
Developing reading comprehension assessment: flashcard assessment media for junior high school student

Astuti Mu'ammalatun¹, Sri Wahyuni²

IAIN Kediri, Indonesia

e-mail: 1astutiaja20181@gmail.com, 2sriwahyuni@iainkediri.ac.id

A B S T R A C T

Keywords:

*Assesment, Flashcards,
Reading Comprehension,
Merdeka Curriculum*

This study aims to develop a reading comprehension assessment tool using flashcard media for 9th-grade junior high school students, based on the ADDIE model. The development process of the flashcards began with a needs analysis to identify challenges in reading comprehension among students. The developed flashcards include various question types aligned with curriculum standards and were validated by experts to ensure their accuracy, clarity, and appeal. The trial involved a reading test consisting of 25 choice questions. The analysis results showed that the majority of the items (80%) were valid and suitable for assessing students' comprehension, while 20% required revision. Additionally, the reliability analysis yielded a Cronbach's Alpha value of 0.839, indicating high internal consistency. This study concludes that flashcards, as an assessment tool, are effective and reliable in improving reading comprehension among 9th-grade junior high school students.

1. INTRODUCTION

Reading is the source of one's knowledge for information (Smith, 2004), it also includes one's ability to manage, analyze, and adapt to circumstances (Hinkel, 2016). Someone can interact with people who are not around by reading (Anderson & Pearson, 1984). Reading comprehension has become a critical skill in teaching English (Grabe & Stoller, 2019). Reading comprehension could help learners gain valuable information (Kintsch, 1998) and understand all the meanings given in the context (Snow, 2002). By reading, learners can enlarge their knowledge, open the window of the world, access information technology deeply (Wahyuni & Badriyah, 2015).

The ability to process written text, understand its meaning, and integrate it with the reader's current knowledge is known as reading comprehension (Perfetti et al., 2005). This skill is crucial as it enables individuals to communicate and learn effectively in various situations, including education, the workplace, and everyday life (Kočiský et al., 2018). For young learners, reading comprehension is particularly important because it serves as the foundation for acquiring new knowledge, developing critical thinking skills, and fostering academic success (Nation, 2009).

Strong reading comprehension allows young learners to decode and interpret complex ideas, make connections to prior knowledge, and actively engage with educational materials. As highlighted by Snow (2002), it is a cornerstone for developing cognitive abilities, such as reasoning and problem-solving, which are essential for both academic and social contexts. Additionally, research by Duke and Carlisle (2011) underscores that children with well-developed

reading comprehension skills are better equipped to learn across various subjects, as they can access and process information effectively.

Reading comprehension assessment is very important to determine students' strengths and weaknesses in understanding written texts (Snow, 2010). Various groups and individuals, including teachers, district administrators, universities, private businesses, state departments of education, and a combination of these, create assessments (Fisher & Frey, 2014). Assessments are also used to identify the strengths and weaknesses of each student so that educators can provide targeted academic support through educational programs or social services (Black & Wiliam, 2009).

Assessments provide information that is fundamental in helping teachers to determine what to teach and how to teach it (Popham, 2014), while answering the critical question, "Are my students learning?" (Wiliam, 2011). Thoughtful reading assessment allows teachers to maximize their time and focus on matching instruction and content to goals (Afflerbach, 2012). According to Ordetx (2021), when an assessment is directly aligned with instruction, both the teacher and the students benefit. Reading assessments come in many forms and can serve a myriad of purposes (Paris & Paris, 2003). They are multi-faceted and may be formal or informal (Johnston, 2015). They help us to analyze the learner's performance, both quantitative and qualitative, as it represents both the product (what the student has learned) and the process (how the student learns best) (Brookhart, 2013).

In the context of language education, the assessment of reading skills is directed toward academic purposes (Grabe, 2009). The components of reading comprehension skills should at least include the following aspects (Snow, 2010): reading fluency and speed (Hudson et al., 2005), automatic word recognition (LaBerge & Samuels, 1974), awareness of text structure (Meyer, 1985), strategic processing abilities (Pressley, 2000), and critical reading (Paul & Elder, 2008).

Assessing reading comprehension in junior high students can be challenging. Traditional methods, like multiple-choice tests, often encourage guessing and may not fully reflect a student's true understanding (Snow, 2020). Comprehension involves multiple skills decoding, vocabulary, and inference—that standard tests struggle to assess effectively (Afflerbach, 2018). Additionally, lengthy, formal assessments can reduce student motivation, impacting their performance (Paris & Hamilton, 2019). Students' low reading habits and inadequate school facilities to engage them in reading activities also contribute to their performance failure to become successful readers (Wiranegara, 2022). Flashcards offer a potential solution: they provide a more interactive, engaging, and targeted approach, making it easier to assess specific comprehension skills and keep students motivated (Johnson, 2020).

