



Original Article

Analysis of Grade Eight Examination Subjects in the Banadir Region of Somalia

Ahmed Mohamud Warsame^{1*}

¹ Ministry of Education, Culture & Higher Education, P. O. Box 1182, Mogadishu Somalia.

* Author for Correspondence ORCID ID: <https://orcid.org/0000-0002-5326-3571>; Email: asaahid@gmail.com

Article DOI: <https://doi.org/10.37284/eajes.6.1.1046>

Date Published: ABSTRACT

13 January 2023

Keywords:

*Weight of Cognitive Skills,
Difficulty Level,
Instructional Objectives,
Design of Question Paper,
Item Analysis.*

Item analysis is essential in the development of assessment tools and in standardised computing measures of student performance. This study was conducted to analyse the questions of grade eight examination subjects in the Banadir region of Somalia with respect to the instructional objectives, the difficulty level, and their impact on the student's achievement. The population of the study was grade 8 students of the Banadir region of Somalia who sat for the national examinations in 2022. A purposive sampling technique was followed for the selection of subjects and the subject experts. A random sampling technique was used to select 10% of question papers for study in every subject examined in Banadir. Also, the view of 16 subject experts was taken into consideration regarding the design of the question paper, instructional objectives, alignment with the syllabus, and difficulty level. The findings of the study revealed the effect of the difficulty level and instructional objectives on students' achievement was apparent. However, the study found a severe mismatch between the distribution of the items in the original table of specifications and the analysed items. The study recommended for post examinations analysis to improve the quality of the questions and to conduct separate analyses for each subject in the syllabus.

APA CITATION

Warsame, A. M. (2023). Analysis of Grade Eight Examination Subjects in the Banadir Region of Somalia *East African Journal of Education Studies*, 6(1), 11-21. <https://doi.org/10.37284/eajes.6.1.1046>.

CHICAGO CITATION

Warsame, Ahmed Mohamud. 2023. "Analysis of Grade Eight Examination Subjects in the Banadir Region of Somalia". *East African Journal of Education Studies* 6 (1), 11-21. <https://doi.org/10.37284/eajes.6.1.1046>

HARVARD CITATION

Warsame, A. M. (2023) "Analysis of Grade Eight Examination Subjects in the Banadir Region of Somalia", *East African Journal of Education Studies*, 6(1), pp. 11-21. doi: 10.37284/eajes.6.1.1046.

IEEE CITATION

A. M. Warsame, "Analysis of Grade Eight Examination Subjects in the Banadir Region of Somalia", *EAJES*, vol. 6, no. 1, pp. 11-21, Jan. 2023.

MLA CITATION

Warsame, Ahmed Mohamud. "Analysis of Grade Eight Examination Subjects in the Banadir Region of Somalia". *East African Journal of Education Studies*, Vol. 6, no. 1, Jan. 2023, pp. 11-21, doi:10.37284/eajes.6.1.1046

INTRODUCTION

Assessment is an integral part of the teaching and learning process. It is intended to influence and affect how teachers teach and students learn. The major function of assessment is to improve the learning process in schools. Kumar & Ii (2014) pointed out that when the assessment is intended to strengthen the quality, it has two purposes; a) to improve the teaching process at school and eventually improve learning, and b) to ascertain the students' achievement levels. Both purposes are equally important in an educational program

High-stakes public examinations exert strong pressure on students, their parents, teachers, and schools. According to Anderson et al. (2001), public examinations have a strong impact on the nature of teaching and learning in schools, stronger than other forms of external assessment. The practice of testing has become increasingly common throughout the twentieth century and reliance on information gained from test scores has made an indelible mark on our culture (Kinsey, 2003). All levels of education, kindergarten through graduate school, most professional license procedures, and many employment avenues place a high reliance on test performance to disseminate opportunities and promote and assure professional standards.

