Evaluating Policy Outcomes in First-Cycle Basic Education (Grades 1–4) through Fourth-Grade Reading Comprehension in South Sharqiyah Governorate, Oman

Turkeya Mohammed Khamis Al Uraimi[1], Ahmad Yunus Mohd Noor (Ph.D)[2], Dr. Zamri Arifin[3]

1 Institute of Islam Hadhari, Universiti Kebangsaan Malaysia (UKM), 43600 Bangi, Selangor, Malaysia; Email: p137389@siswa.ukm.edu.my

2 Associate Senior Research Fellow at the Institute of Islam Hadhari & Associate Professor of Theology & Philosophy, Faculty of Islamic Studies, Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor, Malaysia; Email: a_yunus@ukm.edu.my

3 Institut Islam Hadhari, Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia.

**Corresponding author: Turkeya Mohammed Khamis Al Uraimi; Email: p137389@siswa.ukm.edu.my*

HOW TO CITE:

Turkeya Mohammed Khamis Al Uraimi, Ahmad Yunus Mohd Noor , Zamri Arifin (2026). Evaluating Policy Outcomes in First-Cycle Basic Education (Grades 1–4) through Fourth-Grade Reading Comprehension in South Sharqiyah Governorate, Oman. International Journal of Special Education, 41(17s), 1662-1682

COPYRIGHT STATEMENT:

Copyright: © 2026 Authors. Open access publication underthe terms and conditions of the Creative CommonsAttribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).

ABSTRACT:

This study analyzes the results of policies in first-cycle basic education in Oman, based on the reading comprehension of 4th grade students as a proxy for early literacy outcomes in the South Sharqiyah Governorate. It follows a policy evaluation design of mixed methods, which utilises findings from three categories of documents related to the curriculum collected during the 2025/2026 school year, 200 fourth-grade students, and 20 Arabic language teachers, as well as data from three supervisors/principals. The students' performance was evaluated using a reading comprehension test to determine their performance on a reading test which consisted of literal, inferential, critical, and creative comprehension. The distribution of comprehension activities was determined by analysing curriculum-related documents, and classroom implementation, assessment, feedback, and professional development needs were interpreted through teacher/supervisor evidence. The results indicate that the overall reading comprehension level was moderate (73.5%). The strongest dimension was literal comprehension (82.1%), followed by inferential comprehension (74.8%), and the lowest dimension of critical and creative comprehension was 69.1% and 67.8%, respectively. The results of implementing the intervention indicated that making connections to literal and inferential comprehension was more prevalent than to critical and creative comprehension, while document analysis revealed more focus on these two comprehension skills than on critical and creative comprehension. This study contributes to education policy public administration by demonstrating how grade 4 reading comprehension can function as a practical policy outcome indicator. It offers a governorate-level model for linking learning evidence with curriculum review, assessment reform,

teacher development, school monitoring, and evidence-based educational planning in the Sultanate of Oman.

Keywords: Education policy outcomes, first-cycle basic education, reading comprehension, fourth-grade students, policy evaluation, Oman, South Sharqiyah Governorate.

1. Introduction

Reading comprehension is a known language skill and a strategic measure of effectiveness for educational policy. At the policy level, students' understanding, interpretation, evaluation, and extension of written texts are indicators of the extent to which education systems meet the translation of curriculum standards, teacher preparation, assessment practices, instructional support, and school-level implementation into measurable learning outcomes. This is especially vital in the primary and secondary years of education, when reading comprehension is the basis for the further development of knowledge in subsequent years of education, the acquisition of knowledge-based economies, and the long-term development of human capital. The debate on education policy has started to move from access to school to the quality of learning and the capacity of education systems to monitor and enhance learning outcomes (UNESCO, 2015; World Bank, 2018).

Grade 4 is an important stage at which to assess early literacy as it represents the shift from learning to read to reading to learn. This is a facet highlighted in international assessment frameworks, such as the Progress in International Reading Literacy Study (PIRLS), because readers are expected to do more than decode and render literal interpretations of texts; they are expected to interpret, evaluate, and use text information for wider learning purposes (Mullis & Martin, 2021; Mullis et al., 2023). In this respect, the reading comprehension ability of Grade 4 children serves as a useful proxy measure for determining the effectiveness of early education policy in inducing meaningful learning outcomes.

Oman's first cycle of basic education (Grades 1-4) is the first step in a ladder of literacy, numeracy and learning. By the end of grade 4, students should have developed reading skills that will enable them to read and study more complex material and tasks in later grades. At this stage, the level of reading comprehension will not only be an evidence of achievement for individual students but also will become an evidence of the level of realisation of the curriculum design, development of the textbook, professionalization of the teacher, assessment process and supervision of the school in the achievement of student learning outcomes.

Some studies have been conducted on reading in the Omani context from various perspectives. Al Musawi et al. (2016) investigated the effectiveness of instructional software in increasing fourth-grade students' reading comprehension skills. Al-Mahrooqi and Denman (2016) pointed out the necessity to enhance the culture of reading in Arabic and English. Alajmi and Al Balushi (2024) studied the questions of the 4th grade Arabic textbooks based on the reading processes approach, and Emam and Omara (2025) stressed on the importance of inclusive and curriculum-based evaluation of literacy. These studies offer much insight; however, there is a need for a policy-oriented evaluation that will connect students reading outcomes with curriculum evidence, implementation of the curriculum at the school level, and feedback to the policy.

This study aims to analyse the outcomes of policy implementation in the empirical context of South Sharqiyah Governorate in the field of basic education. While the policy is designed at the central level, its implementation at the governorate level, school practices, classrooms, and assessment processes are what enable the effectiveness of education policy in Oman. Governorate-level analysis may thus provide evidence that is more specific than general national claims and can be used more effectively for targeted interventions, professional

development, school monitoring, and local educational planning.

The research problem is the need to determine whether policy investments in early basic education are evidenced by measurable learning outcomes. Although the Government of Oman has made significant efforts to improve curricula, assessment systems, and early stage basic education, there is little evidence of the impact of these policy initiatives on fourth-grade students' reading comprehension, including higher-order comprehension skills. Literal comprehension is the ability to retrieve information explicitly given; inferential, critical, and creative comprehension is the ability to extract meaning and make inferences, judgments, evaluations, and create meaning. The lower scores in the comprehension section could suggest that education policy provides good results for basic reading but less so for higher-order literacy skills needed for subsequent learning. Thus, the study relies on reading comprehension as a diagnostic policy lens rather than just an achievement score.

This study makes two contributions to the fields of education policy and public administration. First, it frames fourth-grade reading comprehension as a measurable policy product rather than merely a measure of classroom performance. Second, it connects students' performance to curriculum documents and teacher, supervisor, and school leader evidence of curriculum implementation. Third, it provides an evaluation model at the governorate level which can guide evidence-informed curriculum revision, teacher professional development, assessment reform, and early literacy monitoring in Oman (Prøitz, 2015; World Bank, 2018).