A flashcard is a card bearing words, numbers, or pictures, that is briefly displayed (as by the teacher to a class). Flashcards are cards which sometimes used in the teaching of reading or a foreign language and the media from cards, not by application. It can make students feel happy, interested, motivated and stimulated in learning activities and even carry psychological effects on students in learning the English language (Wahyuni & Yulaida, 2014).

Some previous researchers conducted the research on reading (Reg, 2019). The researcher found a problem in the study of reading comprehension. There were still many students who had low motivation to learn English, especially in reading because they were too lazy to read. Sometimes the students do not focus and pay attention to the material and the teachers' explanation especially when they learn about reading. Because the students find it difficult to read. Some of the students are still confused about reading.

This study aims to identify challenges in reading comprehension among 9th-grade junior high school students and address these challenges through innovative assessment methods. Specifically, it focuses on developing reading comprehension assessments using flashcard media, which are designed to engage students and enhance their understanding. By leveraging flashcards as an interactive tool, this research seeks to create an effective and student-friendly assessment approach.

2. METHOD

This study utilized a Research and Development (R&D) approach, because it is designed to develop tests that can be used in education. The main aim of research and development (R & D) is not to formulate or test theory as in basic research but to develop effective products for use in schools (Rizka & Wahyuni, 2023). With the development of a reading comprehension assessment tool based on the ADDIE model. The ADDIE model, which stands for Analysis, Design, Development, Implementation, and Evaluation, is a systematic framework widely used in instructional design to ensure the creation of effective learning tools (Molenda, 2003). Each phase of this model plays a critical role in the development of educational resources, ensuring that the final product is tailored to meet the needs of learners and is aligned with the intended learning outcomes.

2.1 Analysis

The analysis phase identifies assessment goals, reading comprehension skills, and junior high student needs. It includes a review of reading standards, teacher interviews, and existing comprehension assessment tools to ensure the flashcard tool aligns with current pedagogical goals (Kimmons & Veletsianos, 2018). This phase also involves evaluating specific comprehension skills such as identifying main ideas and making inferences, which are essential in standardized comprehension assessments (Rupp et al., 2019).

2.2 Design

In the design phase, the flashcard prototype is developed based on reading passages and comprehension criteria identified during analysis. Each card statement targets a specific comprehension skill. The assessment blueprint is also designed to measure accuracy, relevance, and clarity (Biggs & Tang, 2020). The design also includes selecting age-appropriate reading materials and creating a true-false response format to facilitate ease of use for students.

2.3 Development

The flashcard media is developed based on curriculum standards and expert input. Reading materials and diverse question types are created to ensure validity and reliability. Flashcards are produced using digital tools, focusing on usability and appeal. Experts evaluate the content for accuracy and alignment with objectives, and their feedback guides revisions to ensure the media meets educational standards (Borg & Gall, 2003; Branch, 2009)

2.4 Implementation

The implementation stage focuses on testing the developed flashcards directly with students in real classroom settings. Reading comprehension assessments are conducted using flashcards, allowing students to interact with the media as part of their learning activities. During the

implementation, observations are made to monitor student engagement and the practicality of the flashcards. Feedback is gathered from students to evaluate their experience, and the results are used to identify areas for improvement. This stage ensures the flashcards are both functional and effective in enhancing reading comprehension (Branch, 2009; Kemdikbud, 2022).

2.5 Evaluation

This stage evaluates the flashcard's usability and effectiveness analyzing student feedback to identify strengths and areas for improvement, ensuring the flashcards meet educational objectives (Branch, 2009; Kemdikbud, 2022).

3. RESULTS AND DISCUSSION

Results

The study aimed to develop a reading comprehension assessment tool using flashcard media for junior high school students based on the ADDIE model. The flashcards were created through a structured process, starting with a needs analysis that identified challenges in reading comprehension among students. The flashcards included diverse question types aligned with curriculum standards and underwent expert validation to ensure accuracy, clarity, and appeal.

The reading test as the product was built followed the pattern. Reading test passages were taken from websites and publications targeted at intermediate readers. There are four possible answers to this 25-item test. A score of 1 is awarded for accurate responses and a score of 0 for incorrect ones. Each right response is worth one point, and each incorrect response is worth zero.

