	IMCG, I-8/4	2	1
5013306	5	50133	I 0133	6	2	IMCG, I-8/4	2	1
5013307	5	50133	I 0133	7	2	IMCG, I-8/4	2	1
5013308	5	50133	I 0133	8	2	IMCG, I-8/4	2	1
5013309	5	50133	I 0133	9	2	IMCG, I-8/4	2	1
5013310	5	50133	I 0133	10	2	IMCG, I-8/4	2	1
5013311	5	50133	I 0133	11	2	IMCG, I-8/4	2	1
5013312	5	50133	I 0133	12	2	IMCG, I-8/4	2	1
5013313	5	50133	I 0133	13	2	IMCG, I-8/4	2	1
5013314	5	50133	I 0133	14	2	IMCG, I-8/4	2	1
5013315	5	50133	I 0133	15	2	IMCG, I-8/4	2	1
5013316	5	50133	I 0133	16	2	IMCG, I-8/4	2	1
5013317	5	50133	I 0133	17	2	IMCG, I-8/4	2	1
5013318	5	50133	I 0133	18	2	IMCG, I-8/4	2	1
5013319	5	50133	I 0133	19	2	IMCG, I-8/4	2	1
5013320	5	50133	I 0133	20	2	IMCG, I-8/4	2	1
5013321	5	50133	I 0133	21	2	IMCG, I-8/4	2	1
5013322	5	50133	I 0133	22	2	IMCG, I-8/4	2	1
5013401	5	50134	I 0134	1	2	IMCB, F-7/3	1	1
5013402	5	50134	I 0134	2	2	IMCB, F-7/3	1	1
5013403	5	50134	I 0134	3	2	IMCB, F-7/3	1	1

5013404	5	50134	I 0134	4	2	IMCB, F-7/3	1	1
5013405	5	50134	I 0134	5	2	IMCB, F-7/3	1	1
5013406	5	50134	I 0134	6	2	IMCB, F-7/3	1	1
5013407	5	50134	I 0134	7	2	IMCB, F-7/3	1	1
5013408	5	50134	I 0134	8	2	IMCB, F-7/3	1	1
5013409	5	50134	I 0134	9	2	IMCB, F-7/3	1	1
5013410	5	50134	I 0134	10	2	IMCB, F-7/3	1	1
5013411	5	50134	I 0134	11	2	IMCB, F-7/3	1	1
5013412	5	50134	I 0134	12	2	IMCB, F-7/3	1	1
5013413	5	50134	I 0134	13	2	IMCB, F-7/3	1	1
5013414	5	50134	I 0134	14	2	IMCB, F-7/3	1	1
5013415	5	50134	I 0134	15	2	IMCB, F-7/3	1	1
5013416	5	50134	I 0134	16	2	IMCB, F-7/3	1	1
5013417	5	50134	I 0134	17	2	IMCB, F-7/3	1	1
5013418	5	50134	I 0134	18	2	IMCB, F-7/3	1	1
5013419	5	50134	I 0134	19	2	IMCB, F-7/3	1	1
5013420	5	50134	I 0134	20	2	IMCB, F-7/3	1	1
5013421	5	50134	I 0134	21	2	IMCB, F-7/3	1	1
5013422	5	50134	I 0134	22	2	IMCB, F-7/3	1	1
5013501	5	50135	I 0135	1	2	IMCB, F-8/4	1	1
5013502	5	50135	I 0135	2	2	IMCB, F-8/4	1	1
5013503	5	50135	I 0135	3	2	IMCB, F-8/4	1	1

