



Journal Website:  
<https://theusajournals.com/index.php/ajsshr>

Copyright: Original content from this work may be used under the terms of the creative commons attributes 4.0 licence.

## PEDAGOGICAL MECHANISMS OF PREPARING PRIMARY CLASS STUDENTS FOR PIRLS AND TIMSS INTERNATIONAL ASSESSMENT TESTS

**Submission Date:** October 06, 2024, **Accepted Date:** October 11, 2024,

**Published Date:** October 16, 2024

**Crossref doi:** <https://doi.org/10.37547/ajsshr/Volume04Issue10-14>

**Matyakubova Nuriya Azimbayevna**

Independent researcher of Urganch State University, Trainee-teacher of the "Methodology of Primary Education" department, Uzbekistan

### ABSTRACT

On the basis of the relevant decision of the Cabinet of Ministers of the Republic of Uzbekistan, preparation for international studies in the field of education quality assessment in the public education system was started. programs, scientific research is being conducted to develop and introduce innovative methods of assessment and development of students' reading, mathematical and natural-scientific literacy, creative thinking. Evaluation of the quality of education in the public education system of the Republic, organization of research, establishment of international relations, scientific research and innovation activities of the students, and above all, the creative ideas and creativity of the young generation. Progress in International Reading and Literacy Study (PIRLS) is an international assessment program to assess the level of reading and comprehension of primary 4th grade students based on the decision of the Cabinet of Ministers for the purpose of support and encouragement. is considered The "Roadmap" for the preparatory processes of the participation of the Republic of Uzbekistan in international studies was approved, and systematic work was started on this basis. National survey, research based on the requirements of the TIMSS program for the education quality component within the Human Capital Index - 2020 project, REDS research can be cited as an example. PIRLS assesses the learning of elementary school students in and out of the classroom with two overarching goals: artistic experience and information acquisition and use.



## KEYWORDS

PIRLS, TIMSS, REDS, PISA, OECD, Progress 8, GCCE, BTEC, AC and A Levels, Gold Standard Vocational Qualifications, T-Level, Anachronism and Irrelevance, YOUNGOV Survey.

## INTRODUCTION

The professional significance of this study is to answer the question that has been of great interest in recent years: what explains the huge difference between the results of Russian schoolchildren in two international comparative studies related to the assessment of reading literacy - PIRLS and PISA. It should be noted that the PIRLS assesses reading literacy at the end of elementary school, during the transition to using reading as one of the main tools of post-school education. PISA assesses reading, literacy among 15-year-old students during the transition from using reading as a learning tool to real-life learning guidance. In essence, both tests assess students' readiness for a new way of using written messages to solve problems appropriate for the next age level. In PIRLS, Russian 4th graders twice (2001, 2006) showed high results in reading (statistically significantly higher than the international average) [Tcukermen, 2001]. In 2006, Russia significantly improved its results and led the group of leading countries in research. In the PISA study of reading literacy, the results of 15-year-old Russian students are significantly lower than the OECD average. This result was achieved four times (2000,

2003, 2006, 2009), but the main thing for us is not the rating indicators, but the dynamics of our results.

**Assessment by school system.** This section focuses on system-level evaluations—in other words, evaluations conducted in schools across the country to meet system-level goals. However, it should be noted that daily assessment is an important element of teachers' professional experience and requires all teachers to have a high level of assessment literacy.

A system-level assessment has four main objectives:

1. Monitoring school standards and ensuring coverage of the curriculum;
2. Monitoring of national education standards;
3. Confirmation of students' knowledge;
4. Control access to future opportunities and identify individuals.

The last two can be classified as learner-oriented goals and are less important in the two main stages, while the first two are closely related to educational responsibilities. Currently, all four functions are interrelated, and several of the problems explored in



this section can be at least partially mitigated by reducing the use of multiple targets in the same evaluation. Steps have already been taken in this regard, for example by introducing a national referendum test. Each year, students take a standardized test, which is used to track national standards and set benchmarks.

## METHODOLOGY

International surveys such as PISA, TIMSS and PIRLS are increasingly playing a similar role by tracking national standards and benchmarking them against other jurisdictions. However, these surveys cover only a narrow range of topics and have significant limitations in terms of choice and response rate. (Jerrim 2021 p.). At the same time, accountability reforms mean that Ofsted decisions and school interventions are now less dependent on exam results.

