

Exploring the Use of Claude AI in Reading Comprehension Assessment among Junior High School Students

Febri Yonanda Yovi¹, Ali Mal Zulfahmi², Febria Sri Artika³, Melyann Melani⁴

^{1,2,3,4}UIN Sjech M. Djamil Djambek Bukittinggi, West Sumatera - Indonesia

Email: ¹yonandayovi15@gmail.com, ²zulfahmiali17@gmail.com,

³febriasriartika@uinbukittinggi.ac.id, ⁴melyannmelani@uinbukittinggi.ac.id,

Abstract

Article History:

Received: 10/06/2026

Accepted: 28/06/2026

Published: 10/07/2026

Keywords:

*Claude AI, AI-Assisted
 Assessment, Reading
 Comprehension,
 English Language
 Teaching, Formative
 Assessment*

This study described the use of Claude AI as an assessment tool in analysing Junior High School students' reading comprehension performance. A descriptive research design was employed involving 38 students of SMP Negeri 3 Bukittinggi. Data were collected through a reading comprehension test consisting of one reading passage and five open-ended questions representing four reading indicators: literal comprehension, vocabulary in context, inferential reading, and critical thinking. Students' responses were assessed using Claude AI and analysed descriptively. The finding showed that students performed strongest in literal comprehension, moderately in vocabulary and inferential reading, and weakest in critical thinking, with an overall average score of 68%. The results suggest that students possess adequate foundational reading skills but require further support in developing higher-order comprehension abilities. Claude AI demonstrated potential for providing rapid assessment and feedback, immediate feedback, and systematic analysis of student responses. However, its ability to evaluate complex reasoning and contextual factors remains limited the study concludes that Claude AI can serve as a useful formative assessment tool when combined with teacher judgement. These findings contribute to the growing discussion of AI-assisted assessment in English Language Teaching activities.

Kata Kunci:

*Claude AI, Penilaian
 Berbasis Kecerdasan
 Buatan, Pemahaman
 Membaca,
 Pembelajaran Bahasa
 Inggris, Penilaian
 Formatif*

Abstrak

Penelitian ini bertujuan untuk mengkaji kemampuan pemahaman membaca siswa Sekolah Menengah Pertama serta mengeksplorasi penggunaan Claude AI sebagai alat penilaian dalam pembelajaran Bahasa Inggris. Penelitian ini menggunakan desain penelitian deskriptif yang melibatkan 38 siswa SMPI Islam Bukittinggi. Data dikumpulkan melalui tes

pemahaman membaca yang terdiri atas satu bacaan dan lima pertanyaan esai terbuka yang merepresentasikan empat indikator membaca, yaitu pemahaman literal, kosakata dalam konteks, membaca inferensial, dan berpikir kritis. Respons siswa dinilai menggunakan Claude AI dan dianalisis secara deskriptif. Hasil penelitian menunjukkan bahwa siswa memperoleh hasil terbaik pada pemahaman literal, cukup baik pada kosakata dan membaca inferensial, serta hasil terendah pada berpikir kritis, dengan rata-rata keseluruhan sebesar 68%. Temuan ini menunjukkan bahwa siswa telah memiliki keterampilan dasar membaca yang memadai, tetapi masih memerlukan dukungan lebih lanjut dalam mengembangkan kemampuan pemahaman tingkat tinggi. Penelitian ini juga menemukan bahwa Claude AI efektif dalam memberikan penilaian secara cepat, umpan balik langsung, serta analisis sistematis terhadap respons siswa. Namun, kemampuan Claude AI dalam mengevaluasi penalaran yang kompleks dan faktor-faktor kontekstual masih memiliki keterbatasan. Penelitian ini menyimpulkan bahwa Claude AI dapat berfungsi sebagai alat penilaian formatif yang bermanfaat apabila digunakan bersama dengan pertimbangan profesional guru. Temuan ini berkontribusi pada berkembangnya diskusi mengenai penilaian berbantuan kecerdasan buatan dalam pembelajaran Bahasa Inggris.

INTRODUCTION

Reading is widely recognized as a fundamental skill in language learning, particularly in the context of English as a Foreign Language (EFL). It plays a crucial role in developing learners' vocabulary knowledge, language awareness, critical thinking, and access to academic information. According to (Grabe & Stoller, 2011), reading is not merely decoding written text but involves complex cognitive processes such as interpreting meaning, making inferences, and integrating prior knowledge. Similarly, (Snow, 2002) states that reading comprehension is the process of simultaneously extracting and constructing meaning through interaction with written language. These perspectives make clear that effective reading instruction must be supported by appropriate assessment practices to measure

students' comprehension accurately. However, many educational contexts still rely on traditional assessment methods that emphasize surface-level understanding. Therefore, how can reading assessment be designed to effectively measure students' comprehension skills?

The four indicators of reading comprehension used in this study are drawn from two main theories: Barrett's Taxonomy of Cognitive and Affective Dimensions of Reading Comprehension (Barrett, 1968) and Bloom's Taxonomy of Educational Objectives (Bloom et al., 1956), as later revised by Anderson & Krathwohl (2001). Both theories serve as the theoretical foundation for constructing a cognitive hierarchy of reading abilities, ranging from the most basic level to higher-order thinking skills. The first indicator is literal comprehension, derived from Barrett (1968). At this level, students are required to recognize and recall information that is explicitly stated in the text. This represents the most fundamental level, as students only need to locate answers that are directly written without any need for further interpretation. The second indicator is vocabulary in context, which refers to students' ability to interpret word meanings based on contextual clues found within the text. Nation (2009) emphasizes that this skill is essential in reading comprehension, as insufficient vocabulary knowledge can lead to misunderstandings and inaccurate assumptions about the content of a text. The third indicator is inferential comprehension, drawn from Jufri (2015) with reference to Barrett (1968). Unlike literal comprehension, this level requires students to draw conclusions from information that is implicitly conveyed in the text. To do so, students must utilize contextual clues as well as their prior background knowledge. The fourth and highest indicator is critical thinking, derived from Bloom's Revised Taxonomy by Anderson & Krathwohl (2001). This level falls under the category of Higher Order Thinking Skills (HOTS), as it requires students to evaluate and analyze textual information in depth in order to produce independent and critical judgments or conclusions. Taken together, these four indicators form a graduated cognitive hierarchy progressing from lower-order to higher-order thinking abilities within the context of English reading assessment.

In classroom practice, reading assessment often focuses on students' ability to answer multiple-choice questions or recall explicit information from texts. While such methods are efficient, they may not fully capture deeper levels of comprehension such as critical thinking, interpretation, and evaluation (Alderson, 2000). Furthermore, (Brown, 2004) argues that assessment should align with learning objectives and provide meaningful feedback that supports students' development. In many cases, there is a mismatch between instructional goals and assessment techniques, which can lead to inaccurate representations of students' reading abilities. This issue highlights the importance of developing more comprehensive and valid reading assessment strategies. Thus, what types of reading assessments can better reflect students' higher-order comprehension skills?