From the aforementioned, we can infer that the assessment focuses on the learner's acquisition of knowledge, skills, and values, the ability to apply the skills and knowledge, and to ensure that they have undergone a massive knowledge, attitude, and behavioural change. Criteria used to establish test quality generally focus on the areas of test design, test analysis techniques, and test score interpretation (Kinsey, 2003). Items should transcend the mere

application of rote memory and lower-order thinking questions to higher-order thinking and challenging questions to stimulate their critical thinking skills.

In Somalia, students should go through this public examination to examine their acquisition of the needed knowledge and skills in this early stage of their life. To substantiate this, both Education Policy and National Curriculum Framework emphasise the need for the standardisation of assessment to obviate arbitrary decision-making at various levels and in different institutions. The Curriculum Framework outlines that assessment should focus on ensuring the progress of all students as they have achieved the outcomes outlined in the syllabus (MoECHE, 2017a). The Education Policy (MoECHE, 2017b) indicates that there will be a continuous learning assessment and standardised grade 8th and 12th assessments, and thus, it outlines the purpose of the assessment as follows;

- Enhancement of the overall educational system
- Mastery of essential skills and knowledge
- Diagnosing learners' performance
- Motivating both teachers and learners to work hard
- Awarding qualifications or grades for the achievement

Consequently, the quality of the items must be prioritised. A carefully designed and constructed question can only distinguish between the different abilities of the exam takers and provide information about how teaching and learning were effective. Recent concerns are placed increasingly on the

design, production, administration, analysis, reporting, and interpretation of tests (Kinsey, 2003). Test quality should be a high priority for those who make tests, those who take tests, and those who rely on test scores for decision-making.

Significance of Item Analysis

The quality of the test depends on the quality of the individual question, and to determine the quality of the individual item, item analysis is done after the administration and scoring of the students (Ado Abdu Bichi et al., 2015). After the exam is administered and results released, item analysis has great significance, the importance of which becomes even greater when a sample of answer scripts is examined and analysed. Equally important is the significant impact this analysis has on the teaching and learning process.

Test analysis examines how the test items perform as a set (Susan Matlock-Hetzel, 1997). It is a process that examines student responses to individual test items (questions) to assess the quality of those items and the test as a whole (Shakil, 2008). Item analysis investigates the quality of the individual item set for the students concerning bloom's cognitive skills and test blueprint. Also, this kind of item analysis is done periodically to provide the examination administrators with a critical analysis of the state of affairs of the exams administered. Such analyses can also be employed to revise and improve both items and the test as a whole.

Purpose of Item Analysis

Item analysis is a valuable and powerful technique for teachers and instructors to guide and improve their instruction. It enables them to increase their test construction skills, identify weak areas in the syllabus that need greater emphasis, and improve other classroom practices. Suruchi and Rana (2014) pointed out there are two major purposes of Item analysis; firstly, to identify defective test items and secondly, to indicate the areas where the learners

have or have not mastered. Furthermore, they noted that Item analysis measures the effectiveness of individual test items in terms of their difficulty level and their discrimination power, i.e., to distinguish between high and low achievers in a test. Thus, Item analysis helps in selecting the best test items in the final draft by retaining the good and rejecting poor test items. Similarly, it shows the need to review and modify the items in a test.

This research analyses the grade examination subjects of the Banadir region in Somalia, which was conducted in 2022. The purpose is to measure the effectiveness of the test item in terms of its alignment with the curriculum learning outcomes, difficulty level, and discrimination power, i.e., to distinguish between high and low achievers in a test. This analysis will help in future examinations to craft quality examination questions.

Objectives of the Study

The research sought to answer whether the question papers in every subject are valid in terms of the following:

- Coverage of instructional objectives.
- Appropriate difficulty levels to distinguish the needs of different students.
- Whether the questions papers have appropriate objectivity.
- The weaknesses and strengths of the question papers.

Research Questions

The research sought to answer the following questions to

- Is there an effect of different difficulty levels on students' Pass/Fail in grade eight examinations?