## Analysis of literature on the topic (Literature review).

## ASSESSMENT SOLUTIONS OF RECENT REFORMS.

The English education system has undergone significant reforms in the last ten years. Following a controversial peer review, the New National Curriculum was published in 2013, followed by consistent assessment reforms. These reforms are closely related and have far-reaching implications for what and how students learn in school. In primary school, national curriculum levels were abolished and

new assessments such as the multiplication table test, phonics screening and entry level were introduced over time. At Key Stage 4, the headline measure of school performance moved to Progress 8, along with a measure of the proportion of pupils achieving GCCE in a basket of subjects known as the English Baccalaureate (EBass). In addition, the publication of Wolf's qualitative review in 2011 (Wolf 2011)

It is sometimes thought that GCCE results are norm-referenced in such a way as to cause a certain proportion of students to fail, but this is wrong, because the National Credentials Test, if the evidence shows that standards are rising, allowing the number of students to increase every year. Abolishing many qualifications and revising the 'equivalence' system between GCCEs for some vocational qualifications.

At Key Stage 5, AC and A levels are 'separated' so that A levels are assessed at the end of the two-year course. While BTECs are being phased out, a new Gold Standard qualification, T-level, is also being introduced. Students who fail to achieve a 'good' GCCE in English or maths by the age of 16 will now have to resit the qualification. However, following recommendations from the City of Rochford (Rochford 2016), Level P (CEND), which was previously used to assess the progress of some pupils with special educational needs and disabilities, was withdrawn.



Concurrent systemic reforms, including compulsory academies and increased responsibility for “remediation” of low-performing academies. However, more recently Ofsted has tried to stop the culture of 'teaching to the test' by putting school curricula under the microscope (Ofsted -2019)

Throughout the period, the dominant narrative revolved around "firmness," but the impact of the sentencing changes was unclear. Some refer to English and mathematics scores in international assessments such as PISA, PIRLS, TIMSS. (Sivinini 2020; DFE 2017; Lechain 2023; Schools Week 2019), but others question these arguments. (Jerrim 2019,20) or pointed to England's failure to close the achievement gap between disadvantaged pupils and their peers. (Fair Education Alliance and CFEI 2022; Tuskett et al 2022). As this report shows, there are serious concerns about the wider impact of these reforms.

Criticisms of traditional examinations are numerous and can be traced back to the Asland Report of 1911. (Asland 1911; Richmond 2021). There are three tensions underlying these criticisms. These are:

1. Trade-off between authenticity and reliability
2. How stressful and anxiety-provoking is the assessment?
3. Positive and negative ways that assessment affects teaching and learning.

## HOW TO FIND A BALANCE BETWEEN VALUATION AND PRODUCTION?

### Research methodology. Anachronism and irrelevance.

Many believe that England's exam-based approach to public assessment is no longer appropriate. (Barnard 2019; Burghes 2021; Hazell 2021; Hyman 2021; Rethinking Assessment). Others recognize the importance of exams, but believe that tests are too dominant in assessment, and that what can be fairly assessed trumps what should be assessed. (Pearson 2022). The Assessment Change Limit Review (2023) suggests that the current 'one-size-fits-all' system needs to be reformed and that assessment should be done in more varied ways, such as vivas, open-book exams, personal extended projects and 'Micro-Accounts’” (Lusas 2021). Many teachers, young people, parents and employers are of the same opinion. According to the YOUNGOV survey of 1000 teachers (Txe Edge Foundation 2020), 92% of teachers believe that the evaluation system should fully recognize the strengths and skills of young people, not written exams believes that 48% of teachers and 47% of parents feel that testing and grading. According to a small survey of teachers and parents by Brutain Thinks (Person 2016), it provides a fair assessment of actual student achievement.

In a 2014 survey by the UK's Commission for Employment and Skills, less than half of employers



rated qualifications as 'important' and 'important' in recruitment, while two-thirds said so about work experience (Shury et al. others 2014). Therefore, there is a widespread perception that there is a problem and there is a great appetite for change.

**THE IMPORTANCE OF TENDERNESS.** Exams have a significant advantage over unverified assessments in terms of reliability, and this is critical to the fairness of the system. The high-stakes nature of assessment in a competitive system means that getting a 'wrong' grade due to unreliable assessment has significant consequences. There is currently a margin of error in any estimate. Any expansion of this, or an increase in the degree of skewing of the border towards certain groups or minorities, would be a serious problem.

Therefore, attempts to increase validity must consider three main threats to reliability;

1. Type of question
2. Teacher's verdict
3. Test conditions

### Question type

Ofqual emphasized the balance between reliability and validity, indicating which marker should have the highest response among markers.