Therefore, the general goal of this study is to measure the policy results in first-cycle basic education in Oman in terms of students' reading comprehension in grade four in the South Sharqiyah Governorate. It compares the outcomes of reading comprehension across literal, inferential, critical, and creative dimensions and discusses the findings in relation to curriculum-related documents, teacher and supervisor perceptions, and policy implications related to curriculum, teacher professional development, assessment practices, and literacy monitoring. The study addresses the following research questions: (1) What is the level of fourth-grade students' reading comprehension in the South Sharqiyah Governorate across literal, inferential, critical, and creative comprehension skills? (2) To what extent do fourth-grade students' reading comprehension outcomes reflect the intended policy outcomes of first-cycle basic education in Oman? (3) Which reading comprehension skills represent the strongest and weakest outcomes among fourth-grade students in South Sharqiyah Governorate? (4) What policy implications can be drawn from fourth-grade reading comprehension outcomes for curriculum development, teacher professional development, assessment practices, and literacy monitoring in first-cycle basic education?

2. Literature Review and Policy Context

2.1 Reading comprehension as a public outcome

Reading comprehension is typically studied as an instructional or cognitive ability, but in the context of policy research in education, it is also considered evidence of a public education system. Student reading comprehension reflects curriculum design, teacher preparation, teaching implementation, assessment systems, school support, and monitoring systems. Thus, it is a valid measure of the success of education policy in generating the learning outcomes that it aims to deliver.

In recent years, educational policy has become more focused on learning outcomes than on access to learning. Reading comprehension is at the centre of this agenda, as it underlies learning in all subject areas and becomes the basis for further learning. From a public administration perspective, reading comprehension may be interpreted as a measure of performance indicators for education services, as it is an indicator of how policy intentions are implemented in the classroom and how students are impacted by it (Prøitz, 2015; UNESCO, 2015; World Bank, 2018).

2.2 First-cycle basic education in Oman

Cover Grades 1-4 in Oman is the first cycle of basic education where students develop the core literacy, numeracy and learning skills. This stage is for students who are ready for higher-level

academic work. Grade 4 was found to be an appropriate grade level to examine early literacy outcomes, as noted by Al-Mousawi et al. (2016) about the importance of reading comprehension and reading aloud for fourth grade students in the OCC1 basic education first cycle in the Omani context.

Oman has committed significant resources to curriculum development, support for instruction, digital learning programs, and assessment enhancement. However, evidence of learning is a key aspect of continuous assessment for policy reform. Fourth-grade reading comprehension is a useful metric for evaluating the effectiveness of the first cycle of early education policy in meeting the desired outcomes of literacy learning. Existing studies in Oman have explored literacy assessment practices, reading culture, instructional software, and textbook questions but have not comprehensively investigated the relationship between student performance, documents, implementation evidence, and policy feedback as indicators of Omani reading comprehension policies in grade 4 (Al- Emam & Omara, 2025).

2.3 Fourth-Grade Reading Comprehension And International Assessment Logic

PIRLS emphasises Grade 4 as a pivotal stage in the transition from learning to read to reading to learn. At this level, students should not only retrieve information from texts, but also interpret, integrate, evaluate, and apply meaning to texts (Mullis & Martin, 2021; Mullis et al., 2023). This multidimensional approach to reading is helpful in policy assessment to demonstrate that reading attainment cannot be simply equated with basic reading comprehension.

In this study, reading comprehension is divided into four dimensions that are prevalent in educational settings in Arabic: literal, inferential, critical, and creative reading comprehension. The dimensions allowed this study to determine whether basic education policy is fostering foundational and advanced literacy outcomes in first-cycle programs. These two skills, critical and creative comprehension, are particularly significant as they involve the ability to assess the meaning, justify an interpretation, and make new responses to text to better reflect students' comprehension abilities.

2.4. Dimensions of reading comprehension

Literal comprehension is the student's ability to recognise specific information provided in a text, such as facts, details, sequence, characters, setting, and direct meaning. High literal comprehension suggests that students' reading outcomes are being met through the curriculum and classroom instruction. Inferential comprehension involves students' interpretation of the implicit meanings that may be drawn from a text by relating and connecting text cues and past experiences, determining relationships not explicitly stated, making predictions, and explaining cause and effect. This demonstrates a progression from recall to text analysis.

Critical comprehension is the students' skill in evaluating, judging, and reflecting on textual content. It involves judging the credibility of texts, distinguishing between facts and opinions, judging arguments or characters' actions, interpreting the author's purpose, and making judgements with evidence from the text. The dimension is aligned with PIRLS processes that involve assessing and commenting on content and text features (Mullis & Martin, 2021). Similarly, the findings from textbook analyses indicate the need to balance comprehension processes, such as higher-order evaluative skills (Alajmi & Al Balushi, 2024). Weakness in critical comprehension could be an indicator of building questioning skills, dialogic reading, formative assessment, and evidence-based reasoning (Li & Xiao, 2025; Rojas-Torres et al., 2021).

Creative comprehension is the students' skill to interpret a text and develop their own ideas. Examples of this could include recommending other endings, offering solutions, making alternative suggestions, linking the text to new contexts, or forming their own responses and responses through understanding the text. At the policy level, it is an indicator of the ability of the education system to enable the independent and creative use of knowledge.

2.5 Research gap

The literature reviewed indicates that 4th grade is a critical time to measure early literacy outcomes and that reading comprehension is a critical measure of the quality of learning. However, the majority of current research examines pedagogical, curriculum, and/or assessment aspects individually, and there remains a gap in studies that examine fourth-grade reading comprehension as an integrated policy outcome in Oman. Specifically, limited research has linked student performance data with curriculum-related documents, teacher/supervisor implementation evidence, and governorate-level policy implications. This study addresses this gap by integrating quantitative student performance data with document analysis and implementation evidence to provide a policy evaluation model linking policy inputs, classroom implementation, learning outcomes, and policy feedback.

3. Conceptual Framework

The conceptual framework is grounded in a policy outcome logic linking educational policy design, implementation mechanisms, measurable learning outcomes, and policy feedback. This study does not examine fourth-grade reading comprehension as an isolated classroom achievement but as an empirical indicator of the outcomes of the first-cycle basic education policy. Learning outcomes are increasingly viewed as evidence of how policy intentions are translated into practice through curriculum, assessment, teacher capacity, school resources, and monitoring systems (Prøitz, 2015; Viennet & Pont, 2017; Veselý, 2018).