- Is there a relationship between examining different cognitive levels and students' performance

METHODOLOGY

This is a descriptive study which intends to analyse the questions of grade eight examination subjects in the Banadir region of Somalia with respect to the instructional objectives, the difficulty level, and their impact on the student's achievement. In this research paper, two sampling techniques were used; a) purposive sampling strategy for the experts who have undertaken the research of the subjects. These experts were selected based on their knowledge and experience in analysing the scores of the students and the subject matter. b) Random sampling was used for the exam papers under investigation. 10% out of 27,600 grade eight students were randomly selected. These students were from all levels of the community including government schools, privately owned or managed, IDPs, and special needs students, from the Banadir region.

As the examination commenced in 2020, the researcher took cognisance of the importance of also capturing the performance of the students in the previous years for comparison. To draw valid and reliable conclusions, the research employed teachers with good command of the subjects they are working on along with long experience in teaching, setting and analysing questions, and measuring students' abilities. A total of 16 subject experts were chosen irrespective of their location. The researcher conducted three days of training on the skills required and the tools for data collection.

The research used the table of specifications of the three years examination to analyse differences between the cognitive levels examined in the different years. The past question papers and reports of the previous examination were also used. In addition, the research used MS Excel to analyse the data. Advanced formulas are used to analyse the difficulty and discrimination indexes of the questions.

Research Instrument

The researcher used the following tools to analyse the data;

- Question papers of all 2022 grade examination subjects (Math, English, Afsoomaali, Arabic, Islamic, Science, Social, and ICT)
- Past examination papers for 2020 and 2021
- Reports of the examiners and
- Examination results and student scores
- Questionnaire to collect and analyse different parameters of question papers

In addition to the above tools, the researcher used a simplified version of Bloom's Taxonomy to simplify the work for the teachers and other subject specialists. *Table 1* shows different cognitive objectives, how to measure and their action verbs to guide them to distinguish between the objectives examined and their measures (Anderson et al., 2001).

Table 1: A Summarises the cognitive Dimensions Assessed in the Final Examinations

Cognitive Levels Assessed	How to Measure	Sample of Action Verbs Used
Remembering	Students' ability to recall, retrieve or recognise information, ideas, facts, and principles in the approximate form in which they learned them	define, recognise, recall, select, find, show, reproduce, give, mention, indicate, identify, label, list, memorise, name, outline, tell, point, outline, locate, which, duplicate, underline, state, spell, repeat, record, what, where, when,
Understanding	Students' ability to translate, comprehend or interpret information, facts or ideas based on prior learning	Explain, Summarise, Paraphrase, describe, illustrate, approximate, clarify, complete, convert, extrapolate, fill in the blanks, convert, exemplify, restate, estimate, add, rewrite, insert, intrapolate, interpolate, extend, rephrase, translate, indicate, show, specify, retell, to what extent, why
Applying	Students' ability to select, transfer, use data, ideas, facts, and principles to complete a problem or task in a new situation with minimum direction	Use, compute, solve, apply, manipulate, demonstrate, derive, dramatise, apply, change, calculate, illustrate, employ, assign, express, imitate, articulate, practice, instruct, simulate, make use of
Analysing	Students' ability to distinguish, classify and relate the assumptions, hypotheses, evidence, or structure of a statement or question	Analyse, Categorise, Compare, contrast, separate, break down, select, group, order, distinguish, characterise, relate, discriminate, associate, classify, divide, factorise, sequence, tabulate, prioritise, operate, relate, organise, arrange, inspect, correlate,
Evaluation	Students' ability to appraise, assess, or critique an idea or fact on the basis of specific standards and criteria	Judge, account for, infer, dispute, recommend, critique, assess, conclude, argue, evaluate, defend, support, reframe, predict, rate, verify, prioritise, justify, agree/disagree, comment on, criticise, decide, deduce, diagnose, appraise, disprove, grade, inspect, rule on, suggest
Creating		Create, design, hypothesise, invent, develop, revise, prepare, discover, animate, enhance, improvise, improve, invent, collect, combine, compose, construct, arrange, assemble, devise, formulate, join, modify, plot, draw, plan, rewrite, set up, synthesise, change, generate, redesign, sketch, rearrange, reconstruct, relate, propose, integrate, express, compile, reorganise

FINDINGS AND DISCUSSIONS

From *Table 2*, we can discern that the mean of the students was severely below the average in the

subjects of English, Arabic, and Islamic. The performance of the students in the mother tongue was the highest (77%). While ICT, Science, and Math were above average.