5013504	5	50135	I 0135	4	2	IMCB, F-8/4	1	1
5013505	5	50135	I 0135	5	2	IMCB, F-8/4	1	1
5013506	5	50135	I 0135	6	2	IMCB, F-8/4	1	1
5013507	5	50135	I 0135	7	2	IMCB, F-8/4	1	1
5013508	5	50135	I 0135	8	2	IMCB, F-8/4	1	1
5013509	5	50135	I 0135	9	2	IMCB, F-8/4	1	1
5013510	5	50135	I 0135	10	2	IMCB, F-8/4	1	1
5013511	5	50135	I 0135	11	2	IMCB, F-8/4	1	1
5013512	5	50135	I 0135	12	2	IMCB, F-8/4	1	1
5013513	5	50135	I 0135	13	2	IMCB, F-8/4	1	1
5013514	5	50135	I 0135	14	2	IMCB, F-8/4	1	1
5013515	5	50135	I 0135	15	2	IMCB, F-8/4	1	1
5013516	5	50135	I 0135	16	2	IMCB, F-8/4	1	1
5013517	5	50135	I 0135	17	2	IMCB, F-8/4	1	1
5013518	5	50135	I 0135	18	2	IMCB, F-8/4	1	1
5013519	5	50135	I 0135	19	2	IMCB, F-8/4	1	1
5013520	5	50135	I 0135	20	2	IMCB, F-8/4	1	1
5013521	5	50135	I 0135	21	2	IMCB, F-8/4	1	1
5013522	5	50135	I 0135	22	2	IMCB, F-8/4	1	1
5013601	5	50136	I 0136	1	2	IMCG, F-6/2	2	1
5013602	5	50136	I 0136	2	2	IMCG, F-6/2	2	1
5013603	5	50136	I 0136	3	2	IMCG, F-6/2	2	1

5013604	5	50136	I 0136	4	2	IMCG, F-6/2	2	1
5013605	5	50136	I 0136	5	2	IMCG, F-6/2	2	1
5013606	5	50136	I 0136	6	2	IMCG, F-6/2	2	1
5013607	5	50136	I 0136	7	2	IMCG, F-6/2	2	1
5013608	5	50136	I 0136	8	2	IMCG, F-6/2	2	1
5013609	5	50136	I 0136	9	2	IMCG, F-6/2	2	1
5013610	5	50136	I 0136	10	2	IMCG, F-6/2	2	1
5013611	5	50136	I 0136	11	2	IMCG, F-6/2	2	1
5013612	5	50136	I 0136	12	2	IMCG, F-6/2	2	1
5013613	5	50136	I 0136	13	2	IMCG, F-6/2	2	1
5013614	5	50136	I 0136	14	2	IMCG, F-6/2	2	1
5013615	5	50136	I 0136	15	2	IMCG, F-6/2	2	1
5013616	5	50136	I 0136	16	2	IMCG, F-6/2	2	1
5013617	5	50136	I 0136	17	2	IMCG, F-6/2	2	1
5013618	5	50136	I 0136	18	2	IMCG, F-6/2	2	1
5013619	5	50136	I 0136	19	2	IMCG, F-6/2	2	1
5013620	5	50136	I 0136	20	2	IMCG, F-6/2	2	1
5013621	5	50136	I 0136	21	2	IMCG, F-6/2	2	1
5013622	5	50136	I 0136	22	2	IMCG, F-6/2	2	1
5013701	5	50137	I 0137	1	2	IMSJ (I-V), I-10/1 (No-1)	3	1
5013702	5	50137	I 0137	2	2	IMSJ (I-V), I-10/1 (No-1)	3	1
5013703	5	50137	I 0137	3	2	IMSJ (I-V), I-10/1 (No-1)	3	1