Recent development in Artificial Intelligence (AI) has created new opportunities for educational assessment. AI-powered tools are increasingly being used to support teaching, learning, and evaluating processes by providing automated feedback, analyzing learner performance, and assisting educators in decision-making. In language education, AI technologies have demonstrated potential for improving assessment efficiency while providing detailed feedback to learners. These developments have encouraged researchers and educators to explore how AI can be integrated into classroom assessment practices in meaningful and responsible ways.

One AI tool that has gained attention in educational settings is Claude AI, a large language model developed by Anthropic. Claude AI is capable of analyzing written responses, generating feedback, identifying strengths and weaknesses in learner performance, and applying scoring criteria consistently across multiple responses. Such capabilities suggest that AI may assist teachers in conducting formative assessment more efficiently. Nevertheless, questions remain regarding the extent to which AI-generated assessment accurately reflects students' comprehension abilities and whether AI can adequately evaluate higher-order thinking skills that require interpretation, reasoning, and critical judgement.

Several studies have examined the broader application of AI in language assessment and education. Zawacki-Richter, Marín, Bond, and Gouverneur (2019)

reviewed AI applications in higher education, identifying key uses such as profiling and prediction, intelligent tutoring, and assessment and evaluation. In a related context, research has also explored how AI-assisted tools support specific language skills, such as translation, showing that learners can benefit from AI feedback in developing accuracy and confidence (Almhasees et al., 2024). Other studies have focused on traditional, human-based assessment practices, emphasizing the importance of clear rubrics, rater consistency, and alignment between test design and learning objectives (Bachman & Palmer, 1996; Brown, 2004). However, these studies primarily examine either general AI applications in education or conventional human-rated assessment methods, without specifically addressing how large language models evaluate the distinct cognitive demands of reading comprehension—from literal recall to higher-order critical thinking.

Although previous studies examined reading assessment practices and the use of educational technologies in language learning, research specifically investigating the use of Claude AI in reading comprehension assessment remains limited. Much of the existing literature focuses on traditional assessment methods or general AI applications without examining how AI performs when evaluating students' responses across different reading comprehension indicators. Consequently, there is a need for empirical evidence regarding the strengths, limitations, and practical implications for using Claude AI as a reading assessment tool in English Language Teaching.

Therefore, this study aims to examine students' reading comprehension performance across four indicators-literal comprehension, vocabulary in context, inferential reading, and critical thinking-whole also exploring the role of Claude AI in assessing students' responses. By analyzing the assessment outcomes generated through Claude Ai, this study seeks to contribute to outgoing discussions regarding the interrogation of artificial intelligence into language assessment and provide insights into its potential use in English Language Teaching.

Accordingly, this study addresses the following research questions:

How do Junior High School students perform across different reading comprehension indicators, including literal comprehension, vocabulary in context, inferential reading, and critical thinking by using Claude AI?

METHOD

Descriptive research was used in this study to analyze students' reading comprehension performances and investigate the application of Claude AI to measure their reading performance. Descriptive research is used when researchers intend to describe the phenomena that already exist, the behaviors occurring, or the conditions that prevail (Ary et al., 2010; Creswell, 2014). There were no manipulations of variables in this study as the researcher described the students' performance based on their responses to the reading comprehension test and the evaluation made by Claude AI.

The study analyzed and described students' performance in terms of the four reading comprehension competencies – literal comprehension, vocabulary in context, inferential reading, and critical thinking – as well as the feedback produced by Claude AI. The results were presented using the distribution of scores and percentages in order to give an adequate description of students' reading performance.

The subjects used in this study comprised 38 students from SMP Islam Bukittinggi who participated in a reading comprehension test using Google Forms. This school was selected because it provided accessible permission for classroom-based research and because its students, drawn from Grades 7 to 9, reflect a range of English proficiency levels typical of Junior High School learners in the local context, making it a suitable site for examining reading comprehension performance across multiple cognitive indicators. The sample of 38 students was considered adequate for the purposes of this descriptive study, as it included learners from different grade levels and proficiency backgrounds, thereby allowing the data to capture variation in literal, vocabulary, inferential, and critical thinking performance rather than reflecting a single, narrow ability range. Students' answers were further examined using Claude AI to identify how well students performed in

reading comprehension tests and how effective AI-based assessment is in testing students. Although the sample was not drawn through random probability sampling, its composition across grade levels and proficiency levels provided sufficient variation to meaningfully describe reading comprehension trends within this school population.

The tool employed in this study was a reading comprehension test that involved the use of Google Forms. It comprised a single reading passage titled *My Morning Routine*, which was followed by five essay questions aimed at testing the reading comprehension skills of the participants. The questions were prepared using four indicators of reading comprehension that were identified from the literature on English Language Teaching: literal comprehension, vocabulary in context, inferential reading, and critical thinking (Alderson, 2000; Brown, 2004). The first two questions were created to measure students' literal comprehension of the text, the third one checked the vocabulary knowledge in context, the fourth one estimated the skills of inferential reading, and the last one assessed critical thinking. In order to confirm the appropriateness of the reading comprehension test, validation of the instrument took place in relation to the passage and the five questions in relation to their intended indicators. All the five questions were found to be of the highest level in terms of validation, indicating consistency between the questions and the indicator for each (literal comprehension, vocabulary in context, inferential reading, and critical thinking). The validation process revealed two concerns regarding typographical error in the passage title (from “*My Morning Routine*” to “*My Morning Routine*”) and clarification of Question 4 in relation to character inference from text. Thus, the instrument can be validated for use, although it needs to be noted that only one validator was used.

The answers provided by the respondents were then analyzed using Claude AI and the scoring rubric, allowing for the identification of students' strengths and weaknesses when it comes to the four mentioned above reading comprehension skills.

Data collection occurred via the reading comprehension test conducted online through Google Forms. The test was sent to 38 Junior High School students from SMPI Islam Bukittinggi. Before taking the test, students were asked to read the passage and provide responses to the five questions written in essay format.

After submission of responses, their responses were collated and input into Claude AI for evaluation. Responses were then scored by Claude AI based on the preset grading criteria, providing scores and feedback for each question. This data was then collected and organized.

Descriptive analysis was employed in the processing of data. Responses from the students were first graded using Claude AI with scoring rubrics applied to each reading comprehension indicator. The scores were thereafter collated into tables.

Description of the students' reading comprehension performance involved the computation of their average scores, percentage scores, and score distributions in respect of each indicator such as literal comprehension, vocabulary in context, inferential reading, and critical thinking. Data analysis findings were presented in tables and descriptive paragraphs.