Table 2: Subjects with their mean grade and standard deviation

Subjects	Mean	Standard deviation
Math	51	17
Science	69	17
Social	53	15
English	33	24
Arabic	36	24
Islamic	40	20
Af soomaali	77	16
ICT	64	18

From *Table 3*, we can observe that the distribution of instructional objectives was fair and balanced. Across the subjects, the weightage given to the cognitive skills was meticulously prepared. However, the social subject was different. Only 4%

of the total marks were allotted to the application. In addition, all language subjects (English, Arabic, and Afsoomali) were given similar structures for their similar nature.

Table 3: Weightage given to Instructional objectives as per the approved specification table of the 2022 examinations

Subjects	Instructional objectives			Total
	Knowledge	Comprehension	Application	
ICT	40	33	27	100
English	34	30	36	100
Arabic	34	30	36	100
Somali	34	30	36	100
Science	66	20	14	100
Islamic	52	26	22	100
Math	12	70	18	100
Social	76	20	4	100

Question 1: is there an effect of different difficulty levels on students' Pass/Fail?

Table 4 demonstrates how the different levels of difficulty or easiness affect the performance of the students. It is a common perception that when the difficulty level is lowered, the pass of the students

will jump high. Nevertheless, in some subjects, there is a weak correlation between the difficulty level and passing of the candidates.

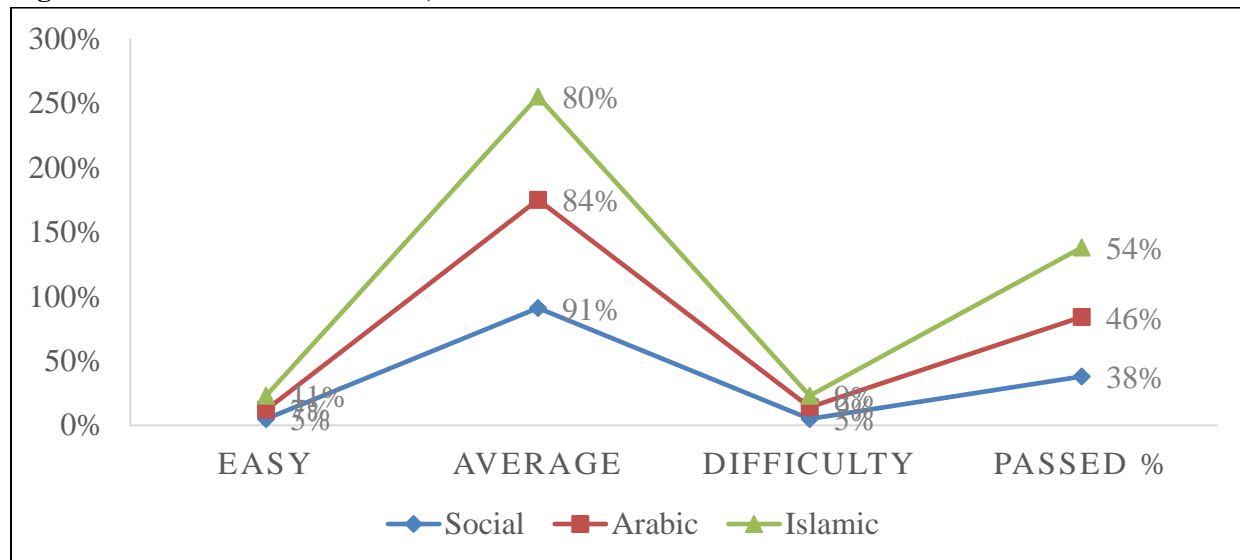
Table 4: weightage given to difficulty level

