Physical Education, Health and Social Sciences
https://journal-of-social-education.org
E-ISSN: 2958-
P-ISSN: 2958-

A Comprehensive Differential Item Functioning Analysis of Physics Test Items in the Secondary School Certificate Examination

Dr. Mobeen Ul Islam[1], Hafiz Muhammad Salman Naveed[2], Afeefa Amjad[3]

[1] Assistant Professor, Department of Education, University of Gujrat,
Email: drmobeen.islam@uog.edu.pk
[2] PhD Scholar, Department of Education, University of Gujrat, Email: salmanvd01@gmail.com
[3] M. Phil. Department of Education, University of Gujrat

DOI: https://doi.org/10.63163/jpehss.v4i1.1111

Abstract
The study was aimed at analyzing test items by means of differential item functioning. The test questions are exercises that learners are to undertake during the test. Differentiation between items functioning analysis is highly significant in test development at both national and international levels to confirm the biased and unbiased items in the test. This research was useful in establishing the validity of the test items and the data regarding the degree of difference in item functioning performance regarding chosen groups. It was a quantitative paradigm descriptive study. Data were analyzed using the Mantel-Haenszel. Differential item functioning was useful to assess the strengths and weaknesses of the tests of the target groups with non-similar traits using the test items based on polytomous items in the science subject, that is, Physics of grade 10th, in the annual examination of BISE Sahiwal 2018. The conclusion was reached that there is a necessity to create the test items, which are to be developed by the experts, in order to conduct an appropriate assessment.

Keywords: DIF, Annual Examination, Mantel-Haenszel, Secondary School Certificate Examination, Physics Test Items

Introduction
Assessment in education is very crucial in determining the knowledge and abilities of the students and influencing their future education, as well as the education policies. The Secondary School Certificate Examination (SSCE), being one of such milestones of secondary education, is really significant in assessing the eligibility of the students to higher education and employment opportunities. Nevertheless, it is important that the examination be fair in assessing the abilities of all the students, and any prejudices in the test questions may create wrong assumptions about the actual capabilities of the students. Different Item Functioning (DIF) analysis is one of the most useful techniques for determining the possibility of biases in test items. It is conducted to determine whether test items perform equally across different groups. The study will strive to do a thorough DI analysis of Physics test items in the SSCE, which will focus on the performance of test items in various groups of students, e.g., gender, school type, and category of students. In most educational frameworks, such as that in Pakistan, the SSCE is a conclusive test that has a profound effect on students in terms of their future education and career prospects (Bashir & Qureshi, 2021). Nonetheless, the unequal distribution of resources, the quality of the teaching process, and the norms that are accepted in society can affect the performance of various groups of students in the exam. There are great differences between the public and the private schools in regard to the

resources and education quality. Although in many cases, the facilities available in private schools are better and children are given more personal consideration, government schools might have problems with congested classes and the lack of infrastructure and available resources (Ali, 2020). Also, the gender bias in society, especially in Physics, may lead to inequality in the learning process of both male and female students (Jabeen et al., 2022). In addition, there could be differences between regular and private students, where the latter may have access to extra coaching and resources that might be a further source of bias to the assessment outcomes (Hassan, 2019). Through the DIF analysis of these various groups, the research paper will seek to examine whether there exist any biases in the test items of SSCE Physics that may interfere with the fairness and accuracy of the test. The inequality between genders in education, especially in the areas of Science, Technology, Engineering, and Mathematics (STEM), is also a major issue in most regions of the world, including Pakistan. Most male students have been traditionally urged to study science and technology, historically, and female students are hindered by their societies and culture, which usually puts them off studying science and technology (Shah et al., 2019). Such gender disparity is especially pronounced in those disciplines, such as Physics, in which the number of male students vastly exceeds the number of female students, not only when it comes to classroom attendance but also in the performance during exams (Farooq, 2020). Although there has been some development in gender equality in education, female students in Pakistan are still underrepresented in the STEM subjects. Studies have demonstrated that, in a learning institution where the main intention is to offer equal opportunities, gender biases can still affect the learning environment and result in different outcomes (Hussain et al., 2021). Through the DIF analysis, the research paper will examine the existence of any gender bias in the test items of Physics, either in the format or in the level of difficulty, which may disadvantage women. It is necessary to determine all these biases so that the environment of encouragement to learn and prosper can be made more inclusive, and both male and female learners can achieve their potential depending on their skills rather than gender expectations. The distinction between the students based on whether they are private or regular also has a key influence on how they perform in standardized tests such as the SSCE. The private students generally have a school that is better equipped, a lower ratio of classes, and more individual attention than the regular students who are normally in the government-funded schools. This gap in access to resources may result in a difference in the readiness of students for the exam (Memon et al., 2018). The probability of attending more coaching centers also increases among the private students, and this further improves their capacity to succeed in the high-stakes assessment. In their turn, ordinary students might not receive the same opportunities to use the additional materials, and this circumstance may affect their results in the Physics test. The question is open to whether the SSCE Physics test items are equally as hard for the two groups of students or whether they unconsciously favor the private students who are better prepared. This paper seeks to analyze the DIF of the private and regular students and whether the test questions are fair to all the students, irrespective of their ability to access extra resources.