In analyzing the responses from the students, the researcher did not only focus on their performance in relation to reading comprehension indicators but also analyzed the feedback from Claude AI. This analysis would assist in highlighting the strengths and limitations associated with the use of AI technology to assess reading comprehension.

FINDINGS AND DISCUSSION

This section presents the findings of the reading comprehension test administered to 38 Junior High School students, followed by a discussion of these results in relation to the four reading comprehension indicators assessed in this study: literal comprehension, vocabulary in context, inferential reading, and critical thinking. The findings are first presented in tabular form to summarize students' overall and indicator-specific performance, after which each indicator is discussed in turn, drawing on students' responses and the assessment generated through

Claude AI. This discussion also considers the extent to which Claude AI's evaluation aligns with established theories of reading comprehension, in order to address the role of AI-assisted assessment raised in the second research question.

Table 1. Students' Reading Comprehension Performance Across Indicators

Indicator	Question	Average Score	Percentage	Category
Literal Comprehension	Q1	1.87 / 2	94%	Excellent
Literal Comprehension	Q2	1.47 / 2	74%	Good
Vocabulary in Context	Q3	1.34 / 2	67%	Fair
Inferential Reading	Q4	2.00 / 3	67%	Fair
Critical Thinking	Q5	1.42 / 3	47%	Needs Improvement
Overall Performance	–	8.11 / 12	68%	Fair

a. Overview of Student Performance

The reading comprehension test was administered to 38 Junior High School Students (Grade 7-9) using a passage titled *My Morning Routine*. The test consisted of five essay questions covering four indicators: Literal Comprehension, Vocabulary in Context, Inferential Reading, and Critical Thinking. The maximum score was 12 points.

The class achieved an overall average score of 8.11 out of 12 (68%), which falls within the Fair category. The grade distribution was as follows: 3 students (8%) achieved Excellent, 14 students (37%) achieved Good, 15 students (39%) were categorised as Fair, and 6 students (16%) fell under Needs Improvement. This distribution shows that most students showed some degree of comprehension from basic up to moderate levels, indicating that there is a need for improvement especially when it comes to higher-order thinking skills. From this pattern of results, it can be deduced that present reading education in this case is more geared toward teaching basic comprehension skills like identifying explicit information and not enough toward developing analytical skills necessary for inferential reading. In other words, students know how to read properly but they lack proper training to go beyond the literal content of the texts. From the point of pedagogy, if most of the students (55%, adding together Fair and Needs Improvement categories) have problems going beyond moderate comprehension, then classroom reading exercises need to emphasize interpretive, judicious and critical activities

instead of memorization alone. The finding is in accordance with existing trends within research on English Language Teaching (ELT), showing that Junior High School level learners demonstrate greater confidence in performing lower-level comprehension tasks than evaluating or analyzing tasks (Grabe, 2009). The findings fit well within the framework of Barrett's taxonomy of cognitive and affective dimensions of reading comprehension. According to Barrett, literal comprehension can be regarded as the most basic level of reading proficiency, as it is developed at the early stage of cognitive development and learning. Inferential and evaluative comprehension skills emerge as learners develop greater cognitive abilities and get more experience in reading (Barrett, 1968). This can be related to the fact that as noted by Alderson (2000), classroom evaluation in many English Foreign Language teaching settings has traditionally focused on retrieving explicit information from the text.

b. Literal Comprehension is a Relative Strength

Literal comprehension was assessed through two questions: identifying the time Budi wakes up (Q1) and identifying what he does before leaving the house (Q2). Q1 yielded the highest-class average of 1.87/2 (94%), making it the strongest-performing indicator overall. Nearly all students were able to retrieve this explicitly stated information from the passage. Q2 scored an average of 1.47/2 (74%), with most students correctly identifying schedule-checking as the key pre-leaving action, though some listed general activities such as bathing and eating breakfast without identifying the most specific and relevant detail.

The strong performance in literal comprehension is consistent with established findings in reading research, which suggest that lower-order skills such as fact retrieval are more accessible to younger learners at early stages of language development (Grabe & Stoller, 2011). This is further supported by Barrett's Taxonomy of Cognitive and Affective Dimensions of Reading Comprehension, which identifies literal comprehension as the foundational level of reading skill, requiring learners only to recognize and recall information explicitly stated in a text rather than to manipulate or interpret it (Barrett, 1968). In addition, Bloom's

Taxonomy of Educational Objectives similarly classifies recall and recognition as lower-order cognitive processes that develop earlier and require less cognitive demand than analysis or evaluation, which helps explain why students across varying proficiency levels were able to perform this task successfully (Bloom et al., 1956). The students in the Junior High level had already learned how to find explicit information in texts, and the linear structure of the text further helped them to do so. Nonetheless, what is notable about the fall from Q1 to Q2 (from 94% to 74%) is the fact that while most students were able to answer correctly, some gave peripheral information, which indicates that selective attention, which refers to finding the most important information from a number of options, is still developing and needs to be taught. This shows that literal comprehension is not one fixed skill but exists on a spectrum depending on the complexity involved in the task. That is to say, finding one isolated piece of information (e.g., the time) is easier than finding the most important information among a series of information in a sentence or a passage. Thus, it can be deduced that there should be an explicit teaching of close reading strategies including re-reading the target parts of the text and comparing the various information to the question asked.

c. Vocabulary Knowledge Requires Deeper Contextual Engagement

Q3 asked students to infer the meaning of the word *punctual* from context, yielding a class average of 1.34/2 (67%). While many students correctly identified "on time" or "not late" as the meaning, a notable number gave imprecise synonyms such as "disciplined," "well-prepared," or "diligent" — words that are associated with punctuality but do not capture its precise definition. Several students answered in Indonesian (*tepat waktu, tidak terlambat*), which were accepted as correct, though this raises questions about students' ability to demonstrate vocabulary knowledge in the target language.

These responses reflect a common challenge in ELT: students often possess a general sense of a word's meaning without being able to articulate its precise definition. This aligns with Nation's (2001) framework of vocabulary depth, which distinguishes between knowing that a word exists, knowing its form-meaning

connection, and knowing how to use it accurately in context. The findings suggest that vocabulary instruction in this context may benefit from greater emphasis on contextual inference strategies — teaching students to systematically use surrounding text clues to derive word meanings — rather than relying solely on rote memorisation or approximate synonymy.

d. *Inferential Reading Reveals a Gap Between Recognition and Analysis*

Q4 required students to infer why Budi checks his schedule and what this reveals about his character, producing a class average of 2.00/3 (67%). Many students successfully identified the surface reason (to avoid forgetting homework) but did not fully develop the character inference, offering single-word traits like "discipline" without elaboration. A smaller group provided rich, well-reasoned answers that named multiple character traits and connected them to textual evidence, demonstrating strong inferential ability.