Year	Weightage given to difficulty level (in %)			Number of students passed %
	Easy	Average	Difficulty	
Math	31%	56%	13%	80%
Science	67%	33%	0%	95%
Social	5%	91%	5%	38%
English	36%	62%	2%	88%
Arabic	7%	84%	9%	46%
Islamic	11%	80%	9%	54%
Soomaali	81%	19%	0%	97%
ICT	42%	58%	0%	93%

For instance, in the Arabic subject, the difficulty level is 9%, and 46% of the students passed. The situation worsened in Social, where the difficulty level was 5%, and only 38% of the students managed to pass. The Islamic subject is not far from

these two subjects; the difficulty level was 9%, whereas the pass rate was 54%. This means even though the difficulty level of the subjects was relatively low (9%, 5%, and 9%), respectively, the majority of the students failed.

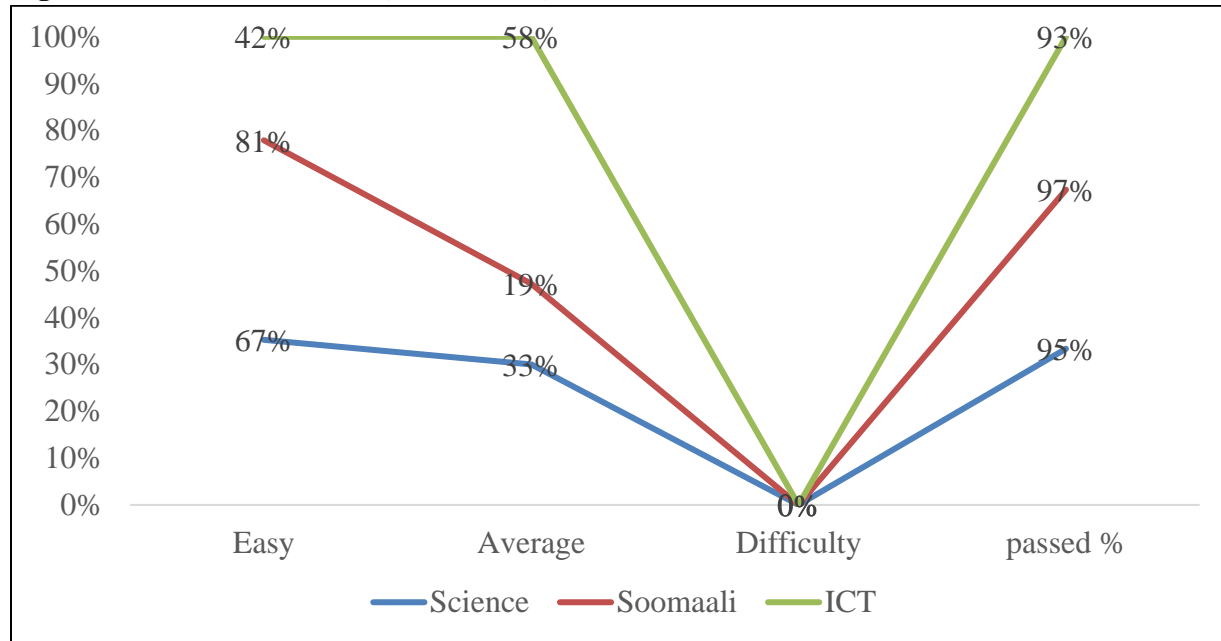
Figure 1: Performance of Islamic, Arabic and Social



The research found that the papers on Science, Afsoomaali, and ICT were not developed to discriminate between students. These papers were found to test only easy and moderate questions giving Zero percentage to the difficulty level, and as a result, the performance of the students was too high (95%, 97%, and 93%), respectively. This shows the great correlation between the difficulty

level of the questions and students passing the examinations. The subject of English is not far from this group. Students performed 80% due to the small percentage given to the difficulty level (2%). However, the most reasonable paper was the Math paper which was broken down into 31% (easy), 56% (moderate), and 13% (difficulty), and the pass of the students was fair (80%).