The contrasts between the government and the non-government schools in Pakistan are very clear, as the former school is attended by the greatest number of students, especially those of lower-income categories. The public schools are often poorly funded, and there are also problems with classrooms being overcrowded, an insufficient number of qualified teachers, and access to educational materials (Ali, 2020). Conversely, better infrastructure, more resources, and the level of teaching are usually better in private schools that serve wealthier families. These differences may have a crucial impact on the performance of students on the SSCE, especially in such topics as Physics, which demand a good level of knowledge of science and the skills of solving problems (Shah & Anwar, 2021). The question also lies in whether the items in the SSCE Physics test are equally difficult for both the students of the public and the private schools, or whether they are more inclined towards the students of the private schools based on their preparedness and accessibility to materials. Through DIF analysis with respect to school type, the study aims at getting any possible bias in the SSCE Physics test items concerning the student background of either being a student of a public or a private school. This paper will seek to determine the occurrence of Differential Item Functioning (DIF) in the Physics questions of the SSCE, namely, it will investigate how the questions operate between various groups of students, in terms of gender, type of school, and student category. This study will be used to add to the overall discussion of fairness and equity in education measurement by employing the performance of male and female students and those students of both public and private schools. Also, the research will enhance the validity and equity of the SSCE by eliminating possible prejudice in the test questions so that it can be a true representation of the actual ability of students, irrespective of gender, school, or resource availability. The study is not only important for its chances of improving the fairness of the SSCE, but also for the general implications of the study on educational assessment worldwide. Through a thorough DIF analysis, the research study is adding to current efforts of making large-scale educational assessment more equitable. It holds the view that all the students are assessed and placed on the same level. Finally, the research will encourage a more equal and accommodating education system, whereby the items in examinations are truly relevant to what students know and can do, devoid of the interference of other external variables like gender, type of school, and category of students.

Research Objectives

The research objectives of the study were:

To investigate whether physics test items used by 10th-grade students in secondary school examinations exhibit differential item performance for different traits of students (gender, student type, school type).

Research Questions

The following research questions were developed to achieve the research objectives of the study:

A) Is there any Physics test item used in *the Board of Intermediate and Secondary Education that shows differential item functioning across gender groups*?

B) Is there any Physics test item used in *the Board of Intermediate and Secondary Education that shows* differential item functioning among regular and private students?

C) Is there any Physics test item used in *the Board of Intermediate and Secondary Education that shows* differential item functioning for students' institution type?