5013704	5	50137	I 0137	4	2	IMSJ (I-V), I-10/1 (No-1)	3	1
5013705	5	50137	I 0137	5	2	IMSJ (I-V), I-10/1 (No-1)	3	1
5013706	5	50137	I 0137	6	2	IMSJ (I-V), I-10/1 (No-1)	3	1
5013707	5	50137	I 0137	7	2	IMSJ (I-V), I-10/1 (No-1)	3	1
5013708	5	50137	I 0137	8	2	IMSJ (I-V), I-10/1 (No-1)	3	1
5013709	5	50137	I 0137	9	2	IMSJ (I-V), I-10/1 (No-1)	3	1
5013710	5	50137	I 0137	10	2	IMSJ (I-V), I-10/1 (No-1)	3	1
5013711	5	50137	I 0137	11	2	IMSJ (I-V), I-10/1 (No-1)	3	1
5013712	5	50137	I 0137	12	2	IMSJ (I-V), I-10/1 (No-1)	3	1
5013713	5	50137	I 0137	13	2	IMSJ (I-V), I-10/1 (No-1)	3	1
5013714	5	50137	I 0137	14	2	IMSJ (I-V), I-10/1 (No-1)	3	1
5013715	5	50137	I 0137	15	2	IMSJ (I-V), I-10/1 (No-1)	3	1
5013716	5	50137	I 0137	16	2	IMSJ (I-V), I-10/1 (No-1)	3	1
5013717	5	50137	I 0137	17	2	IMSJ (I-V), I-10/1 (No-1)	3	1
5013718	5	50137	I 0137	18	2	IMSJ (I-V), I-10/1 (No-1)	3	1
5013719	5	50137	I 0137	19	2	IMSJ (I-V), I-10/1 (No-1)	3	1
5013720	5	50137	I 0137	20	2	IMSJ (I-V), I-10/1 (No-1)	3	1
5013721	5	50137	I 0137	21	2	IMSJ (I-V), I-10/1 (No-1)	3	1
5013722	5	50137	I 0137	22	2	IMSJ (I-V), I-10/1 (No-1)	3	1
5013801	5	50138	I 0138	1	2	IMCG, F-10/2	2	1
5013802	5	50138	I 0138	2	2	IMCG, F-10/2	2	1
5013803	5	50138	I 0138	3	2	IMCG, F-10/2	2	1

5013804	5	50138	I 0138	4	2	IMCG, F-10/2	2	1
5013805	5	50138	I 0138	5	2	IMCG, F-10/2	2	1
5013806	5	50138	I 0138	6	2	IMCG, F-10/2	2	1
5013807	5	50138	I 0138	7	2	IMCG, F-10/2	2	1
5013808	5	50138	I 0138	8	2	IMCG, F-10/2	2	1
5013809	5	50138	I 0138	9	2	IMCG, F-10/2	2	1
5013810	5	50138	I 0138	10	2	IMCG, F-10/2	2	1
5013811	5	50138	I 0138	11	2	IMCG, F-10/2	2	1
5013812	5	50138	I 0138	12	2	IMCG, F-10/2	2	1
5013813	5	50138	I 0138	13	2	IMCG, F-10/2	2	1
5013814	5	50138	I 0138	14	2	IMCG, F-10/2	2	1
5013815	5	50138	I 0138	15	2	IMCG, F-10/2	2	1
5013816	5	50138	I 0138	16	2	IMCG, F-10/2	2	1
5013817	5	50138	I 0138	17	2	IMCG, F-10/2	2	1
5013818	5	50138	I 0138	18	2	IMCG, F-10/2	2	1
5013819	5	50138	I 0138	19	2	IMCG, F-10/2	2	1
5013820	5	50138	I 0138	20	2	IMCG, F-10/2	2	1
5013821	5	50138	I 0138	21	2	IMCG, F-10/2	2	1
5013822	5	50138	I 0138	22	2	IMCG, F-10/2	2	1
5013901	5	50139	I 0139	1	2	ICB, G-6/3	1	1
5013902	5	50139	I 0139	2	2	ICB, G-6/3	1	1
5013903	5	50139	I 0139	3	2	ICB, G-6/3	1	1

5013904	5	50139	I 0139	4	2	ICB, G-6/3	1	1
5013905	5	50139	I 0139	5	2	ICB, G-6/3	1	1
5013906	5	50139	I 0139	6	2	ICB, G-6/3	1	1
5013907	5	50139	I 0139	7	2	ICB, G-6/3	1	1
5013908	5	50139	I 0139	8	2	ICB, G-6/3	1	1
5013909	5	50139	I 0139	9	2	ICB, G-6/3	1	1
5013910	5	50139	I 0139	10	2	ICB, G-6/3	1	1
5013911	5	50139	I 0139	11	2	ICB, G-6/3	1	1
5013912	5	50139	I 0139	12	2	ICB, G-6/3	1	1
5013913	5	50139	I 0139	13	2	ICB, G-6/3	1	1
5013914	5	50139	I 0139	14	2	ICB, G-6/3	1	1
5013915	5	50139	I 0139	15	2	ICB, G-6/3	1	1
5013916	5	50139	I 0139	16	2	ICB, G-6/3	1	1
5013917	5	50139	I 0139	17	2	ICB, G-6/3	1	1
5013918	5	50139	I 0139	18	2	ICB, G-6/3	1	1
5013919	5	50139	I 0139	19	2	ICB, G-6/3	1	1
5013920	5	50139	I 0139	20	2	ICB, G-6/3	1	1
5013921	5	50139	I 0139	21	2	ICB, G-6/3	1	1
5013922	5	50139	I 0139	22	2	ICB, G-6/3	1	1
5066901	5	50669	I0669	1	1	CMS Mera Jafer Fazal Elahi Market G-12 Islamabad	3	3
5066902	5	50669	I0669	2	1	CMS Mera Jafer Fazal Elahi Market G-12 Islamabad	3	3