Figure 2: Performance of ICT, Afsoomaali, and Science



Question 2: Is there a relationship between examining different cognitive levels and students' performance

Table 5 presents information about the weightage given to the instructional objectives in each subject. In its analysis, the percentage of knowledge, comprehension, and application in each subject. If you compare Table 5 with Table 3 (weights of the instructional objectives), we can realize that teachers did not give any attention to the specification table. Without exception, all subjects appeared totally different and far from the table of

specifications. It was the responsibility of the examination staff to address this matter during item development and moderation and make sure that teachers reflected the table of specification (TOS) at every step. TOS is developed to guide teachers to match topics, objectives, and skill levels. The following table demonstrates the huge difference between the approved table of specification and the item analysis TOS.

Table 5: Weightage of cognitive levels

Subjects	Weightage of cognitive levels (in %)			Student performance
	Knowledge %	Comprehension %	Application %	Number of students passed %
Math	20%	4%	76%	80%
Science	60%	33%	7%	95%
English	41%	52%	7%	38%
Social	80%	18%	2%	88%
Arabic	37%	51%	12%	46%
Islamic	56%	33%	11%	54%
Soomaali	23%	56%	21%	97%
ICT	43%	45%	11%	93%

It appears that the number of passed students is more in Science and Social owing to the weight

given to the knowledge area (60% and 80%) respectively. The research shows that the weightage

given to comprehension in English and Arabic subjects has severely affected the passing of the students (38% and 46%). Both subjects gave more than 50% of the total marks to comprehension and undermined the nature of the language, which requires more applications. Although Afsoomaali is a mother language, nevertheless, the passing of the students was very high (97%), and this is due to the high proportion of knowledge (56%). The search found that the allocation of cognitive abilities in Math was not appropriately distributed. Only 4% of the questions measured comprehension. This allocation upended the table of specifications of the subject which was clearly stating the distribution of

cognitive abilities in each paper. Therefore, the subject needed an extensive review to align it with TOS. The research also found that across the subjects, teachers did not follow the TOS strictly and this can be seen in the same-nature subjects like; Math, Science, and languages. Thus, it can be noted that there is an effect of different cognitive levels in different subjects on students' performance. Further, it can be true that if more weight is given to the "Knowledge" level, more students tend to pass as is the case in Science and Social. In contrast, if more weight is given to comprehension and/or application, more students are expected to fail as is the case in English and Arabic.

Table 6: Comparison between the approved specification table and the actual specification table found in the item analysis.

Subjects	Instructional objectives	Approved specification table	Actual distribution	Percentage increase/decrease
Math	Knowledge	12	20	66.7%
	Comprehension	70	4	-94.3%
	Application	18	76	322.2%
Arabic	Knowledge	34	37	8.8%
	Comprehension	30	51	70.0%
	Application	36	12	-66.7%
Afsoomaali	Knowledge	30	23	-23.3%
	Comprehension	35	56	60.0%
	Application	36	21	-41.7%
Social	Knowledge	76	80	5.3%
	Comprehension	20	18	-10.0%
	Application	4	2	-50.0%
English	Knowledge	34	41	20.6%
	Comprehension	30	57	90.0%
	Application	36	2	-94.4%
Islamic	Knowledge	52	56	7.7%
	Comprehension	26	33	26.9%
	Application	22	11	-50.0%
ICT	Knowledge	40	43	7.5%
	Comprehension	33	45	36.4%
	Application	27	11	-59.3%

CONCLUSION

The research sought to answer two questions related to the weights of instructional objectives in the exam papers and the difficulty levels of the questions. The research found the following realities;

There is a good relationship between the difficulty level of the questions and the students' passing rate. This is evidenced in Science, Afsoomaali, and ICT, where the performance of the students in these three subjects was incredibly high. These subjects measured only the "easy" and "moderate" parts of the syllabus. In contrast, Arabic, Islamic, and Social were the least performed subjects due to the high percentages given to the "average" level and rising "difficulty" level (9%); hence, the passing of the students was very low. There is a strong correlation between the students' pass/fail and the weights given to the instructional objectives of the syllabus. The research shows that when we give more weight to the "knowledge", more students will tend to pass and vice versa. This is evidenced in Science and Social, where the pass rate was above 88%.