### “Objective questions”, such as multiple choice

### “Questions (whereas) that require longer responses have lower marker agreement within a person”

This is why inter-marker agreement is lower in subjects containing essay-style responses (Ofqual 2019). Questions involving judgments against general descriptors are particularly susceptible to uncertainty and limitations of human judgment (Christodoulou 2017).

“Targeted” multiple-choice questions can test the application of knowledge and skills to real-world scenarios, and it's easy to overstate the extent to which exams now focus on "regurgitating facts. “Indeed, assessments are often constructed to include progression from factual recall to the application of that knowledge to a real-world scenario. provide” (Chind and Ellis 2021). However, validity is not a property of the test, but the conclusions drawn from it. (Christodoulou 2013). Claims of genuine impairment, then, are said to be meaningless because the assessment does not support valuable conclusions (such as whether the student has gained a good education, a valuable education, or is eligible for further education or employment).

### Teacher's verdict.

Many of the proposed reforms to assessment increase the role of teacher input. The risk is that it will have a negative impact on students from disadvantaged groups, including those with low incomes, parents with



no tertiary education, certain ethnic groups and those with special educational needs and disabilities. possible (Bur GECC and Greaves 2013; Campbell 2015). These letters showed during the pandemic, when teacher-assessed grades widened the gap between students eligible for free school meals compared to ineligible students at the same previous reading level; as well as Roma, Gypsy and Traveler pupils, compared to white British pupils and public school pupils compared to their independent school peers (Beynon 2021).

A comparison of students' "centered grades" and "reported grades" shows that students whose parents are graduates are more likely to receive "relatively higher" "centered grades" than their "reported grades". 'more educated parents (Anders et al. 2021). A recent qualitative study of students affected by the abolition of testing, as well as ethnic minority and disadvantaged students, found that when assessments are based on teacher judgment, biases and structural inequities hinder their achievement. determined that it will do. As one youth said, "I want to be judged on what I do, not what my teachers think I can do" (Bhopal and Meers 2023). Therefore, any reform in assessment must be combined with reliability and validity to avoid further disadvantaging vulnerable students. As Christodoulou (2015) points out, the inhuman nature of tests may be the best system for fair assessment. Some argue that large-scale moderation

processes can alleviate human reasoning problems, but there is no evidence that this will be enough. As Johnson (2013) notes, based on a review of assessment practices in a number of global jurisdictions:

"Applying to moderation is supposed to ensure validity, consistency, and comparability in teacher evaluations. "Unfortunately, there is currently no reliable evidence to support this assumption."

Extensive moderation processes also create additional workloads for overworked teachers.

## RESULTS

### Controlled conditions.

Many of the proposed alternatives to the current examination system would result in assessment taking place outside the examination hall in less controlled settings.

Assessment in England has not always been exam-heavy In the early 1990s, most 16-year-olds sat English GCCEs without exams and based entirely on coursework (QCA 2006). A number of concerns led to its replacement by 'supervised assessment'. (QCDA 2005). These issues include: plagiarism, students receiving an unfair amount of aid (through multiple referrals by parents or teachers)

Campbell (2015) shows that teachers have systematic biases in their assessment of student abilities, based on



income level, gender, special educational needs status, ethnicity, and spoken language.

Similarly, based on a repeated cross-sectional analysis of achievement at age 11, BurGECC and Greaves concluded that teachers from poor families and some ethnic minorities may have a significant disadvantage in the recorded achievement of children (BurGECC and Greaves 2013). However, there is inconsistency between the results of different studies (Lee and Newton 2021).

- direct malpractice in limited cases
- questions about the correctness of the course work, especially in science practice.

Therefore, coursework was replaced in 2009 by 'controlled assessment' (CA). This involved students completing coursework-style assignments under standardized, controlled conditions. However, CA was soon abandoned because, according to some, it was "difficult" and did not allow for the research skills and independent study that the coursework should have envisaged - in fact it was no longer became a kind of examination (Oates 2013; Ofqual). On the other hand, according to some data, teachers' views are more positive (Ipsos Mori 2011).

Given that students have varying levels of support both at school and at home, conducting more assessments outside of test settings may skew results. As the

pandemic has highlighted, there are huge differences in students' learning environments at home, and this can be particularly detrimental to vulnerable students.

Developments from large language models such as ChatGPT may also have increased the need for more carefully controlled test conditions.