The framework assumes that policy outcomes are not produced by curriculum documents alone. They emerge from the interaction between policy inputs and implementation mechanisms. The curriculum standards, Arabic language curriculum, teacher preparation, assessment guidelines, school resources, and supervision practices were considered policy inputs in this study. These inputs impact classroom teaching, reading, formative assessment, teacher feedback, and monitoring at the school level. These mechanisms influence students' measurable reading comprehension outcomes on the literal, inferential, critical and creative aspects of reading (Bhatti & Iqbal, 2026).

Figure 1 presents the conceptual framework that guided this study. It shows how policy inputs are translated through implementation mechanisms into learning outcomes and how these outcomes generate policy feedback for curriculum revision, teacher training, assessment improvement, targeted interventions, and resource allocation.

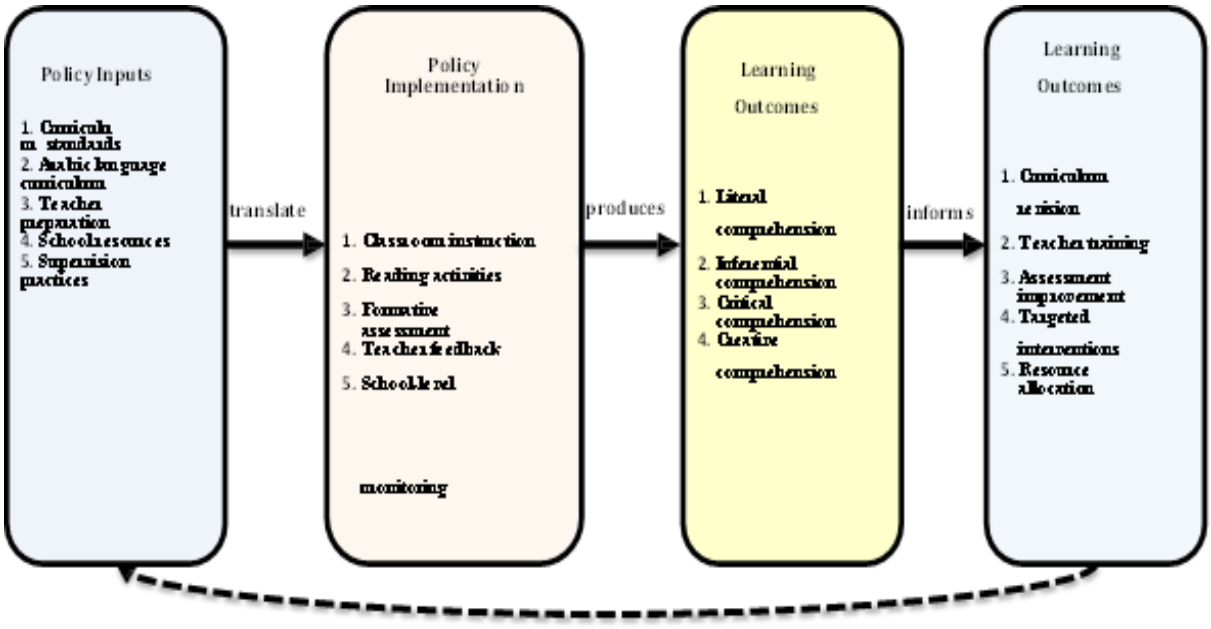

Figure 1: Conceptual Framework of the Study

Note. Policy inputs are translated through implementation mechanisms into measurable reading comprehension. These outcomes provide policy feedback for curriculum revision, teacher training, assessment improvement, targeted interventions, and resource allocation.

3.1 Policy Inputs

Policy inputs are the formal and institutional elements through which education policies are designed and communicated to schools. If curriculum standards, textbooks, assessment tools, teacher preparation, and supervision practices are weakly aligned, the learning outcomes are likely to be uneven. Conversely, coherent inputs increase the likelihood that reading comprehension outcomes will reflect the intended goals of basic education in the first cycle.

Implementation mechanisms are mechanisms of classroom and school level in which policy inputs are converted into learning. Whether or not curriculum expectations are implemented into classrooms is determined by classroom instruction, reading activities, formative assessment, teacher feedback, and monitoring by the school. Policy feedback relates to evidence-based learning to enhance policy design, implementation and resource allocation When students have a positive trend in literal comprehension and a negative trend in critical or creative comprehension, this should be considered a policy signal that calls for curriculum, assessment, and professional development responses.

The feedback part of the framework is especially relevant to the applicability of the study. In many educational systems, the data collected for learning purposes are used for reporting but not necessarily for revising curriculum, teacher development and allocation of resources. In this framework, assessment evidence is only relevant when it is utilized to identify gaps in implementation and passed on to the corrective action to be taken at the school, governorate, and ministry level.

4. Methodology

4.1 Research Design

This study adopts a mixed-methods policy evaluation design to examine the extent to which first-cycle basic education policy in Oman is reflected in fourth-grade reading comprehension outcome in South Sharqiyah Governorate. The design was appropriate because the study aimed not only to measure students' reading performance, but also to interpret that performance in relation to curriculum-related documents, teacher/supervisor implementation evidence, and policy feedback.

A mixed-methods design was also necessary because no single data source could fully explain policy outcomes. Student test results indicate the level of learning achieved, but they do not explain why some dimensions are stronger than others. Document analysis shows what types of comprehension tasks are emphasized in curriculum-related materials, while teacher and supervisor evidence clarifies how those materials are implemented in practice. Integration of these sources provides a stronger basis for policy interpretation than reliance on test scores alone.

The quantitative component consisted of a reading comprehension test administered to fourth-grade students. The documentary and qualitative components included analysis of curriculum-related documents and implementation evidence collected from Arabic language teachers and supervisors/principals. This design enabled the study to integrate learning outcome with policy inputs and implementation mechanisms (Creswell & Creswell, 2018; Creswell & Plano Clark, 2018).

Table 1. Research Alignment Matrix

Research focus	Data source	Analysis	Policy output
Overall level reading comprehension level	**Test administered to 300 Grade 4 students**	**Mmean SDs, percentage and performance level**	**Evidence on overall literacy outcomes**
Strongest and weakest dimensions	Scores for literal, inferential, critical, and creative comprehension	Dimension-level descriptive analysis and ranking	Diagnostic evidence on balanced learning outcomes
Curriculum support for comprehension outcome	Textbook, teacher's guide, activities, and remedial / enrichment plans	Content analysis by four comprehension dimensions	Evidence on curriculum alignment
Implementation factors	Teachers, supervisors, and principals	Descriptive and thematic analysis	Evidence on instruction, assessment, feedback and professional development
Policy improvement	Integrated quantitative, documentary and implementation evidence	Mixed-methods integration matrix	Recommendations for curriculum, assessment, training and resource allocation

4.2 Population and Sampling Procedures

The study was conducted during the 2025/2026 academic year in South Sharqiyah Governorate. Data were collected from 300 fourth-grade students, 20 Arabic language teachers, three supervisors/principals and four categories of curriculum-related documents. The student sample was drawn from 15 mixed government schools offering Grades 1-4.