Recommendation

While analysing the question papers on the different subjects, many interesting aspects were observed. As some of the questions papers on different subjects were dissimilar in approaches, it was difficult to compare them. The analysis showed that there was a need to review and improve the questions during moderation. Some subjects needed to balance the instructional objectives and abilities measured, some needed to balance the levels of difficulty, and some needed to reflect the test blueprint closely. However, there was a substantial need to improve the quality of the questions. Following are some suggestions regarding this:

- Across the subjects, the items analysed were knowledge-based. In fact, only a few subjects can be exempted. So, questions related to testing higher thinking skills were not

examined. Questions such as problem-solving, critical thinking, imaginative thinking, interpretative, creative thinking reasoning, etc., should be included more in numbers.

- The design and construction of the exam papers were not coherent and balanced. Some papers tested some difficult questions while others did not. Some mainly tested average questions, while others gave more than 80% to easy questions. This shows how the exam papers lacked coherence. Hence, it is recommended that the same policy is applied to all subjects and strict follow-up mechanism of the item setters be established
- There is no rule of thumb for listing or grouping questions. The criteria to be followed is from easy to difficult questions. Start first with easy questions as an incentive for students and gradually proceed to the difficult questions.
- The application of the test blueprint is essential in test development. Carefully designed questions not only measure the progress of the student's knowledge and skill but also measure the overall development of the students. Thus, it is recommended that continuous strict follow-ups be made to ensure that questions are in line with the standards and guidelines for item development.
- The examinations board should carry out a post-examination analysis of the question papers for further improvement. This is done through the evaluation of the papers by trustworthy subject experts to examine them question by question and submit the final draft of the paper.
- Finally, evaluation is a technical area where skills matter. Item setters and other evaluators should be regularly trained in this task. They should be trained according to recent trends and developments in the assessment world.

- As this is the first analytical report to analyse the examinations, the research suggests conducting separate reports for each subject in the curriculum. This will give a full picture of the quality of education.

REFERENCES

- Ado Abdu Bichi, Rahimah Embong, & Mustafa Mamat. (2015). Classical Item Analysis of Science Achievement Test. *Proceedings of the IniSZA Research Conference 2015, 2015*(April), 1–11.
- Anderson, L. W., Krathwohl, D. R., Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., Raths, J., & Wittrock, M. C. (2001). *A Taxonomy for Learning, Teaching and Assessing*. 302.
- Shakil, M. (2008, March). Assessing Student Performance Using Test Item Analysis and its Relevance to the State Exit Final Exams of MAT0024 Classes-An Action Research Project. In *A Paper presented on MDC Conference Day*.
- Kinsey, T. L. (2003). *A comparison of IRT and Rasch procedures in a mixed-item format test*. University of North Texas.
- Kumar, A., & Ii, S. P. (2014). An analytical study of mathematics question papers of madhya pradesh board with respect to its patterns/design, content coverage, difficulty level, objectivity in marking and its impact on achievement of students. In *Research Article International Journal of Advancement in Education and Social Sciences*, 2(2).
- MoECHE. (2017a). Somalia National Curriculum Framework. *MoECHE, Mogadishu, Somalia*.
- MoECHE. (2017b). Somlia National Education Policy. *MoECHE, Mogadishu, Somalia*.
- Suruchi, S., & Rana, S. S. (2014). Test Item Analysis and Relationship Between Difficulty Level and Discrimination Index of Test Items in an Achievement Test in Biology. *Paripex - Indian Journal of Research*, 3(6), 56–58. <https://doi.org/10.15373/22501991/june2014/18>
- Matlock-Hetzel, S. (1997). Basic Concepts in Item and Test Analysis Susan Matlock-Hetzel Texas. *Texas A&M University*. <https://www.ptonline.com/articles/how-to-get-better-mfi-results>.