These bimodal patterns point to a key challenge in developing inferential reading at the Junior High level — the ability to move from *what the text says* to *what the text implies*. Many students could identify the surface reason but did not take the additional step of connecting that behaviour to a broader character analysis. This finding is consistent with Alderson's (2000) observation that inferential comprehension requires not only linguistic competence but also the cognitive habit of reading beyond the literal level — a habit that typically develops through guided practice, questioning strategies, and explicit instruction in how to read for implication rather than just information. It also fits with Jufri's (2015) account of inferential reading, which describes it as a step up from literal comprehension: learners need to draw on contextual clues and their own background knowledge to fill in what the text doesn't say outright. Bloom's Revised Taxonomy points to something similar — inferring and analysing sit higher up the thinking ladder than simply remembering or understanding, which means students naturally need more support and practice before they can do it confidently (Anderson & Krathwohl, 2001). Put together, these ideas help explain why some students in this study could spot the surface-level reason in the text but struggled to take that extra step toward

a fuller character interpretation: moving from literal to inferential reading isn't something that just happens on its own, it's a skill that needs to be deliberately taught and practiced.

e. Critical Thinking Remains the Most Underdeveloped Skills

Critical thinking was the weakest-performing indicator, with a class average of 1.42/3 (47%). Q5 asked students to state their position on the importance of morning routines and support it with at least two reasons. Many students gave very brief responses — some as short as "yes" or "yes, because it is important" — without developing their reasoning. Others listed reasons without connecting them to a clear, structured argument. Only a small number of students produced well-supported responses that demonstrated genuine critical engagement.

This result reflects a broader challenge in Indonesian ELT contexts, where reading instruction has historically emphasised comprehension of explicit content over the development of evaluative and argumentative skills (Cahyono & Widiati, 2006). Several factors may explain this gap: limited exposure to opinion-based writing tasks in English, low confidence in expressing personal views in the target language, and a classroom culture that does not consistently reward divergent or evaluative thinking. One standout response drew on the Indonesian cultural concept of *gotong royong* as a parallel to collective discipline, showing both creativity and cultural awareness — demonstrating that when students are given the space and confidence to draw on their own cultural knowledge, the quality of critical thinking improves markedly.

The findings that have been mentioned above answer the first research question, which is related to describing the level of students' performance in terms of their abilities to perform four types of reading comprehension tasks such as literal comprehension, vocabulary in context, inferential reading, and critical thinking. On the whole, it can be concluded that the level of performance in literal comprehension tasks was higher than in the case of vocabulary, inferential, and critical thinking. Students showed better results when performing tasks aimed at identifying explicitly written information. However, problems emerged in the case

of the interpretation, evaluation, and critical response to textual information. The use of Claude AI in assessing students' reading comprehension responses proved to be appropriate because reading comprehension assessment fundamentally involves the interpretation and evaluation of written language, which aligns with the core capabilities of Large Language Models (LLMs). Claude AI is trained on extensive linguistic data and is designed to recognize meanings, contextual relationships, logical connections, and argumentative structures within texts. These capabilities enable it to evaluate responses related to literal comprehension, vocabulary in context, inferential comprehension, and critical thinking.

For literal comprehension, Claude AI can identify whether students accurately retrieve explicit information presented in the text. For vocabulary in context, it can analyze whether students understand word meanings based on surrounding textual cues. For inferential comprehension, Claude AI is capable of examining the logical relationship between textual information and students' conclusions, allowing it to determine whether inferences are supported by evidence from the text. In assessing critical thinking, Claude AI can evaluate the relevance, coherence, and justification of students' judgments and opinions regarding the text.

Furthermore, the use of Claude AI offers a level of scoring consistency that may be difficult to achieve through manual assessment alone. Human raters can be influenced by fatigue, personal bias, or variations in interpretation, whereas Claude AI applies the same assessment criteria across all responses. As long as a clear scoring rubric is provided, the evaluation process remains systematic and standardized. Therefore, the findings suggest that Claude AI can function as a reliable assessment assistant for reading comprehension, particularly in studies involving large amounts of qualitative response data.

The results also support the growing body of research indicating that Large Language Models can perform educational assessment tasks that require language understanding and analytical reasoning. Since reading comprehension assessment is primarily a language-based cognitive task, the ability of Claude AI to process semantic meaning, contextual information, and textual relationships makes it a suitable tool for evaluating students' comprehension performance. However,

Claude AI should be viewed as a supportive assessment instrument rather than a replacement for human judgment. The final interpretation of students' performance should remain grounded in established reading comprehension theories and validated scoring rubrics.

The use of Claude AI in this study is justified because the assessment of reading comprehension requires the evaluator to understand language, interpret meaning, identify textual evidence, and analyze reasoning. These are precisely the functions that Large Language Models such as Claude AI are designed to perform. Since the four indicators assessed in this study—literal comprehension, vocabulary in context, inferential comprehension, and critical thinking—are language-based constructs, Claude AI possesses the necessary linguistic and analytical capabilities to evaluate students' responses systematically. Therefore, Claude AI is not merely a technological convenience but a theoretically relevant assessment tool whose core competencies directly correspond to the cognitive processes involved in reading comprehension assessment.

CONCLUSION

. The purpose of the current descriptive study is to investigate the performance of the students in reading comprehension using four reading parameters. Moreover, the application of Claude AI in assessing their reading comprehension skills in the English Language Teaching classroom has been investigated in this study. The results suggest that the students have shown good performance in literal comprehension, fairly performed in vocabulary and inferential reading, but showed poor performance in critical thinking with an average class performance of 68%.

There are some implications regarding the importance of this study in both teaching and assessments. Reading should not just focus on comprehension but also should include activities that involve inferential and evaluative skills. When dealing with vocabulary lessons, it is important not only to teach students the meanings of the words but also how they can be used contextually. The development of critical

thinking should include activities such as opinion writing, text response, discussion, and debate.

In terms of AI-powered assessment, it was found out that Claude AI can serve as an efficient formative assessment tool by utilizing the scoring rubric, providing instant feedback, and recognizing performance trends in reading comprehension skills. This means that Claude AI can help teachers recognize the performance level and areas for improvement more effectively. Nevertheless, there are some limitations that should also be mentioned. First of all, Claude AI cannot evaluate oral performance, classroom participation, and other skills requiring human observation. Furthermore, assessing critical thinking skills through open-ended questions may not be as accurate as when conducted by humans. Thus, Claude AI should be treated as an auxiliary tool rather than an alternative to teachers. Its main role is to support teachers in their work and provide them with additional feedback to enhance their efficiency.