Literature Review

The differential item functioning (DIF) analysis is relevant to provide fairness in the educational measures with the aim of determining the functioning of items in a test by various groups of students, despite their being of similar ability level. DIF happens when a test question is biased against some groups of people, i.e., gender, socioeconomic status, or type of school, thus giving undue advantages or disadvantages. Such biases are identified through such a technique, and in this way, the test actually measures what it is expected to do without the interference of external factors (Hambleton & Swaminathan, 1985). The use of DIF has been extensive over the years at both international and local levels in order to improve the validity of educational tests and foster equality in standardized testing. DIF has been used to test the system fairness of test items internationally. A large amount of literature has been dedicated to DIF based on gender, especially in the STEM areas such as Mathematics and Physics. As an example, a study conducted by Kaigama et al. (2025) compared gender differences in Basic Science and Technology items and discovered that some test items were biased towards the male population. This bias was explained by the manner in which the items were framed and occasionally themed around male experiences or views. In the same vein, research in Turkey has indicated that test items in science and mathematics assessments have DIF between the genders, with male students, in many cases, having an advantage since many of the items usually have cultural and social biases (Turn, 2017). These results indicate that in the so-called gender-neutral systems, it is still possible that certain biases play a role in the outcomes, which raises significant doubts related to the applicability of the test to female students. The other significant feature of the DIF analysis is the consideration of the functionality of test items depending on the school type, namely, the public and private schools. Inequality in these two kinds of schools in terms of resources, teacher qualification, and the readiness of students tends to create variations in the standardized test results. A study by Ali (2020) in Pakistan has shown that students in the private schools, usually better funded and equipped, have better outcomes in the Secondary School Certificate Examination (SSCE) compared to those in the public schools. This gap leads to a question of whether test contents in the SSCE are just for students in the public schools, because they might not be equally equipped to access the resources. The latter was also highlighted in one research by Memon et al. (2018), which indicated that according to DIF analysis, some of the test items were biased towards the students of private schools, as students in this group have more opportunities to take advantage of additional learning materials, including coaching and extracurricular activities. These findings also bring about the necessity of DIF as a part of the process of making sure that standard tests are not used to disadvantage students in the public schools when they need to be provided with the same level of support and preparation. Besides the type of school, socioeconomic status (SES) has also been established to be a key determinant of the academic performance of students. SES tends to be associated with resource availability, and higher-SES students have a superior chance at academic enrichment out of the school environment. The students who are regular in Pakistan and usually attend the public school are more likely to belong to lower-income families than students who attend the private schools. It has been demonstrated that SES-related differences may contribute to DI, in which test items might be skewed toward students with higher SES histories that had access to private tutors, superior learning environments, and more academic assistance (Hassan, 2019). Utilizing DIF analysis between regular and private students, the researchers can conclude that some of the items may unintentionally favor students who have higher access to academic resources, giving them unfair advantages in the process of examination.

In Pakistan, gender differences in the education of Physics are very pronounced, as male students have long been doing better than female students in the STEM subjects. Women students are usually burdened by the demands and prejudices of society, which makes them reluctant to study science and technology, and, thereby, they are not enrolled in physics and other science programs as much (Shah et al., 2019). Such gender differences are manifested in standardized testing, whereby male students generally perform better than female students in Physics tests. Research has demonstrated that this difference is not only because of the difference in ability but also because of gender biases that exist in test questions. As an example, test questions on Physics can occasionally rely on examples or situations that are more relatable to the male students, thereby disadvantaging the female students (Farooq, 2020). Then, DIF analysis can play a vital role in determining whether one gender finds test items more challenging and in providing a test that can be considered fair and just to all the students. Islam, Naveed, and Amjad (2025) conducted a study that was published in the Advanced Social Science Archive Journal to investigate DIF in the Biology test items in the SSCE. They followed the Mantel-Haenszel to test the functioning of items, and discovered that some items in the exam were characterized by DIF favoring students of the private schools as compared to the public schools. The research implied that the questions of the biology test were not equally difficult for all students and that it was necessary to make certain adjustments to make it fairer. Despite the fact that the study concentrated on Biology, its methodology and results are very applicable in the present study, where the researcher intends to examine DIF in Physics test items. In this case, the DIF analysis underlines the significance of considering the issue of whether the test items can be effective in the case of the various demographic categories, such as gender, school type, and student category. The necessity of DIF analysis in educational tests is not so much a requirement to reveal and solve the bias, depending on the school type or gender. DIF is also significant in the provision of convergent and discriminant validity of tests. Convergent validity is whether or not items on a test actually measure what they are intended to measure. Conversely, discriminant validity prevents items from measuring other unrelated constructs accidentally (Brown, 2020). DIF analysis is a serious approach to assessing such forms of validity, whereby items can be considered to be trustworthy and just among students of varying groups. The significance of the DIF analysis is especially visible in large-scale testing, like the SSCE, where the questions on the test should be valid in terms of students, as the questions should correspond with their actual skills, and should not be affected by some additional factors, like gender or socioeconomic background. Conclusively, DIF analysis plays an important role in keeping educational tests like the SSCE fair and equitable. DF analysis can be used to enhance the validity of educational tests by determining test items that show varying performance between gender, type of school, and type of student. Research on DIF in science classes, such as Biology and Mathematics, has revealed that one should recognize item bias and make appropriate amendments that would put all students, irrespective of their background, on equal grounds. Through examining DIF in Physics test questions in the SSCE, the research aims to make a contribution to the larger endeavor to both make standardized tests more accommodating and representative of the real abilities of the students.