### Revise the case for an unverified assessment

Non-exam-based assessment is not closed and must be continually revised for each topic as further research emerges and technology evolves. Beiton's (2016) analysis suggests that GECC coursework grades are a better predictor of A-level grades than exams, and Pinot de Moira suggests that concerns about assessment bias may be overstated and that "coursework has no advantage it doesn't bring a disadvantage either," he says. People with low social and economic status (2020:26). Wyness (2021) therefore argues that coursework should be used alongside exams, but that it should be externally set and assessed. Richmond and Regan (2023) examine the advantages and disadvantages of written examinations, coursework/supervised assessments, teacher assessments, oral assessments, portfolios, extended essays/projects, and performance-based assessments.

They concluded that there may be scope for greater use of oral assessment, extended essays and projects in the 5 key stages. They therefore encourage all



students in the class to take an extended project qualification and take an additional course at AC level, a standard assessed by oral examination. However, at the high school level, they note:

“written examinations should remain the primary means of assessing student knowledge and understanding... (because) an increased focus on coursework and other forms of 'teacher assessment' increases the workload of teachers and can be biased which leads to unreliable grades, against students from disadvantaged backgrounds.

### Richmond and Regan 2023

Focusing and finding clearly displayed information:

- ❖ Identifying and searching for information related to the specific purpose of study;
- ❖ Search for specific opinions;
- ❖ Search for the definition of words or phrases in the text;
- ❖ Determining the place and time of the story;
- ❖ Finding the main topic of the text (if it is clearly stated);
- ❖ Determining certain information in the drawing.
- ✓ **Making direct conclusions:**
- ❖ Reasoning that one event causes another to happen;

- ❖ Justifying the actions of the hero of the work;
- ❖ Describe the relationship between the two heroes of the work;
- ❖ Determining which part of the text serves a specific purpose;
- ✓ **Check and evaluate content, language and text elements:**
- ❖ Discuss the completeness or accuracy of the information given in the text;
- ❖ Estimating the probability that the events described in the text will actually happen;
- ❖ To assess the extent to which the opinion put forward by the author can change people's thoughts and lifestyle;

### CONCLUSIONS

The importance of the teaching process in the development of student learning indicators in the PIRLS and TIMSS international programs, joint effectiveness in the organization of the teaching process, the relationship between school staff, parents and students mutual trust is considered a factor that serves to create a positive learning process at school, and these factors, in turn, create a positive learning environment.

**High literacy results in these countries are associated with the following factors:**

**-Availability of books and digital devices at home**





## - Intelligent parents and their love for books

## -The fact that the child has learned to read and write before going to school and the high rate of coverage in preschool educational institutions

Based on the above factors, it is determined that we can certainly become one of the top ten countries if we work on the basis of scientifically based evidence, while developing several other factors to achieve high results of PIRLS-reading literacy.

This section focuses on system-level assessment – in other words, there is a strong focus on international assessment conducted in schools across the country to achieve system-level goals, but it is important to note that everyday assessment is an important element of teachers' professional experience, and requires all teachers to have a high level of assessment literacy.

## REFERENCES

1. Sh. Ismailov and others. (2019). Materials of the seminar-training on the topic "On international research in the field of educational quality assessment (in the example of mathematics in the PISA International Assessment Program)".
2. A.A. Ismailov and others. (2019). Assessment of students' mathematical literacy in international studies (Methodological manual for mathematics teachers, methodologists and experts in the field).

National Center for the Implementation of International Studies on the Evaluation of the Quality of Education under the Education Inspectorate.

3. A.A. Ismailov and others. (2020). Newsletters designed to prepare students for international studies (newsletters for mathematics, natural sciences, native language and literature teachers, methodologists and specialists in these fields). National Center for the Implementation of International Research on the Evaluation of the Quality of Education under the Inspectorate of Education.
4. A.B.Radjiyev et al. (2019) International research program on the assessment of students' literacy. Manual. Tashkent. National Center for International Research on Educational Quality Assessment.
5. Mullis, I.V.S, Martin. M. O. (Eds). (2017). TIMSS (2019) Assessment Frameworks. Retrieved from Boston College, TIMSS & PIRLS International Study Center.
6. A.A. Ismailov and others. (2021). National Center for the Implementation of International Studies on the Evaluation of the Quality of Education, BULLETIN 2021-Nº3 for the preparation of students for international studies (Elementary teachers of general secondary educational institutions and a set of tasks intended for students).