When it comes to the effectiveness of Claude AI as an instrument for performing formative assessments with the help of AI, it is necessary to mention that Claude AI can prove to be effective in assessing students' skills thanks to the use of a scoring rubric, immediate feedback, and performance trends. This means that Claude AI will assist teachers in determining students' skills and performance better. Still, it should be noted that there are certain limitations of such an approach. For example, Claude AI will not be able to assess oral performance, classroom participation, and any skills which require human observation to be evaluated. Moreover, Claude AI will not be as precise as humans when assessing critical thinking skills in open-ended questions. Therefore, it should be treated only as an auxiliary tool.

REFERENCES

- Alderson, J. C. (2000). *Assessing reading*. Cambridge University Press.
- Anderson, L. W., & Krathwohl, D. R. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Longman.
- Almhasees, M., Al-Saqqar, A., & Omar, H. (2024). The impact of AI-assisted translation on English learners. *International Journal of Language and AI*, 8(1), 44–59.

- Ary, D., Jacobs, L. C., Sorensen, C., & Walker, D. (2010). *Introduction to research in education* (8th ed.). Wadsworth Cengage Learning.
- Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice*. Oxford University Press.
- Barrett, T. C. (1968). *Taxonomy of cognitive and affective dimensions of reading comprehension*. In T. Clymer (Ed.), *What is reading?* University of Chicago Press.
- Blin, F., & Munro, M. (2008). Why hasn't technology disrupted academics' teaching practices? *Computers & Education*, 51(4), 1530–1543.
- Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). *Taxonomy of educational objectives: The classification of educational goals. Handbook I: Cognitive domain*. Longman.
- Brown, H. D. (2004). *Language assessment: Principles and classroom practices*. Pearson Education.
- Cahyono, B. Y., & Widiati, U. (2006). The teaching of EFL reading in the Indonesian context. *TEFLIN Journal*, 17(1), 36–58.
- Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). SAGE Publications.
- Grabe, W., & Stoller, F. L. (2011). *Teaching and researching reading* (2nd ed.). Routledge.
- Harmer, J. (2007). *How to teach English* (new ed.). Pearson Longman.
- Hudson, T. (2007). *Teaching second language reading*. Oxford University Press.
- Jufri. (2015). Reading strategies for comprehending literal, inferential, and critical meanings in reading. *Proceedings of ISELT FBS Universitas Negeri Padang*.
- Nation, I. S. P. (2009). *Teaching ESL/EFL reading and writing*. Routledge.
- O'Malley, J. M., & Pierce, L. V. (1996). *Authentic assessment for English language learners: Practical approaches for teachers*. Addison-Wesley.
- Snow, C. E. (2002). *Reading for understanding: Toward a research and development program in reading comprehension*. RAND Corporation.
- Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). *Systematic review of research on artificial intelligence applications in higher education*. *International Journal of Educational Technology in Higher Education*, 16(1).

APPENDIX

APPENDIX 1

READING COMPREHENSION TEST

Junior High School (Grade 7–9)

Subject: English Language | Time: 45 Minutes

Student Information

Name :

Class :

Date :

Passage

Read the following passage carefully.

My Morning Routine

Every morning, Budi wakes up at five o'clock. He starts his day by washing his face and brushing his teeth. After that, he helps his mother prepare breakfast in the kitchen. They usually eat rice, eggs, and vegetables together as a family.

After breakfast, Budi gets dressed in his school uniform. He packs his bag with his books and stationery. Before leaving the house, he always checks his schedule to make sure he does not forget any homework.

Budi walks to school with his neighbour, Rina. The school is about ten minutes away from their house. On the way, they talk about their lessons and sometimes stop to buy a drink from a small stall. Budi always arrives at school five minutes before the bell rings.

Budi believes that a good morning routine helps him stay organised and focused during the day. His teacher often praises him for being punctual and well-prepared.

Questions

1. **[Literal]** What time does Budi wake up every morning?
2. **[Literal]** What does Budi do before leaving the house?
3. **[Vocabulary]** The passage uses the word "punctual." Based on the context, what does "punctual" most likely mean?
4. **[Inferential]** Why do you think Budi checks his schedule before leaving the house? What does this tell us about his character?
5. **[Critical Thinking]** Do you think having a morning routine is important for students? Give at least two reasons to support your answer.

APPENDIX 2
Score Assessment By Claude.AI
READING COMPREHENSION TEST
Student Response Analysis Report
Junior High School | English Language
Total Respondents: 38 Students

CLASS SUMMARY

Class Average Score: 8.11 / 12 (68%)

Indicator	Question	Max Score	Class Average	Avg %
Literal (Q1 — Wake-up time)	Q1	2	1.87	94%
Literal (Q2 — Pre-leaving action)	Q2	2	1.47	74%
Vocabulary (Q3 — 'Punctual')	Q3	2	1.34	67%
Inferential (Q4 — Schedule check)	Q4	3	2.00	67%
Critical Thinking (Q5 — Morning routine)	Q5	3	1.42	47%
TOTAL		12	8.11	68%

Grade Distribution

Grade	Count	Percentage	Range
Excellent	3	8%	85–100%
Good	14	37%	70–84%
Fair	15	39%	55–69%
Needs Improvement	6	16%	0–54%

INDIVIDUAL STUDENT RESULTS

1. ZELDA ASYIVA — 5/12 (42%) **Needs Improvement**

Student	Q1	Q2	Q3	Q4	Q5	Total
ZELDA ASYIVA	2/2	2/2	1/2	0/3	0/3	5/12

Feedback:

- Q1: Correct — identified the exact wake-up time.
- Q2: Correct — identified schedule-checking/homework preparation.
- Q3: Partial — gave a related concept but missed the precise meaning of 'on time'.
- Q4: Incorrect — did not address the inferential question.
- Q5: Incorrect — did not answer the critical thinking question.

2. KHANZA AQILA PUTRI — 8/12 (67%) **Fair**

Student	Q1	Q2	Q3	Q4	Q5	Total
KHANZA AQILA PUTRI	2/2	2/2	0/2	2/3	2/3	8/12

Feedback:

- Q1: Correct — identified the exact wake-up time.
- Q2: Correct — identified schedule-checking/homework preparation.
- Q3: Incorrect — did not recognise 'punctual' as meaning 'on time'.
- Q4: Good — mentioned reason or character, but did not fully develop both aspects.
- Q5: Good — gave a position and at least one reason, but needed more development.

3. Agustia Ramadhani — 9/12 (75%) **Good**

Student	Q1	Q2	Q3	Q4	Q5	Total
Agustia Ramadhani	2/2	2/2	2/2	2/3	1/3	9/12

Feedback:

- Q1: Correct — identified the exact wake-up time.
- Q2: Correct — identified schedule-checking/homework preparation.
- Q3: Correct — accurately defined 'punctual' as being on time.
- Q4: Good — mentioned reason or character, but did not fully develop both aspects.
- Q5: Weak — minimal response, position unclear or reasons insufficient.