A stratified cluster sampling strategy was used because schools represent natural educational clusters while the study required representation across different school contexts. Schools were selected according to geographical distribution, size, and accessibility. Within each school, 20 Grade 4 students were selected, whit proportional selection across classes where necessary. A separate pilot sample of 30 fourth-grade students from a similar school outside the main sample was used to examine item clarity, timing, difficulty, and scoring consistency. Teacher and supervisor/principal participants were selected purposively because of their direct role in teaching, supervising, and monitoring Arabic reading comprehension.

Table 2. Study Context and Data Sources

Component	Description
Governorate	South Sharqiyah Governorate, Oman
School type	Mixed government schools offering first-cycle basic education, Grades 1-4
Number of participating schools	15 schools
Student participants	300 fourth-grade students
Teacher participants	20 Arabic language teachers
Supervisors/principals	Three supervisors/principals
Academic year	2025/2026
Curriculum-related documents analyzed	Grade 4 Arabic textbook, teacher's guide, expression and reading activities, and remedial or enrichment plans

4.3 Instruments, Validity and Reliability

Three instruments were used: a reading comprehension test, a curriculum-related document analysis checklist, and a teacher/supervisor implementation instrument. The reading test consisted of 40 items distributed equally across four dimensions: literal, inferential, critical, and creative comprehension. Each dimension included 10 items, allowing direct comparison among comprehension skills.

Table 3. Structure of the Reading Comprehension Test

Dimension	Items	Sscore	Main skill focus
Literal comprehension	10	10	Retrieving explicit information and direct meanings
Inferential comprehension	10	10	Interpreting implicit meanings and making text-based inferences
Critical comprehension	10	10	Evaluating ideas, judging meaning, and justifying responses
Creative comprehension	10	10	Extending meaning and producing creative interpretations
Total	**40**	**40**	**Overall Grade 4 reading comprehension**

Each item was scored on a one-point basis. Correct responses received one point, while incorrect or missing responses received zero. Short-answer items, especially in critical and creative comprehension, were scored a rubric; response was credited when they were relevant, text-based, and clearly expressed. Student. Performance was classified as high (80% and above), moderately high (70%–79.9%), moderate (60%–69.9%), low (50%–59.9%) or very low (below 50%).

The use performance levels also helped translate numerical scores into policy-relevant categories. Percentages and means provide statistical summaries, but classification labels such as high, moderately high, and moderate make the findings more interpretable for curriculum planners, supervisors, and school leaders. This was important because the study aimed to support policy feedback rather than present descriptive results only.

The document analysis checklist classified reading activities and questions in this textbook, teacher's guide, expression and reading activities and remedial/enrichment plans according to the four comprehension dimensions. The teacher/supervisor instrument gathered evidence on curriculum clarity, textbook alignment, reading activities, formative assessment, feedback, questioning strategies, reading resources and professional development needs.

Validity was addressed through expert review and pilot testing. Seven experts in Arabic curriculum, reading comprehension, educational assessment, first-cycle basic education and education policy reviewed the instruments for relevance, clarity, age appropriateness, linguistic accuracy, and alignment. Items that were unclear, repetitive, or weakly aligned were modified. The pilot test with 30 students examined test clarity, timing, item difficulty, and suitability of passages. Reliability results were strong: KR-20 was 0.86 for the reading test, Cronbach's was 0.88 for the teacher/supervisor instrument, and scorer agreement for open-ended responses and document analysis reached 0.91 (Boateng et al., 2018; Taber, 2018).

These procedures were especially important because the study measured both basic and higher-order comprehension. Literal and inferential items are relatively direct to classify, whereas critical and creative items require careful rubric design to avoid subjective scoring. The use of expert review, pilot testing, and scorer agreement therefore strengthened the credibility of the test and ensured that higher-order responses were evaluated consistently.

4.4 Data Collection and Analysis Procedures

Data collection was conducted after obtaining institutional permissions. The reading comprehension test was administered under standardized classroom conditions. Curriculum-related documents were collected and analyzed separately, and teacher/supervisor data were collected using the implementation instrument.

Quantitative data from the student test and teacher/supervisor instrument were analyzed using IBM SPSS statistics, version 26. Descriptive statistics, including frequencies, percentages, means, standard deviations, performance levels, and rankings, were used to analyze student performance and implementation responses. This descriptive approach was appropriate because the study aimed to evaluate the level and pattern of reading outcomes rather than test causal relationships.

Document analysis was conducted through content analysis. Reading activities and question were coded according to literal, Microsoft Excel was used to organize the document-analysis matrix, calculate frequencies and percentages, and compare distributions across document categories. Open-ended responses and qualitative notes were coded manually using a structured thematic coding sheet. The final stage involved mixed-methods integration, where student results, document analysis, and implementation evidence were combined in an integrated findings matrix.

4.5 Ethical Considerations

The study followed ethical principles for educational research involving children and school personnel. Approval was obtained from

the relevant educational authority and schools. Because the study involved fourth-grade students, parental consent and student assent were obtained. Teacher, supervisor and principal participation was voluntary. Participants were informed of the study purpose, confidentiality protections, and their right to withdraw. All data were coded, stored securely, and reported in aggregate form. No student, teacher, supervisor, principal, or school was identified (British Educational Research Association, 2018). A formal ethical approval statement is also provided in the Declarations section.

5. Findings

The findings follows the conceptual logic shown in Figure 1 and the methodological alignment Table 1. The analysis begins with student reading outcomes, followed by curriculum analysis, implementation evidence, and an integrated policy interpretation.

5.1 Student Reading Comprehension Performance

The overall results show that fourth-grade students achieved a moderately high level of reading comprehension. The overall mean was 29.38 out of 40, with a standard deviation of 5.82 and a performance percentage of 73.5%. This suggests that first-cycle basic education policy in South Sharqiyah Governorate is producing generally positive early literacy outcomes, although dimension-level results are needed for diagnostic interpretation.

Table 4. Student reading comprehension performance

Dimension	Items	Max score	Mean	SD	Percentage	Level	Rank
Literal comprehension	10	10	8.21	1.34	82.1%	High	1
Inferential comprehension	10	10	7.48	1.51	74.8%	Moderately high	2
Critical comprehension	10	10	6.91	1.63	69.1%	Moderate	3
Creative comprehension	10	10	6.78	1.70	67.8%	Moderate	4
Overall reading comprehension	40	40	29.38	5.82	73.5%	Moderately high	-

Literal comprehension was strongest dimension at 8.21%, indicating that students were generally able to retrieve explicit information, identify stated details and recognise direct meanings. Inferential comprehension ranked second at 74.8%, showing that students were able to interpret implied meanings and make text-based connections to a reasonably positive extent.