5066903	5	50669	I0669	3	1	CMS Mera Jafer Fazal Elahi Market G-12 Islamabad	3	3
5066904	5	50669	I0669	4	1	CMS Mera Jafer Fazal Elahi Market G-12 Islamabad	3	3
5066905	5	50669	I0669	5	1	CMS Mera Jafer Fazal Elahi Market G-12 Islamabad	3	3
5066906	5	50669	I0669	6	1	CMS Mera Jafer Fazal Elahi Market G-12 Islamabad	3	3
5066907	5	50669	I0669	7	1	CMS Mera Jafer Fazal Elahi Market G-12 Islamabad	3	3
5066908	5	50669	I0669	8	1	CMS Mera Jafer Fazal Elahi Market G-12 Islamabad	3	3
5066909	5	50669	I0669	9	1	CMS Mera Jafer Fazal Elahi Market G-12 Islamabad	3	3
5066910	5	50669	I0669	10	1	CMS Mera Jafer Fazal Elahi Market G-12 Islamabad	3	3
5066911	5	50669	I0669	11	1	CMS Mera Jafer Fazal Elahi Market G-12 Islamabad	3	3
5066912	5	50669	I0669	12	1	CMS Mera Jafer Fazal Elahi Market G-12 Islamabad	3	3
5066913	5	50669	I0669	13	1	CMS Mera Jafer Fazal Elahi Market G-12 Islamabad	3	3
5066914	5	50669	I0669	14	1	CMS Mera Jafer Fazal Elahi Market G-12 Islamabad	3	3
5066915	5	50669	I0669	15	1	CMS Mera Jafer Fazal Elahi Market G-12 Islamabad	3	3
5066916	5	50669	I0669	16	1	CMS Mera Jafer Fazal Elahi Market G-12 Islamabad	3	3

5066921	5	50669	I0669	21	1	CMS Mera Jafer Fazal Elahi Market G-12 Islamabad	3	3
5066922	5	50669	I0669	22	1	CMS Mera Jafer Fazal Elahi Market G-12 Islamabad	3	3
5067001	5	50670	I0670	1	2	Al Firdous Model School Bhara Kahu	3	4
5067002	5	50670	I0670	2	2	Al Firdous Model School Bhara Kahu	3	4
5067003	5	50670	I0670	3	2	Al Firdous Model School Bhara Kahu	3	4
5067004	5	50670	I0670	4	2	Al Firdous Model School Bhara Kahu	3	4
5067005	5	50670	I0670	5	2	Al Firdous Model School Bhara Kahu	3	4
5067006	5	50670	I0670	6	2	Al Firdous Model School Bhara Kahu	3	4
5067007	5	50670	I0670	7	2	Al Firdous Model School Bhara Kahu	3	4
5067008	5	50670	I0670	8	2	Al Firdous Model School Bhara Kahu	3	4
5067009	5	50670	I0670	9	2	Al Firdous Model School Bhara Kahu	3	4
5067010	5	50670	I0670	10	2	Al Firdous Model School Bhara Kahu	3	4
5067011	5	50670	I0670	11	2	Al Firdous Model School Bhara Kahu	3	4
5067012	5	50670	I0670	12	2	Al Firdous Model School Bhara Kahu	3	4