Research Design

The research design that was embraced by the researcher in this study was a descriptive research design. The researcher explained, validated, and described the research findings. The research question in the provided study was to establish that there existed differences in item functioning (DIF) of the test items being used in the target population of Physics of the first group of test takers of BISE Sahiwal in 2018. Differentiation Item Functioning is also a form of analysis that is used to establish biased items in an assessment. The article by Zieky (2003) also mentions that DIF is a powerful approach to establishing the potential unfairness and assessing the reasons for achievement discrepancies, rather than contrasting the total score. Essentially, it is given between two groupings, namely the focal group and the reference group of test takers, and is used to decrease the test bias.

Population and Sample

The present research sample size comprised all the test items in the Physics subject, as it was given to 80000 students who were the test takers in the 2018 examination conducted at BISE Sahiwal.

Research Tool

The research instrument of the study was the 2018 Physics exams MCQs test items of the Board of Intermediate and Secondary Education, Sahiwal. The instrument that was applied in this study was a group of multiple-choice objective test questions in the year 2018 in Physics. The physics test included 15 items.

Data Collection

This data was gathered on the Board of Intermediate Secondary Education, Sahiwal. The chairman of the Sahiwal board was persuaded that the information gathered during the test would not be utilized in any other way other than research, and the purpose of the research was clarified to him.

Data Analysis
The experiment aimed at determining the DIF in test items. First, the data were entered into an Excel sheet, and it was classified into different groups. It was then analyzed in terms of DIF in items and by the use of the SPSS software to derive descriptive statistics for its analyzed version of the Item Response Theory 4.4. In this section, data analysis was done to identify the differential item functioning of the items that were used in the 2018 Sahiwal annual exams of the secondary certificate. This section consisted of the analysis and interpretation of data. The statistical method employed by the researcher to analyze the data is the Mantel-Haenszel statistic approach, as well as the descriptive statistics approach.

Question A: Is there any Physics test item used in the Board of Intermediate and Secondary Education that shows differential item functioning across gender groups?

Table 4.1: DIF Analysis among the Gender Group

Item No.	Mean		S.D.		Rpbis	DeltaMH	P-value
	Female	Male	Female	Male			
MCQ1	.30	.28	.460	.447	0.203	-0.1989 A	0.8819
MCQ2	.44	.49	.497	.500	0.332	-0.8848 A	0.0506
MCQ3	.50	.44	.501	.497	0.631	0.7331 A	0.0512
MCQ4	.28	.26	.448	.440	0.257	-0.4659 A	0.5450
MCQ5	.29	.26	.453	.440	0.579	0.1575 A	0.8667
MCQ6	.24	.21	.425	.406	0.376	0.2792 A	0.6804
MCQ7	.27	.26	.446	.438	0.624	-0.2056 A	0.8107
MCQ8	.42	.46	.494	.499	0.193	-0.5827 A	0.1642
MCQ9	.27	.23	.444	.424	0.675	0.0601 A	0.8541
MCQ10	.47	.41	.500	.493	0.487	0.7120 A	0.1206
MCQ11	.41	.42	.493	.494	0.532	-0.3511 A	0.5451
MCQ12	.27	.22	.442	.416	0.469	0.5353 A	0.4920