4. Qeyla zahrani ulzi — 3/12 (25%) **Needs Improvement**

Student	Q1	Q2	Q3	Q4	Q5	Total
Qeyla zahrani ulzi	2/2	1/2	0/2	0/3	0/3	3/12

Feedback:

- Q1: Correct — identified the exact wake-up time.
- Q2: Partial — mentioned some activities but missed the key action (checking schedule).
- Q3: Incorrect — did not recognise 'punctual' as meaning 'on time'.
- Q4: Incorrect — did not address the inferential question.
- Q5: Incorrect — did not answer the critical thinking question.

5. Hafizh sony al akbar — 8/12 (67%) **Fair**

Student	Q1	Q2	Q3	Q4	Q5	Total
Hafizh sony al akbar	1/2	0/2	2/2	2/3	3/3	8/12

Feedback:

- Q1: Partially correct — mentioned time but imprecise or added irrelevant detail.
- Q2: Incorrect — did not identify the correct pre-leaving action.
- Q3: Correct — accurately defined 'punctual' as being on time.
- Q4: Good — mentioned reason or character, but did not fully develop both aspects.
- Q5: Excellent — stated a clear position with well-developed reasons.

6. Aqila Faiqa — 9/12 (75%) **Good**

Student	Q1	Q2	Q3	Q4	Q5	Total
Aqila Faiqa	2/2	2/2	0/2	2/3	3/3	9/12

Feedback:

- Q1: Correct — identified the exact wake-up time.
- Q2: Correct — identified schedule-checking/homework preparation.
- Q3: Incorrect — did not recognise 'punctual' as meaning 'on time'.
- Q4: Good — mentioned reason or character, but did not fully develop both aspects.
- Q5: Excellent — stated a clear position with well-developed reasons.

7. Junior Yuliadri — 10/12 (83%) **Good**

Student	Q1	Q2	Q3	Q4	Q5	Total
Junior Yuliadri	2/2	2/2	2/2	3/3	1/3	10/12

Feedback:

- Q1: Correct — identified the exact wake-up time.
- Q2: Correct — identified schedule-checking/homework preparation.
- Q3: Correct — accurately defined 'punctual' as being on time.
- Q4: Excellent — explained both the reason and character trait with good inferential reasoning.
- Q5: Weak — minimal response, position unclear or reasons insufficient.

8. Tianro pradipta alvaro — 7/12 (58%) **Fair**

Student	Q1	Q2	Q3	Q4	Q5	Total
Tianro pradipta alvaro	2/2	1/2	0/2	2/3	2/3	7/12

Feedback:

- Q1: Correct — identified the exact wake-up time.
- Q2: Partial — mentioned some activities but missed the key action (checking schedule).
- Q3: Incorrect — did not recognise 'punctual' as meaning 'on time'.
- Q4: Good — mentioned reason or character, but did not fully develop both aspects.
- Q5: Good — gave a position and at least one reason, but needed more development.

9. Kenzie Sinaru Mustaqim — 6/12 (50%) **Needs Improvement**

Student	Q1	Q2	Q3	Q4	Q5	Total
Kenzie Sinaru Mustaqim	2/2	1/2	0/2	2/3	1/3	6/12

Feedback:

- Q1: Correct — identified the exact wake-up time.
- Q2: Partial — mentioned some activities but missed the key action (checking schedule).

- Q3: Incorrect — did not recognise 'punctual' as meaning 'on time'.
- Q4: Good — mentioned reason or character, but did not fully develop both aspects.
- Q5: Weak — minimal response, position unclear or reasons insufficient.

10. Asha Rafeliko — 9/12 (75%) Good

Student	Q1	Q2	Q3	Q4	Q5	Total
Asha Rafeliko	2/2	2/2	2/2	2/3	1/3	9/12

Feedback:

- Q1: Correct — identified the exact wake-up time.
- Q2: Correct — identified schedule-checking/homework preparation.
- Q3: Correct — accurately defined 'punctual' as being on time.
- Q4: Good — mentioned reason or character, but did not fully develop both aspects.
- Q5: Weak — minimal response, position unclear or reasons insufficient.

11. Alvaro fabio rizky — 7/12 (58%) Fair

Student	Q1	Q2	Q3	Q4	Q5	Total
Alvaro fabio rizky	2/2	1/2	0/2	2/3	2/3	7/12

Feedback:

- Q1: Correct — identified the exact wake-up time.
- Q2: Partial — mentioned some activities but missed the key action (checking schedule).
- Q3: Incorrect — did not recognise 'punctual' as meaning 'on time'.
- Q4: Good — mentioned reason or character, but did not fully develop both aspects.
- Q5: Good — gave a position and at least one reason, but needed more development.

12. alfiano Aprilia RS Fatmawati — 7/12 (58%) Fair

Student	Q1	Q2	Q3	Q4	Q5	Total
alfiano Aprilia RS Fatmawati	2/2	1/2	2/2	1/3	1/3	7/12

Feedback:

- Q1: Correct — identified the exact wake-up time.
- Q2: Partial — mentioned some activities but missed the key action (checking schedule).
- Q3: Correct — accurately defined 'punctual' as being on time.
- Q4: Weak — gave a vague answer with minimal inference.
- Q5: Weak — minimal response, position unclear or reasons insufficient.

13. shaqilla Aisha valerie — 12/12 (100%) Excellent

Student	Q1	Q2	Q3	Q4	Q5	Total
shaqilla Aisha valerie	2/2	2/2	2/2	3/3	3/3	12/12

Feedback:

- Q1: Correct — identified the exact wake-up time.
- Q2: Correct — identified schedule-checking/homework preparation.
- Q3: Correct — accurately defined 'punctual' as being on time.
- Q4: Excellent — explained both the reason and character trait with good inferential reasoning.
- Q5: Excellent — stated a clear position with well-developed reasons.

14. Aidah Saluun Sasmita — 7/12 (58%) Fair

Student	Q1	Q2	Q3	Q4	Q5	Total
Aidah Saluun Sasmita	2/2	2/2	0/2	2/3	1/3	7/12

Feedback:

- Q1: Correct — identified the exact wake-up time.
- Q2: Correct — identified schedule-checking/homework preparation.
- Q3: Incorrect — did not recognise 'punctual' as meaning 'on time'.
- Q4: Good — mentioned reason or character, but did not fully develop both aspects.
- Q5: Weak — minimal response, position unclear or reasons insufficient.