5067013	5	50670	I0670	13	2	Al Firdous Model School Bhara Kahu	3	4
5067014	5	50670	I0670	14	2	Al Firdous Model School Bhara Kahu	3	4
5067015	5	50670	I0670	15	2	Al Firdous Model School Bhara Kahu	3	4
5067016	5	50670	I0670	16	2	Al Firdous Model School Bhara Kahu	3	4
5067017	5	50670	I0670	17	2	Al Firdous Model School Bhara Kahu	3	4
5067018	5	50670	I0670	18	2	Al Firdous Model School Bhara Kahu	3	4
5067019	5	50670	I0670	19	2	Al Firdous Model School Bhara Kahu	3	4
5067020	5	50670	I0670	20	2	Al Firdous Model School Bhara Kahu	3	4
5067021	5	50670	I0670	21	2	Al Firdous Model School Bhara Kahu	3	4
5067022	5	50670	I0670	22	2	Al Firdous Model School Bhara Kahu	3	4
5067101	5	50671	I0671	1	2	PBF International College	3	2
5067102	5	50671	I0671	2	2	PBF International College	3	2
5067103	5	50671	I0671	3	2	PBF International College	3	2
5067104	5	50671	I0671	4	2	PBF International College	3	2
5067105	5	50671	I0671	5	2	PBF International College	3	2
5067106	5	50671	I0671	6	2	PBF International College	3	2
5067107	5	50671	I0671	7	2	PBF International College	3	2

5067108	5	50671	I0671	8	2	PBF International College	3	2
5067109	5	50671	I0671	9	2	PBF International College	3	2
5067110	5	50671	I0671	10	2	PBF International College	3	2
5067111	5	50671	I0671	11	2	PBF International College	3	2
5067112	5	50671	I0671	12	2	PBF International College	3	2
5067113	5	50671	I0671	13	2	PBF International College	3	2
5067114	5	50671	I0671	14	2	PBF International College	3	2
5067115	5	50671	I0671	15	2	PBF International College	3	2
5067116	5	50671	I0671	16	2	PBF International College	3	2
5067117	5	50671	I0671	17	2	PBF International College	3	2
5067118	5	50671	I0671	18	2	PBF International College	3	2
5067119	5	50671	I0671	19	2	PBF International College	3	2
5067120	5	50671	I0671	20	2	PBF International College	3	2
5067121	5	50671	I0671	21	2	PBF International College	3	2
5067122	5	50671	I0671	22	2	PBF International College	3	2
5067201	5	50672	I0672	1	2	Telecom Foundation School System.	3	2
5067202	5	50672	I0672	2	2	Telecom Foundation School System.	3	2
5067203	5	50672	I0672	3	2	Telecom Foundation School System.	3	2
5067204	5	50672	I0672	4	2	Telecom Foundation School System.	3	2

5067205	5	50672	I0672	5	2	Telecom Foundation School System.	3	2
5067206	5	50672	I0672	6	2	Telecom Foundation School System.	3	2
5067207	5	50672	I0672	7	2	Telecom Foundation School System.	3	2
5067208	5	50672	I0672	8	2	Telecom Foundation School System.	3	2
5067209	5	50672	I0672	9	2	Telecom Foundation School System.	3	2
5067210	5	50672	I0672	10	2	Telecom Foundation School System.	3	2
5067211	5	50672	I0672	11	2	Telecom Foundation School System.	3	2
5067212	5	50672	I0672	12	2	Telecom Foundation School System.	3	2
5067213	5	50672	I0672	13	2	Telecom Foundation School System.	3	2
5067214	5	50672	I0672	14	2	Telecom Foundation School System.	3	2
5067215	5	50672	I0672	15	2	Telecom Foundation School System.	3	2
5067216	5	50672	I0672	16	2	Telecom Foundation School System.	3	2
5067217	5	50672	I0672	17	2	Telecom Foundation School System.	3	2
5067218	5	50672	I0672	18	2	Telecom Foundation School System.	3	2

5067219	5	50672	I0672	19	2	Telecom Foundation School System.	3	2
5067220	5	50672	I0672	20	2	Telecom Foundation School System.	3	2
5067221	5	50672	I0672	21	2	Telecom Foundation School System.	3	2
5067222	5	50672	I0672	22	2	Telecom Foundation School System.	3	2