The outcome of the Differentiated Items Functioning (DIF) analysis is presented in Table 4.1, which assesses the presence of gender-based biases in the Physics test items by comparing how female and male students perform using different items in the tests. The table presents the means, standard deviations (S.D.), point-biserial correlation (Rpbis), DeltaMH values, and the p-values of the individual items. A large p-value (less than 0.05) would signify a possible DIF, which implies that the item is going to be different in male and female conditions. When the results were examined, it can be seen that the p-value of most of the items is significantly above 0.05, meaning that there is no statistically significant gender difference in the performance of these test items. As an example, the p-values of MCQ1 ($p = 0.8819$), MCQ5 ($p = 0.8667$), and MCQ6 ($p = 0.6804$) all indicate the lack of a gender bias. The MCQ2 ($p = 0.0506$) and MCQ3 ($p = 0.0512$) have p-values that are near the significance threshold; however, they are more than the conventional cut-off, and any difference in these items observed is not statistically significant.
Moreover, the DeltaMH values, which are an indicator of the magnitude of DIF, are rather low across all items, which once again indicates that the test items have little to no gender-based biases. One example is MCQ1 has a DeltaMH of -0.1989, and MCQ3 has a value of 0.7331 that lies within a range of values that show there is no significant gender bias. Overall, these findings of this DIF analysis indicate that the Physics test items are not heavily biased by gender, as indicated by the

large value of p-values and small DeltaMH. This means that the test must be fair and that it must work effectively on both female and male students.
Question B: Is there any Physics test item used in the Board of Intermediate and Secondary Education that shows differential item functioning among regular and private students?

Table 4.2, the Differentiating Items Functioning (DIF) test, is used to compare the scores of regular and private students in different items on the Physics test. The table shows the mean scores, standard deviations (S.D.), point-biserial correlations (Rpbis), DeltaMH values, and the p-values of every item. A p-value under 0.05 will mean that there is significant DIF, implying that the test item might discriminate against regular and private students.

Table 4.2: DIF Analysis among the Regular and Private Students

Item No.	Mean		S.D.		Rpbis	DeltaMH	P-value
	Regular student	Private student	Regular student	Private student			
MCQ1	.27	.37	.447	.485	0.798	-0.6666 A	0.5652
MCQ2	.46	.48	.499	.502	0.551	0.5056 A	0.4247
MCQ3	.47	.47	.499	.501	0.272	-0.0809 A	0.9570
MCQ4	.26	.30	.441	.460	0.762	0.9421 A	0.3982
MCQ5	.25	.39	.436	.489	0.731	-1.6324 C	0.0118
MCQ6	.22	.23	.415	.420	0.619	0.3518 A	0.7256
MCQ7	.26	.28	.440	.453	0.723	0.8964 A	0.2298
MCQ8	.44	.47	.496	.501	0.399	-0.4087 A	0.5393
MCQ9	.24	.32	.427	.469	0.878	-0.8183 A	0.6785
MCQ10	.43	.49	.496	.502	0.559	-0.5342 A	0.4710
MCQ11	.41	.47	.492	.501	0.653	-0.0046 A	0.8793
MCQ12	.23	.29	.424	.456	0.759	-0.7354 A	0.5247

After analysis of test results, it is clear that the majority of the test items are significant (all those with p-values above 0.05); therefore, no significant difference between regular and private students in terms of DIF. To give an example, MCQ1 (p = 0.5652), MCQ2 (p = 0.4247), and MCQ4 (p = 0.3982) all have p-values that indicate that there is no bias between the two groups of students. Even such items as MCQ8 (p = 0.5393) and MCQ9 (p = 0.6785), where there are only slight differences, are still significantly above the significance level. MCQ5, however, has a p-value of 0.0118, which is less than the 0.05 level; hence, significant DIF between regular and private students. The DeltaMH of MCQ5 is -1.6324, implying that the two sets of the population respond differently to this item, and this may favor one group as opposed to the other. Conclusively, most test items demonstrate no significant difference in DIF between regular and private students, and thus are usually fair between the two populations. Nevertheless, MCQ5 is a question that should be given additional consideration because it has a great DIF that can be interpreted to mean that the question is biased to a certain group of students and, thus, it is necessary to revise the item to ensure fairness.