15. FELISA ASHIL FAUZH — 7/12 (58%) Fair

Student	Q1	Q2	Q3	Q4	Q5	Total
FELISA ASHIL FAUZH	2/2	2/2	1/2	2/3	0/3	7/12

Feedback:

- Q1: Correct — identified the exact wake-up time.
- Q2: Correct — identified schedule-checking/homework preparation.
- Q3: Partial — gave a related concept but missed the precise meaning of 'on time'.
- Q4: Good — mentioned reason or character, but did not fully develop both aspects.
- Q5: Incorrect — did not answer the critical thinking question.

16. raissa sadiya — 10/12 (83%) Good

Student	Q1	Q2	Q3	Q4	Q5	Total
raissa sadiya	2/2	2/2	2/2	2/3	2/3	10/12

Feedback:

- Q1: Correct — identified the exact wake-up time.
- Q2: Correct — identified schedule-checking/homework preparation.
- Q3: Correct — accurately defined 'punctual' as being on time.
- Q4: Good — mentioned reason or character, but did not fully develop both aspects.
- Q5: Good — gave a position and at least one reason, but needed more development.

17. VANIA RAISSA PUTRI — 10/12 (83%) Good

Student	Q1	Q2	Q3	Q4	Q5	Total
VANIA RAISSA PUTRI	2/2	2/2	2/2	3/3	1/3	10/12

Feedback:

- Q1: Correct — identified the exact wake-up time.
- Q2: Correct — identified schedule-checking/homework preparation.
- Q3: Correct — accurately defined 'punctual' as being on time.
- Q4: Excellent — explained both the reason and character trait with good inferential reasoning.
- Q5: Weak — minimal response, position unclear or reasons insufficient.

18. Alfiqah almahira — 8/12 (67%) Fair

Student	Q1	Q2	Q3	Q4	Q5	Total
Alfiqah almahira	2/2	1/2	1/2	2/3	2/3	8/12

Feedback:

- Q1: Correct — identified the exact wake-up time.
- Q2: Partial — mentioned some activities but missed the key action (checking schedule).
- Q3: Partial — gave a related concept but missed the precise meaning of 'on time'.
- Q4: Good — mentioned reason or character, but did not fully develop both aspects.
- Q5: Good — gave a position and at least one reason, but needed more development.

19. HANYFA MIFTAHUL JANNAH — 9/12 (75%) Good

Student	Q1	Q2	Q3	Q4	Q5	Total
HANYFA MIFTAHUL JANNAH	2/2	2/2	2/2	2/3	1/3	9/12

Feedback:

- Q1: Correct — identified the exact wake-up time.
- Q2: Correct — identified schedule-checking/homework preparation.
- Q3: Correct — accurately defined 'punctual' as being on time.
- Q4: Good — mentioned reason or character, but did not fully develop both aspects.
- Q5: Weak — minimal response, position unclear or reasons insufficient.

20. BAGAS AIDAN SYAM — 10/12 (83%) Good

Student	Q1	Q2	Q3	Q4	Q5	Total
BAGAS AIDAN SYAM	2/2	1/2	2/2	2/3	3/3	10/12

Feedback:

- Q1: Correct — identified the exact wake-up time.
- Q2: Partial — mentioned some activities but missed the key action (checking schedule).
- Q3: Correct — accurately defined 'punctual' as being on time.
- Q4: Good — mentioned reason or character, but did not fully develop both aspects.
- Q5: Excellent — stated a clear position with well-developed reasons.

21. KHAIRINISA DWI ULVA — 5/12 (42%) Needs Improvement

Student	Q1	Q2	Q3	Q4	Q5	Total
KHAIRINISA DWI ULVA	0/2	0/2	0/2	2/3	3/3	5/12

Feedback:

- Q1: Incorrect — did not correctly state the wake-up time.
- Q2: Incorrect — did not identify the correct pre-leaving action.
- Q3: Incorrect — did not recognise 'punctual' as meaning 'on time'.
- Q4: Good — mentioned reason or character, but did not fully develop both aspects.
- Q5: Excellent — stated a clear position with well-developed reasons.

22. KANZA AZIZAH PUTRI — 10/12 (83%) Good

Student	Q1	Q2	Q3	Q4	Q5	Total
KANZA AZIZAH PUTRI	2/2	2/2	2/2	3/3	1/3	10/12

Feedback:

- Q1: Correct — identified the exact wake-up time.
- Q2: Correct — identified schedule-checking/homework preparation.
- Q3: Correct — accurately defined 'punctual' as being on time.
- Q4: Excellent — explained both the reason and character trait with good inferential reasoning.
- Q5: Weak — minimal response, position unclear or reasons insufficient.

23. Aisya Calista syainda — 8/12 (67%) Fair

Student	Q1	Q2	Q3	Q4	Q5	Total
Aisya Calista syainda	2/2	2/2	2/2	2/3	0/3	8/12

Feedback:

- Q1: Correct — identified the exact wake-up time.
- Q2: Correct — identified schedule-checking/homework preparation.
- Q3: Correct — accurately defined 'punctual' as being on time.

- Q4: Good — mentioned reason or character, but did not fully develop both aspects.
- Q5: Incorrect — did not answer the critical thinking question.

24. Mikayla Alysia Zetta — 11/12 (92%) Excellent

Student	Q1	Q2	Q3	Q4	Q5	Total
Mikayla Alysia Zetta	2/2	2/2	2/2	3/3	2/3	11/12

Feedback:

- Q1: Correct — identified the exact wake-up time.
- Q2: Correct — identified schedule-checking/homework preparation.
- Q3: Correct — accurately defined 'punctual' as being on time.
- Q4: Excellent — explained both the reason and character trait with good inferential reasoning.
- Q5: Good — gave a position and at least one reason, but needed more development.

25. KEVIN ANANDA — 9/12 (75%) Good

Student	Q1	Q2	Q3	Q4	Q5	Total
KEVIN ANANDA	2/2	1/2	2/2	2/3	2/3	9/12

Feedback:

- Q1: Correct — identified the exact wake-up time.
- Q2: Partial — mentioned some activities but missed the key action (checking schedule).
- Q3: Correct — accurately defined 'punctual' as being on time.
- Q4: Good — mentioned reason or character, but did not fully develop both aspects.
- Q5: Good — gave a position and at least one reason, but needed more development.

26. M.Ikbal subagri — 6/12 (50%) Needs Improvement

Student	Q1	Q2	Q3	Q4	Q5	Total
M.Ikbal subagri	0/2	0/2	2/2	3/3	1/3	6/12

Feedback:

- Q1: Incorrect — did not correctly state the wake-up time.
- Q2: Incorrect — did not identify the correct pre-leaving action.
- Q3: Correct — accurately defined 'punctual' as being on time.
- Q4: Excellent — explained both the reason and character trait with good inferential reasoning.
- Q5: Weak — minimal response, position unclear or reasons insufficient.