ANNEX-VII

CERTIFICATION OF TOOL VALIDATION

LARGE SCALE ASSESSMENT IN PAKISTAN: A CRITICAL ANALYSIS OF DEVELOPMENT AND APPLICATION OF NATIONAL ACHIEVEMENT TEST

LARGE SCALE ASSESSMENT IN PAKISTAN: A CRITICAL ANALYSIS OF DEVELOPMENT AND APPLICATION OF NATIONAL ACHIEVEMENT TEST

By

SAIMA ABBAS
PhD. Education, National University of Modern Languages, Islamabad

This certifies that the questionnaire titled "Teacher Perception Regarding Large Scale Assessment Test," developed by Don Klinger and adapted by the researcher for her study titled **"LARGE SCALE ASSESSMENT IN PAKISTAN: A CRITICAL ANALYSIS OF DEVELOPMENT AND APPLICATION OF NATIONAL ACHIEVEMENT TEST,**" to identify teacher perceptions regarding the National Achievement Test (Nat) in Pakistan, has been assessed and found appropriate for the data collection process. All items in the tool meet the objectives and address the research question. Face and content validity are also assured, and it may be used by the researcher for the data collection process.

Name: Dr. Tanveer Afzal

Designation: Assistant Professor

Institution: AIOU

Signature:

Date Jul 30, 2022

CERTIFICATION OF TOOL VALIDATION

LARGE SCALE ASSESSMENT IN PAKISTAN: A CRITICAL ANALYSIS OF DEVELOPMENT AND APPLICATION OF NATIONAL ACHIEVEMENT TEST

LARGE SCALE ASSESSMENT IN PAKISTAN: A CRITICAL ANALYSIS OF DEVELOPMENT AND APPLICATION OF NATIONAL ACHIEVEMENT TEST

By

SAIMA ABBAS

PhD. Education, National University of Modern Languages, Islamabad

This certifies that the questionnaire titled "Teacher Perception Regarding Large Scale Assessment Test," developed by Don Klinger and adapted by the researcher for her study titled **"LARGE SCALE ASSESSMENT IN PAKISTAN: A CRITICAL ANALYSIS OF DEVELOPMENT AND APPLICATION OF NATIONAL ACHIEVEMENT TEST,"** to identify teacher perceptions regarding the National Achievement Test (Nat) in Pakistan, has been assessed and found appropriate for the data collection process. All items in the tool meet the objectives and address the research question. Face and content validity are also assured, and it may be used by the researcher for the data collection process.

Name: ...Dr. Nasreen Bano...

Designation: ......Headmistress..........................

Institution: ....IMCB, G.10/4: Islamabad........................

Signature: ....Nasreen Bano

Date Jul 30, 2022

CERTIFICATION OF TOOL VALIDATION

LARGE SCALE ASSESSMENT IN PAKISTAN: A CRITICAL ANALYSIS OF DEVELOPMENT AND APPLICATION OF NATIONAL ACHIEVEMENT TEST

LARGE SCALE ASSESSMENT IN PAKISTAN: A CRITICAL ANALYSIS OF DEVELOPMENT AND APPLICATION OF NATIONAL ACHIEVEMENT TEST

By

SAIMA ABBAS

PhD. Education, National University of Modern Languages, Islamabad

This certifies that the questionnaire titled "Teacher Perception Regarding Large Scale Assessment Test," developed by Don Klinger and adapted by the researcher for her study titled **"LARGE SCALE ASSESSMENT IN PAKISTAN: A CRITICAL ANALYSIS OF DEVELOPMENT AND APPLICATION OF NATIONAL ACHIEVEMENT TEST,**" to identify teacher perceptions regarding the National Achievement Test (Nat) in Pakistan, has been assessed and found appropriate for the data collection process. All items in the tool meet the objectives and address the research question. Face and content validity are also assured, and it may be used by the researcher for the data collection process.

Name: Dr Zafar Iqbal

Designation: Principal

Institution: IMCB

Signature:

Date Jul 30, 2022