Critical comprehension recorded a moderate level at 69.1%, indicating that students were less confident in evaluating ideas, identifying author purpose, judging textual content, and justifying responses with evidence. Creative comprehension was the weakest dimensions at 6.78%, suggesting difficulty in extending textual meaning, generating alternative responses, and producing creative interpretations. Overall, outcomes were positive but uneven, with stronger performance in foundational and interpretive comprehension than in higher-order comprehension.

5.2 Curriculum-Related Document Analysis

To interpret student results, four categories of curriculum-related documents were analysed.

The analysis examined distribution of reading activities and questions across literal, inferential, critical, and creative comprehension.

Table 5. Distribution of reading comprehension skills in analyzed documents

Document type	Total activities / questions analyzed	Literal	Inferential	Critical	Creative
Grade 4 Arabic textbook	120	43 (35.8%)	37 (30.8%)	22 (18.3%)	18 (15.1%)
Teacher's guide	55	17 (30.9%)	16 (29.1%)	12 (21.8%)	10 (18.2%)
Expression and reading activities	45	16 (35.6%)	13 (28.9%)	8 (17.8%)	8 (17.8%)
Remedial/enrichment plans	20	10 (50.0%)	4 (20.0%)	3 (15.1%)	3 (15.0%)
Total	240	86 (35.8%)	70 (29.2%)	45 (18.8%)	39 (16.2%)

Literal comprehension received the greatest emphasis, representing 35.8% of all activities analyzed and questions. Inferential comprehension represented 29.2%, while critical and creative comprehension accounted for 18.8% and 16.2%. respectively. This distribution is consistent with student performance: the dimensions most emphasized in curriculum-related documents were also the dimensions in which students performed more strongly. From a policy perspective, the lower representation of critical and creative comprehension suggests the need for curriculum review to ensure a more balanced distribution of reading tasks.

5.3 Teacher/Supervisor Implementation Evidence

Table 6. Teacher and supervisor implementation evidence

Implementation factor	Mean	SD	Level
Clarity of reading comprehension objectives in the curriculum	4.12	0.61	High
Alignment between textbook activities and reading objectives	3.98	0.68	High
Use of classroom reading activities	4.05	0.64	High
Use of formative assessment in reading	3.91	0.72	High

Implementation factor	Mean	SD	Level
Teacher feedback on students' reading errors	4.10	0.59	High
Use of inferential questioning strategies	3.76	0.74	Moderately high
Use of critical reading strategies	3.54	0.81	Moderate
Use of creative reading activities	3.47	0.84	Moderate
Availability of reading resources	3.73	0.79	Moderately high
Need for further professional development	4.36	0.56	High

Implementation evidence indicates generally positive conditions for teaching reading comprehension. Teachers and supervisors/principals reported high levels of curriculum clarity, textbook alignment, classroom reading activities, formative assessment, and teacher feedback. However critical reading strategies and creative reading activities were rated at a moderate level, corresponding with lower student performance in those dimensions. The highest mean was recorded for further professional development, indicating a need for targeted training in higher-order reading strategies.

5.4 Integrated Findings

Table 7. Integrated findings matrix

Dimension	Student performance level	Document emphasis	Policy interpretation	Recommended response
Literal	High, 82.1%	Highest emphasis, 35.8%	Foundational reading outcomes are well supported.	Maintain basic reading support while increasing higher-order tasks.
Inferential	Moderately high, 74.8%	Strong emphasis, 29.2%	Students show positive interpretive ability but need systematic support.	Expand guided questioning, prediction and text-based inference activities.
Critical	Moderate, 69.1%	Moderate emphasis, 18.8%	Evaluative reading remains less developed.	Increase evaluative questions, reasoning tasks and teacher training.

Dimension	Student performance level	Document emphasis	Policy interpretation	Recommended response
Creative	Moderate, 67.8%	Lowest emphasis, 16.2%	Creative reading outcomes require further development.	Expand open-ended, imaginative and enrichment-based reading tasks.
Overall	Moderately high, 73.5%	Positive but uneven alignment	Policy outcomes are positive but uneven.	Use dimension-specific monitoring to guide curriculum, assessment and training.

The integrated findings show that fourth-grade reading comprehension outcomes in South Sharqiyah Governorate are generally positive but uneven. Literal and inferential comprehension are better supported by curriculum materials and classroom practices, whereas critical and creative comprehension requires stronger policy attention. The convergence between student results, document analysis, and teacher/supervisor evidence indicates that reading outcomes are shaped by curriculum balance, instructional practices, assessment routines, and professional support.

The findings also suggest that improvements should be sequenced rather than implemented as a single broad reform. The first step is to make higher-order comprehension more visible in the curriculum materials and assessment tools. Second, teachers should be equipped with practical strategies for teaching and evaluating these skills. Third, progress is monitored through school and governorate-level reading profiles. This sequence allows policy actors to move from diagnosis to implementation and then to continuous feedback.

This convergence is one of the most important empirical contributions of this study. The lower performance in critical and creative comprehension isn't as isolated student-level result; it is supported by the lower presence of these dimensions in documents and by moderate implementation ratings for critical and Creative reading activities. Therefore, the improvement priority is not based on a single data source but on a consistent pattern across quantitative, documentary, and practitioner evidence.

6. Discussion

This section discusses the findings from a policy-oriented perspective. The aim isn't only to interpret student performance as an educational achievement result, but also to examine what the results reveal about the effectiveness and future direction of first-cycle basic education policy in South Sharqiyah Governorate. The discussion follows the relationship between policy inputs, implementation mechanisms, learning outcomes and policy feedback shown in Finger 1.

The overall performance rate of 73.5% provides positive evidence that early literacy policy in South Sharqiyah Governorate is producing generally acceptable learning outcomes by the end of Grade 4. This is important as reading comprehension forms a crucial component of learning which influences students' learning ability in subsequent learning stages. This finding is in line with the international focus on Grade 4 as a transition between learning to read and

reading to learn (Mullis & Martin, 2021; Mullis et al., 2023).

The dimension-level results, however, show varying policy outcome. Students' performance in literal comprehension was best, followed by their performance in inferential comprehension and lowest in critical and creative comprehension. The results indicate that first cycle basic education policy performs better on achieving basic and interpretive comprehension of the texts than on increasing higher order reading skills. The findings show both insight and need: there is good understanding of the basics; but there is a need to develop further understanding of more complex outcomes in the areas of evaluation, reasoning, judgment and creative extension.

This trend can be attributed to the document analysis. Materials that were linked to the curriculum most emphasized literal and inferential comprehension, and critical and creative comprehension were less represented. This distribution is similar to that of student performance; this indicates that reading outcomes are influenced by the organization of curriculum materials, the tasks assigned in class and the expectations held for assessment of reading performance. The curriculum-related documents serve as policy tools due to their impact on the focus of teachers and on the opportunities provided for students to practice them (Viennet & Pont, 2017; Veselý, 2018).