27. FADLURRAHMAN IKRAM — 9/12 (75%) Good

Student	Q1	Q2	Q3	Q4	Q5	Total
FADLURRAHMAN IKRAM	2/2	1/2	2/2	3/3	1/3	9/12

Feedback:

- Q1: Correct — identified the exact wake-up time.
- Q2: Partial — mentioned some activities but missed the key action (checking schedule).
- Q3: Correct — accurately defined 'punctual' as being on time.
- Q4: Excellent — explained both the reason and character trait with good inferential reasoning.
- Q5: Weak — minimal response, position unclear or reasons insufficient.

28. Shakila cantika — 10/12 (83%) Good

Student	Q1	Q2	Q3	Q4	Q5	Total
Shakila cantika	2/2	2/2	2/2	2/3	2/3	10/12

Feedback:

- Q1: Correct — identified the exact wake-up time.
- Q2: Correct — identified schedule-checking/homework preparation.
- Q3: Correct — accurately defined 'punctual' as being on time.
- Q4: Good — mentioned reason or character, but did not fully develop both aspects.
- Q5: Good — gave a position and at least one reason, but needed more development.

29. Aditya Zikri hamdani — 7/12 (58%) Fair

Student	Q1	Q2	Q3	Q4	Q5	Total
Aditya Zikri hamdani	2/2	2/2	0/2	2/3	1/3	7/12

Feedback:

- Q1: Correct — identified the exact wake-up time.
- Q2: Correct — identified schedule-checking/homework preparation.
- Q3: Incorrect — did not recognise 'punctual' as meaning 'on time'.
- Q4: Good — mentioned reason or character, but did not fully develop both aspects.
- Q5: Weak — minimal response, position unclear or reasons insufficient.

30. Trivani Novelianty — 8/12 (67%) Fair

Student	Q1	Q2	Q3	Q4	Q5	Total
Trivani Novelianty	2/2	2/2	2/2	2/3	0/3	8/12

Feedback:

- Q1: Correct — identified the exact wake-up time.
- Q2: Correct — identified schedule-checking/homework preparation.
- Q3: Correct — accurately defined 'punctual' as being on time.
- Q4: Good — mentioned reason or character, but did not fully develop both aspects.
- Q5: Incorrect — did not answer the critical thinking question.

31. Indah Lestari Putri — 12/12 (100%) Excellent

Student	Q1	Q2	Q3	Q4	Q5	Total
Indah Lestari Putri	2/2	2/2	2/2	3/3	3/3	12/12

Feedback:

- Q1: Correct — identified the exact wake-up time.
- Q2: Correct — identified schedule-checking/homework preparation.
- Q3: Correct — accurately defined 'punctual' as being on time.
- Q4: Excellent — explained both the reason and character trait with good inferential reasoning.
- Q5: Excellent — stated a clear position with well-developed reasons.

32. Afiqah Mei Syasya — 7/12 (58%) Fair

Student	Q1	Q2	Q3	Q4	Q5	Total
Afiqah Mei Syasya	2/2	2/2	1/2	1/3	1/3	7/12

Feedback:

- Q1: Correct — identified the exact wake-up time.
- Q2: Correct — identified schedule-checking/homework preparation.
- Q3: Partial — gave a related concept but missed the precise meaning of 'on time'.
- Q4: Weak — gave a vague answer with minimal inference.
- Q5: Weak — minimal response, position unclear or reasons insufficient.

33. Muhammad alfarel — 10/12 (83%) **Good**

Student	Q1	Q2	Q3	Q4	Q5	Total
Muhammad alfarel	2/2	1/2	2/2	3/3	2/3	10/12

Feedback:

- Q1: Correct — identified the exact wake-up time.
- Q2: Partial — mentioned some activities but missed the key action (checking schedule).
- Q3: Correct — accurately defined 'punctual' as being on time.
- Q4: Excellent — explained both the reason and character trait with good inferential reasoning.
- Q5: Good — gave a position and at least one reason, but needed more development.

34. GIBRAN IBRAHIM — 9/12 (75%) **Good**

Student	Q1	Q2	Q3	Q4	Q5	Total
GIBRAN IBRAHIM	2/2	2/2	2/2	2/3	1/3	9/12

Feedback:

- Q1: Correct — identified the exact wake-up time.
- Q2: Correct — identified schedule-checking/homework preparation.
- Q3: Correct — accurately defined 'punctual' as being on time.
- Q4: Good — mentioned reason or character, but did not fully develop both aspects.
- Q5: Weak — minimal response, position unclear or reasons insufficient.

35. MUHAMMAD RASYID AL FAHRI — 7/12 (58%) **Fair**

Student	Q1	Q2	Q3	Q4	Q5	Total
MUHAMMAD RASYID AL FAHRI	2/2	2/2	0/2	2/3	1/3	7/12

Feedback:

- Q1: Correct — identified the exact wake-up time.
- Q2: Correct — identified schedule-checking/homework preparation.
- Q3: Incorrect — did not recognise 'punctual' as meaning 'on time'.
- Q4: Good — mentioned reason or character, but did not fully develop both aspects.
- Q5: Weak — minimal response, position unclear or reasons insufficient.

36. Cheryl Olivia — 8/12 (67%) **Fair**

Student	Q1	Q2	Q3	Q4	Q5	Total
Cheryl Olivia	2/2	1/2	2/2	2/3	1/3	8/12

Feedback:

- Q1: Correct — identified the exact wake-up time.
- Q2: Partial — mentioned some activities but missed the key action (checking schedule).
- Q3: Correct — accurately defined 'punctual' as being on time.
- Q4: Good — mentioned reason or character, but did not fully develop both aspects.
- Q5: Weak — minimal response, position unclear or reasons insufficient.

37. Azka dzikra yrinza — 3/12 (25%) **Needs Improvement**

Student	Q1	Q2	Q3	Q4	Q5	Total
Azka dzikra yrinza	2/2	0/2	1/2	0/3	0/3	3/12

Feedback:

- Q1: Correct — identified the exact wake-up time.
- Q2: Incorrect — did not identify the correct pre-leaving action.
- Q3: Partial — gave a related concept but missed the precise meaning of 'on time'.

- Q4: Incorrect — did not address the inferential question.
- Q5: Incorrect — did not answer the critical thinking question.

38. MAURA RIZKY DWI PUTRI — 8/12 (67%) **Fair**

Student	Q1	Q2	Q3	Q4	Q5	Total
MAURA RIZKY DWI PUTRI	2/2	1/2	2/2	1/3	2/3	8/12

Feedback:

- Q1: Correct — identified the exact wake-up time.
- Q2: Partial — mentioned some activities but missed the key action (checking schedule).
- Q3: Correct — accurately defined 'punctual' as being on time.
- Q4: Weak — gave a vague answer with minimal inference.
- Q5: Good — gave a position and at least one reason, but needed more development.