Question C: Is there any Physics test item used in *the Board of Intermediate and Secondary Education that shows* differential item functioning for students' institution type?

Table 4.3: DIF Analysis among the Public and Private Institutes

Item No	Mean		S.D.		Rpbis	DeltaMH	P-value
	Public school	Private school	Public school	Private school			
MCQ1	.28	.30	.449	.460	0.798	0.5797 A	0.4900
MCQ2	.46	.47	.499	.500	0.551	0.2583 A	0.6209
MCQ3	.48	.45	.500	.498	0.272	0.2138 A	0.6194
MCQ4	.26	.28	.440	.449	0.762	0.6875 A	0.3645
MCQ5	.26	.30	.438	.459	0.731	-0.5262 A	0.4277
MCQ6	.21	.24	.409	.425	0.619	-0.3496 A	0.5927
MCQ7	.26	.28	.438	.449	0.723	0.2149 A	0.8008
MCQ8	.43	.46	.496	.499	0.399	-0.2703 A	0.5601
MCQ9	.24	.27	.428	.443	0.878	0.4575 A	0.9120
MCQ10	.43	.46	.495	.499	0.559	-0.3470 A	0.5015
MCQ11	.41	.43	.493	.495	0.653	0.4690 A	0.4278
MCQ12	.23	.27	.419	.444	0.759	-0.9878 A	0.1717

The analysis of the Differential item functioning (DIF) in Table 4.3 involves the comparison of the results of students in the Physics test with the performance of students in both the school types, either public or private schools. The table reflects the average scores, standard deviations (S.D.), point-biserial correlations (Rpbis), values of DeltaMH, and p-values of individual test items. A p-value of less than 0.05 would have been a sign of any significant DIF, and therefore, the item would have been acting differently on the students of public and private schools. The analysis indicates that most of the test items do not depict any significant DIF between the two groups, as indicated by the p-values of all the test items, which have a p-value greater than 0.05. An example is MCQ1 (p = 0.4900), MCQ2 (p = 0.6209), and MCQ4 (p = 0.3645); they all have no significant difference in the responses of the students in the public and private schools, implying that they are all performing well in that case. On the same note, most of the DeltaMH values are near zero, meaning that there is very little bias. MCQ12, however, has a p-value of 0.1717, which is above the significance threshold but not significantly so, indicating that the item may have a slight difference between the two groups, although not significantly so. The DeltaMH of MCQ12 is -0.9878, which implies that the approach to this item differs slightly between the students of the public school and those of the private school, but, at the same time, it does not imply that it is highly biased. On the whole, the DIF analysis indicates that the majority of the test items in the Physics exam are equal in both the public and the private schools, with MCQ12 being the one with the slight difference. The fact that there are no large DIF values in the majority of items implies that the test is serving students of the various types of schools fairly, which is also significant to the consideration of fairness in the assessment. Nevertheless, MCQ12 might need more assessment so that it could be neutral across the two types of schools.