Not only does this finding highlight the significance of curriculum materials as policy instruments, it also underscores the necessity of considering them in that light. Teachbooks, guides and activity plans convey policy priorities to teachers and influence the opportunities given to students to learn. The less visible the application of higher order comprehension skills, the fewer the prompts will be available in these materials for the teacher's use when developing evaluation, justification and creative response skills. It is not to suggest that such skills are not addressed by the teacher, but rather that policy documents might be made more explicit on how to address these skills.

The evidence of the teacher and supervisor is consistent with this interpretation. The overall curriculum clarity, the alignment of curriculum with the textbooks, the classroom reading activities and formative assessment and teacher feedback were scored high. The results suggest a generally positive implementation setting. The lower rating in critical reading strategies and creative reading activities, however, indicate that it has been implemented well for basic comprehension, but not for higher-order comprehension. This indicates that there is some evidence of implementation of policy, but that this is not consistent for all literacy outcomes to be achieved.

The results suggest a strong linkage between students' achievement and instructional focus presented in documents of the curriculum. A document analysis was conducted and revealed that the highest level of emphasis was on literal comprehension then inferential comprehension. In contrast, the critical and creative comprehension were under-represented. This distribution resembles the student's performance in terms of the higher level of literal and inferential comprehension as compared to critical and creative comprehension.

This alignment implies that reading outcomes are influenced by the organisation of reading materials, tasks in the classroom and the demands for assessment. The Curriculum-related documents are policy documents in the sense that they affect what teachers express in their lessons, and what students are offered a chance to do (Viennet & Pont, 2017; Veselý, 2018). If students are encouraged to attend to literal and inferential tasks, they will be more likely to be good at these tasks. On the other hand, if the critical and creative aspects of a task are not stressed, students may have fewer opportunities to evaluate, to justify, and to extend the creative aspects.

Teacher and supervisor evidence corroborates the meaning. Overall, respondents indicated favorable conditions for implementing the curriculum in terms of: clarity of the curriculum, alignment with the curriculum books, use of

classroom reading, formative assessment, and teacher feedback. The lower ratings for critical reading strategies and creative reading activities indicate that implementation is more robust in the basic comprehension skills than in higher order comprehension skills. This is in line with Alajmi and Al Balushi (2024), who found that not enough balance was observed in the distribution of the questions in the Omani fourth-grade Arabic textbooks.

The weaker performance in critical and creative comprehension must be considered a policy message, not an issue in the classroom. These are all elements of "higher order" comprehension that must be purposefully planned, taught, and assessed. International reading frameworks not only focus on information retrieval but also interpretation, integration, evaluation, and critique of reading (Mullis & Martin, 2021). Other studies also reveal that reading instruction strategies used by teachers and that which fosters inference, self-monitoring, evaluation and reflective practice have a relationship with reading achievement (Li & Xiao, 2025; Rojas-Torres et al., 2021).

The results are also directly applicable to the governance of public administration and education. First of all they facilitate the systematical monitoring of policies. Overall achievement is a good indicator of general achievement, while dimension-level achievement is more diagnostic. Secondly, they enable developmental accountability, using data to enhance the system, not just to assess performance. Third, they facilitate evidence-based decision making at the Governorate level as students' performance and document analysis and implementation evidence are parallel and have the same priority: the positive outcomes for foundational comprehension are not sufficient enough, whereas critical and creative comprehension requires more support. Lastly they note that teacher professional development is a governance mechanism that should be linked to the curriculum expectations, design of assessment, classroom observations, and evidence of student learning (OECD, 2019).

The study also highlights from a governance perspective the benefits of shifting from achieving monitoring to dimension specific monitoring. One overall mark could be good, but may indicate uneven progress in comprehension skills. Dimension-specific data offer more actionable information for both the supervisor and the curriculum planner or school leader as they will indicate the specific learning processes that need support.

Overall, the results indicated that the results of first-cycle basic education (live) in the South Sharqiyah Governorate are having positive but mixed impact on literacy. That current policy is failing is not the appropriate conclusion and that the overall score is a measure of its success is not the answer. Instead, the current literacy policy has been implemented and it serves as a good starting point, which can be enhanced and improved to promote deeper comprehension. More critical and creative tasks should be included in the curriculum policy; all the dimensions of comprehension should be included in the assessment policy; dimension-specific outcomes should be monitored at the school level; and the planning at the governorate level should be based on reading data for training, supervision and resource distribution.

7. Policy Implications

The findings have implications for curriculum policy, teacher professional development, assessment reform, school-level monitoring, and governorate-level educational planning. The central policy implication is that fourth-grade reading comprehension should be used not only as a student achievement measure but also as a practical indicator for evaluating the outcomes of first-cycle basic education policies.

The dimensions of reading comprehension should be more balanced in Grade 4 Arabic language materials, as per the curriculum policy. Textbook questions, teacher guide activities, classroom tasks, and enrichment plans should offer more opportunities for students to analyse and determine their interpretation, make comparisons and contrasts, justify author purpose, make alternative endings, and generate

creative responses with evidence from the text. Literal comprehension is not diminished but rather serves as a foundation to build upon for the interpretation, evaluation, and creative application of texts ((Rawashdeh & Idris, 2025).

Therefore, the focus of the practical policy framework should be on coordination and not on a list of individual recommendations. There is one body of evidence, and it is for everyone involved in curriculum planning, assessment units, training units, supervisors, and school leaders. This coordinated response is important to ensure that assessment reform does not stand alone and impact classroom instruction, teacher training is not in line with curriculum gaps, and school monitoring is not linked to real learning needs.

This coordination is especially crucial at the governorate level. If curriculum revision, assessment design, and professional development are not coordinated, efforts to improve the curriculum may be fragmented. The framework allows for a more unified approach to prioritising training, providing support in the classroom, monitoring the school, and allocating resources in response to the same identified learning needs (Lamin, 2026).

Teacher professional development is not just about teaching reading but also about the strategies used to comprehend reading. Training should support teachers in developing questioning techniques, dialogic reading, using evidence from the text, providing formative feedback, implementing differentiated instruction, and open-ended reading tasks to create inferential, critical, and creative reading tasks. It is also important that professional learning is ongoing and that it is supported by classroom coaching, peer observation, model lessons, school-based professional learning communities, and supervisor feedback (Alabay, 2026).

The aim of the assessment policy should be to offer a fair and diagnostic reading comprehension assessment. Grade 4 assessments should contain a mix of literal, inferential, critical and creative items to prevent overrating students' reading ability from literal recall only. Governorate-level reading diagnostics could be implemented periodically to focus the curriculum on the needs of students, inform teacher-training priorities, and aid in identifying schools for support. These evaluations should be formative rather than summative.