Discussions

According to the Differentiating item functioning (DIF) analysis, there is no significant bias related to gender in terms of Physics test items that are applied to the SSCE, as the p-values of all the items are high. The p-values of most items, such as MCQ1, MCQ5, and MCQ6, are significantly more than 0.05, indicating that the performance of female and male students does not significantly differ in these items. Although the p-values of MCQ2 and MCQ3 are near the threshold, they are not statistically significant, which adds to the conclusion that the differences that are observed are not significant. Also, the DeltaMH values, which measure the degree of bias, are not high, supporting the fact that items are functionally equivalent between the sexes. These findings indicate that the test item Physics in this research paper is just and impartial, giving both a female and a male student an equal chance to be assessed, which is a significant factor in ensuring the validity and equity of the test. The Differentiating Item Functioning (DIF) analysis outcomes between regular and private students show that the majority of the test items are fair and can be applied to both categories equally, as the p-values of most of the items are high (more than 0.05). As an example, MCQ1, MCQ2, and MCQ4 do not show any significant difference in the performance of regular and private students, implying that the items are not biased and give an equal assessment to both groups. Nevertheless, MCQ5 has a p-value of 0.0118, which is less than the significance level, showing that this item might not work in the same way that it does between regular and private students. The negative value of DeltaMH of -1.6324 is also large, which indicates that the item can be biased towards one group as compared to the other. This observation shows the need to renew MCQ5 so that it does not favor or deprive either of the groups. Generally, the findings indicate that the test items are rather fair, although MCQ5 should be focused on in order to eliminate the bias of the assessment. The findings of the Differentiate item functioning (DIF) comparing the performance of the students in the public school and the private school show that most of the test items have the same performance, as indicated by high p-values (Above 0.05) and low DeltaMH values, which indicate no significant DIF. Questions like MCQ1, MCQ2, and MCQ4, which do not differ significantly in the performance of the students in the public schools and in the private school, therefore suggest that the questions are operating equitably between the two types of schools. Nonetheless, MCQ12 seems to be the most interesting with a p-value of 0.1717 and DeltaMH of -0.9878, which means that the difference between the performance of students in the two schools (public and private) in this specific item is rather minor but is not significant. Such results indicate that the test materials are mostly unbiased, and MCQ12 can be considered with some attention, so that it can be completely neutral in both groups of students. The findings, in general, justify the fairness of the Physics test items, and there is only weak evidence to prove that one of the items contains some DIF, which further confirms the necessity to review the item to ensure equity in the assessment thoroughly.

References

Ali, A. (2020). *Educational disparities in Pakistan: Public vs. private schools*. Journal of Education Policy, 35(4), 12-25.

Bashir, M., & Qureshi, R. (2021). *The role of the Secondary School Certificate Examination in shaping students' future opportunities*. Educational Assessment Review, 29(2), 33-47.

Farooq, M. (2020). *Gender disparities in science education: The case of Physics in Pakistan*. Journal of Educational Research, 14(3), 56-67.

Hassan, S. (2019). *The influence of supplementary coaching on students' performance in national examinations*. International Journal of Educational Research, 58(1), 45-60.

Hussain, Z., Rizvi, S., & Asghar, N. (2021). *Breaking barriers: Gender disparities in STEM education in Pakistan*. International Journal of Gender Studies in Education, 9(2), 120-134.

Memon, A., Khan, F., & Haider, Z. (2018). *Socioeconomic status and academic performance: The impact of private and public schooling in Pakistan*. Journal of Social Sciences, 42(1), 89-101.

Shah, S., & Anwar, S. (2021). *The role of school type in student performance: Evidence from Pakistan's SSCE results*. Pakistan Journal of Education, 45(2), 74-90.

Shah, P., Ali, M., & Zafar, S. (2019). *Women in science education: Examining the barriers and opportunities in Physics education*. International Journal of Science Education, 41(4), 367-379.

Ali, A. (2020). Educational disparities in Pakistan: Public vs. private schools. *Journal of Education Policy*, *35*(4), 12-25.

Farooq, M. (2020). Gender disparities in science education: The case of Physics in Pakistan. *Journal of Educational Research*, *14*(3), 56-67.

Hassan, S. (2019). The influence of supplementary coaching on students' performance in national examinations. *International Journal of Educational Research*, *58*(1), 45-60.

Islam, M. U., Naveed, H. M. S., & Amjad, A. (2025). Examining differential item functioning of Biology test items in the Secondary School Certificate Examination. *Advance Social Science Archive Journal*, *4*(02), 3901–3911. Retrieved from https://assajournal.com/index.php/36/article/view/1361

Kaigama, A., et al. (2025). An analysis of gender-based differential item functioning in Basic Science and Technology. *Global Journal of Educational Research*.

Martinková, P. (2017). Checking equity: Why differential item functioning matters. *Frontiers in Education Methods*.

Shah, P., Ali, M., & Zafar, S. (2019). Women in science education: Examining the barriers and opportunities in Physics education. *International Journal of Science Education*, *41*(4), 367-379.

Shanmugam, S. (2018). Determining gender differential item functioning for mathematics in a coeducational school culture. *ERIC*.

Turn, R. (2017). Differential item functioning analysis of science and mathematics items in university entrance examinations in Turkey. *Journal of Educational Measurement*, *24*(2), 85-92.