At school, students' reading profiles should be created, indicating student performance in the four dimensions. These profiles would be useful for principals, senior teachers, and Arabic language teachers to pinpoint areas needing support. Governorate-level planning should leverage periodic Grade 4 reading data to understand trends by school, determine resource allocation, design and plan interventions, and monitor curriculum implementation.

Table 8. Practical Policy Framework for Improving Fourth-Grade Reading Comprehension Outcomes

Policy area	Evidence from the study	Recommended policy action	Responsible level	Expected outcome
Curriculum policy	Critical and creative comprehension were less represented in documents (18.8% and 16.2%).	Revise Grade 4 Arabic materials to balance literal, inferential, critical and creative tasks.	Ministry curriculum planners and Arabic curriculum specialists	More balanced development of basic and higher-order reading skills

Policy area	Evidence from the study	Recommended policy action	Responsible level	Expected outcome
Teacher professional development	Teachers and supervisors reported a high need for further development (M = 4.36).	Provide targeted training on inferential questioning, critical reading, creative response, formative feedback and differentiated instruction.	Governorate training units, supervisors and school-based professional learning communities	Improved teacher capacity to develop higher-order comprehension
Assessment policy	Student results were stronger in literal comprehension (82.1%) than in critical and creative comprehension.	Develop a Grade 4 reading assessment framework aligned with national goals and PIRLS logic.	Ministry assessment units, curriculum departments and governorate teams	More accurate diagnosis of reading strengths and weaknesses
School-level monitoring	Implementation evidence showed moderate use of critical and creative reading activities.	Establish school-level reading profiles tracking performance across the four dimensions.	Principals, senior Arabic teachers and school improvement teams	More targeted remedial and enrichment planning
Governorate planning	The study provides evidence from 15 schools and 300 students.	Use periodic Grade 4 reading data to guide training, supervision, resource allocation and school support.	South Sharqiyah Directorate of Education and policy planners	Evidence-based regional planning
Targeted interventions	Critical and creative comprehension recorded moderate levels.	Introduce enrichment programs, reading clubs and structured higher-order comprehension tasks.	Schools, supervisors and Arabic language teachers	Improved critical and creative comprehension
Policy feedback	Student, document and practitioner evidence point to the same priorities.	Create a feedback loop linking outcomes, curriculum review, teacher training, assessment reform and monitoring.	Ministry, governorate directorate, supervisors and school leaders	Institutionalized use of learning evidence for continuous improvement

The framework converts reading comprehension evidence into practical governance actions. It reinforces the argument that Grade 4 reading comprehension can be used as a public-sector performance indicator for improving first-cycle basic education policy.

8. Conclusion

The current study aimed to assess the outcome of the policies in first-cycle basic education in Oman by studying the reading comprehension achievement level of fourth-grade students in South Sharqiyah Governorate. Data from 300 students, 20 Arabic language teachers, three supervisors/principals, and four types of documents related to the curriculum were used in the study, which focused on the four levels of comprehension (literal, inferential, critical, and creative comprehension). The heart of its thesis is that 4th grade reading comprehension is not a quantifiable measure of the translation of education policy to learning.

The results indicate that students' reading comprehension was moderate at 73.5%. This offers positive evidence for the production of meaningful literacy outcomes by the end of grade four. However, the results of the dimensions are more complicated. Literal and inferential comprehension were better than critical and creative comprehension [insert reference]. Document analysis and evidence of implementation verified that the same is true for foundational comprehension, which is more supported by curriculum materials and classroom practice than higher-order comprehension.

The study finds that the basic education policy for the first cycle is generally effective in the South Sharqiyah Governorate, but it is not well balanced in terms of all reading outcomes. The right policy answer is not a policy change but policy refinement. There are opportunities for improvements to be made in relation to curriculum balance, professional development, diagnostic assessment, school-level reading profiles, targeted interventions, and governorate-level planning based on the evidence of learning. The findings outlined in Table 8 offer a set of practical policy recommendations for coordinated action in curriculum policy, teacher development, assessment reform, monitoring, resource allocation, and policy feedback.

However, the significance of this conclusion for decision-makers is that the evidence is specific. The study does not prescribe the improvement of literacy in general, but specifies the dimensions that need to be supported and relates this to curriculum representation, classroom practice, curriculum assessment design, and teacher development. This is highly actionable for further school improvement and governorate planning.

This study contributes to education policy administration by developing the concept of reading comprehension as a policy outcome indicator, combining student performance with document and implementation evidence, and providing a model for the improvement of literacy policies based on evidence at the governorate level. The use of reading comprehension data in the institutionalisation of curriculum review, teacher training, assessment policy, and planning at the governorate level can shift the education system from a general indication of a literacy goal to a more specific and evidence-based approach to improving early learning outcomes.

9. Limitations and Future Research

There are several caveats to consider. First, the study was conducted in only one governorate, South Sharqiyah; therefore, the results of the study cannot be extrapolated to other governorates without moderation. Comparative studies of similar patterns across other governorates should be conducted in the future to determine whether there is a pattern of similarity throughout the country.

Second it does not demonstrate the progression of reading comprehension from grades 1-4, which is the focus of basic education. Longitudinal studies could follow students' reading development over the first cycle and determine when differences in the dimensions of comprehension start to occur.

Third, the cross-sectional design cannot claim causal relationships between curriculum materials

and teacher practices on the one hand, and student outcomes on the other, as it only offers a snapshot from the 2025/2026 school year. Longitudinal or quasi-experimental designs can be used in future studies to examine the impact of curriculum revision, teacher training, or assessment intervention.

Fourth, curriculum-related materials were selected for document analysis, and implementation evidence was collected from teachers, supervisors, and principals. Classroom observations, student and parent feedback, national policy documents, supervision reports, and school improvement plans may be included in additional studies to provide a broader picture of the implementation of reading policies.

Fifth, word knowledge, fluency, motivation, prior knowledge, home literacy environment, classroom interaction, and digital exposure are factors that may affect reading comprehension. It would be interesting to find out in future research how these factors are related to the comprehension results, as well as whether or not they are able to account for differences between schools, students, and governorates.

These restrictions do not diminish the significance of the study but simply set the proper limits for interpretation. The study is best seen as a policy-level analysis carried out at the governorate level and diagnostic information for South Sharqiyah, and a model that can be adapted in other locations. Its value is not in providing a snapshot of learning outcomes, curriculum materials, and evidence of implementation, but in offering a glimpse of how learning outcomes, curriculum materials, and evidence of implementation can be linked to inform policy improvement. Second, this study limited the consideration to Grade 4. This is an appropriate focus as Grade 4 is the end of first-

References

Al Musawi, A. S., Kazem, A. M., Al Hashmi, A., & Al Busaidi, F. Y. (2016). The effectiveness of instructional software in reading comprehension skills and reading aloud of Omani fourth basic schools' students. *Technology, Innovation and Education, 2*, Article 6. https://doi.org/10.1186/s40660-016-0018-0

Alabay, S. (2026). Majestic Multidisciplinary Science Insight Teacher Preparedness for Inclusive Classrooms: Evaluating Competencies, Challenges, Training Needs and Barriers to Educational Inclusion in Initial Teacher Training. *Https://Cmpublisher.Com/Wp-Content/Uploads/2026/03/Sercan_Alabay.Pdf*, *1*, 1–14. https://doi.org/10.65080/inprocess

Alajmi, M. S., & Al Balushi, S. M. M. A. (2024). Reading comprehension questions in fourth grade Arabic language textbooks in Oman: An analysis from the perspective of Progress in International Reading Literacy Study. *Theory and Practice in Language Studies, 14*(11). https://doi.org/10.17507/tpls.1411.04

Alghafri, Y., & Yusuf, M. S. (2024). The impact of using the interactive board on developing academic achievement in reading comprehension and its relationship to reading comprehension levels from the perspective of fifth-grade students in Muscat Governorate, Sultanate of Oman. *Ibn Khaldun Journal for Studies and Research, 4* (8). https://doi.org/10.56989/benkj.v4i8.1146

Al-Mahrooqi, R., & Denman, C. (2016). Establishing a reading culture in Arabic and English in Oman. *Arab World English Journal, 7*(1), 5–17. https://doi.org/10.24093/awej/vol7no1.1

Bhatti, M., & Iqbal, I. (2026). Competency-based curriculum models for higher education: A comparative study across global frameworks. *Majestic Multidisciplinary Science Insight*, *1*, 1. https://doi.org/10.65080/msi.v1.cm2601107005

Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quinonez, 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*, Article 149. https://doi.org/10.3389/fpubh.2018.00149

Braun, V., & Clarke, V. (2021). *Thematic analysis: A practical guide*. SAGE Publications. https://uk.sagepub.com/en-gb/eur/thematic-analysis/book248481

British Educational Research Association. (2018). *Ethical guidelines for educational research* (4th ed.). https://www.bera.ac.uk/publication/ethical-guidelines-for-educational-research-2018

Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). SAGE Publications. https://us.sagepub.com/en-us/nam/research-design/book255675

Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research* (3rd ed.). SAGE Publications. https://us.sagepub.com/en-us/nam/designing-and-conducting-mixed-methods-research/book241842

Emam, M. M., & Omara, E. M. N. (2025). Decolonising literacy assessments in Oman: Curriculum-based measurements and power dynamics. *British Journal of Special Education*. https://doi.org/10.1111/1467-8578.70029

Gill, B., Lerner, J. S., & Meosky, P. (2016). Reimagining accountability in K–12 education. *Behavioral Science & Policy, 2*(1), 57–70. https://doi.org/10.1353/bsp.2016.0007

Hartati, M., Kurniady, D. A., Sa'ud, U. S., & Nurdin, D. (2025). CIPP evaluation on the literacy program in improving reading proficiency in senior high school. *Al-Ishlah: Jurnal Pendidikan, 17*(3). https://doi.org/10.35445/alishlah.v17i3.7412

Honorato-Errázuriz, J., & Ramírez-Montoya, M. S. (2020). Innovative public policies in the development of reading skills in basic education: Regarding the systematic mapping protocol. In *Proceedings of the Eighth International Conference on Technological Ecosystems for Enhancing Multiculturality*. Association for Computing Machinery. https://doi.org/10.1145/3434780.3436575

https://doi.org/10.1596/978-1-4648-1096-1

Krippendorff, K. (2019). *Content analysis: An introduction to its methodology* (4th ed.). SAGE Publications. https://us.sagepub.com/en-us/nam/content-analysis/book258450

Lamin, M. (2026). Applying design thinking for achieving national education policy objectives through AI-driven solutions. *Majestic International Journal of AI Innovations*, *1*. https://doi.org/10.65080/inprocess

Li, Z., & Xiao, Y. (2025). Unpacking the influence of teachers' reading instruction strategies on primary school students' reading achievement: Evidence from PIRLS 2021. *Reading Research Quarterly*. https://doi.org/10.1002/rrq.70064

Mullis, I. V. S., & Martin, M. O. (Eds.). (2021). *PIRLS 2021 assessment frameworks*. International Association for the Evaluation of Educational Achievement. https://pirls2021.org/frameworks/

Mullis, I. V. S., von Davier, M., Foy, P., Fishbein, B., Reynolds, K. A., & Wry, E. (2023). *PIRLS 2021 international results in reading*. International Association for the Evaluation of Educational Achievement. https://pirls2021.org/results

OECD. (2019). *TALIS 2018 results (Volume I): Teachers and school leaders as lifelong learners*. OECD Publishing. https://doi.org/10.1787/1d0bc92a-en

Prøitz, T. S. (2015). Learning outcomes as a key concept in policy documents throughout policy changes. *Scandinavian Journal of Educational Research, 59*(3), 275–296. https://doi.org/10.1080/00313831.2014.904418

Rawashdeh, A., & Idris, M. (2025). Synchronized Curriculum Adaptation: A Strategy for Enhancing Skill Development in Accounting Education. *Journal of International Education in Business*, *18*(2), 269–288. https://doi.org/10.1108/jieb-01-2024-0001

Rojas-Torres, L., Ordóñez, G., & Calvo, K. (2021). Teacher and student practices associated with performance in the PISA reading literacy evaluation. *Frontiers in Education, 6*, Article 658973. https://doi.org/10.3389/feduc.2021.658973

Taber, K. S. (2018). The use of Cronbach's alpha when developing and reporting research instruments in science education. *Research in Science Education, 48*, 1273–1296. https://doi.org/10.1007/s11165-016-9602-2

UNESCO. (2015). *Education 2030: Incheon Declaration and Framework for Action for the implementation of Sustainable Development Goal 4*. UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000245656

Veselý, A. (2018). Conceptual framework for analysis of relationship between educational policy and educational outcomes. *Orbis Scholae, 5*(1), 23–52. https://doi.org/10.14712/23363177.2018.73

Viennet, R., & Pont, B. (2017). *Education policy implementation: A literature review and proposed framework*

World Bank. (2018). *World development report 2018: Learning to realize education's promise*. World Bank.

Author Declaration

Ethical approval: Ethical approval was obtained from the relevant educational authority before data collection. Permission was also obtained from participating schools. Because the study involved fourth-grade students, parental consent and student assent were obtained. Teachers, supervisors, and principal participation was voluntary. All data were coded, treated confidentially, and reported in aggregate form to protect the identity of participants and schools.

Data availability statement: The data supporting the findings of this study are available from the corresponding author upon reasonable request, subject to institutional approval and confidentiality restrictions related to student and school data.

Conflict of interest: The author declares that they have no conflict of interest.

Funding: This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.