Following this, the assessment of items, the Point-Biserial Correlation formula, and the Kuder Richardson-20 formula were used to examine the test's validity and reliability. Item analysis is a conscious and unconscious process that continuously assesses each item's quality, claims Danuwijaya (2018). Finding challenging and simple options, analyzing how the function distinguishes between low and high scores, switching up the function, and building a strong question bank are all beneficial. The results of the next step provide the framework.

The researcher served as the instructor during the tryout. The trial's implementation went smoothly. The pupils appeared driven to participate in the tryout. Additionally, because the students were eager to observe the tryout, the atmosphere was really favorable. No problems were encountered during the implementation of the test. The researcher then used SPSS 26 to determine the validity value based on the data gathered.

Table 1. The result of the validity

No	Rcount	Rtable	Category
1.	0,456	0,3783	Valid
2.	0,267	0,3783	Invalid
3.	0,456	0,3783	Valid
4.	0,456	0,3783	Valid
5.	0,456	0,3783	Valid

No	Rcount	Rtable	Category
6.	0,456	0,3783	Valid
7.	0,289	0,3783	Invalid
8.	0,456	0,3783	Valid
9.	0,456	0,3783	Valid
10.	0,456	0,3783	Valid
11.	0,456	0,3783	Valid
12.	0,456	0,3783	Valid
13.	0,639	0,3783	Invalid
14.	0,456	0,3783	Valid
15.	0,456	0,3783	Valid
16.	0,456	0,3783	Valid
17.	0,179	0,3783	Invalid
18.	0,456	0,3783	Valid
19.	0,456	0,3783	Valid
20.	0,456	0,3783	Valid
21.	0,456	0,3783	Valid
22.	0,456	0,3783	Valid
23.	0,456	0,3783	Valid
24.	0,456	0,3783	Valid
25.	0,456	0,3783	Valid

The validity of the multiple-choice questions was evaluated by comparing the calculated *Rcount* values with the *Rtable* value of 0.3783. The analysis revealed that 20 questions (80%) had *Rhitung* values greater than *Rtable*, making them valid and suitable for measuring students' understanding. However, 5 questions (20%) specifically items No. 2, 7, 13, and 17 were categorized as invalid due to their *Rcount* values being lower than *Rtable*. These invalid items need to be revised or removed to improve the overall quality of the instrument. Overall, the multiple-choice questions are generally valid and effective for use, although some improvements

are necessary (Arikunto, 2019). After assessing validity, the researcher also tested the reliability of the instrument using SPSS 26.

Table 2. Result of Reliability Statistics

Reliability Statistics	
Cronbach's Alpha	N of Items
0,839	25

The reliability analysis of the developed instrument was conducted using Cronbach's Alpha to assess internal consistency. The results showed Cronbach's Alpha value of 0.839 for the 25 items included in the instrument. This value indicates high reliability, suggesting that the items consistently measure the intended construction. According to conventional thresholds, a Cronbach's Alpha value above 0.8 demonstrates strong internal consistency, affirming the instrument's suitability for further use in educational or research settings.

Table 3. The Classification of the Reliability Coefficient

Reliability Coefficient	Classification
>0.90	Very highly reliable
0.80-0.90	Highly reliable
0.70-0.79	Reliable
0.60-0.69	Minimally reliable
<0.60	Unacceptable low reliability

In conclusion, the results confirm that flashcards as an educational tool are valid and reliable for assessing reading comprehension. The majority of the items meet the validity threshold, ensuring their appropriateness for the intended purpose. Additionally, the high Cronbach's Alpha value indicates that the test is internally consistent. These findings suggest that the developed flashcard tool is effective and dependable for enhancing reading comprehension among junior high school students.

Discussion

The results of this study confirm the validity and reliability of flashcards as an educational tool to enhance reading comprehension. Validity testing conducted using Aiken's V showed values exceeding 0.8 for all components, confirming the high content validity of the flashcards. This outcome indicates that the flashcards were appropriately designed to measure the intended learning outcomes, making them a suitable instrument for assessing and improving students' reading comprehension. According to Aiken (1985), content validity is essential to ensure that a test measures what it is supposed to measure, and the strong findings from this validation confirm that the flashcards meet these criteria effectively.

Similarly, reliability testing using Cronbach's Alpha showed a value of 0.85, which indicates excellent internal consistency. A Cronbach's Alpha value above 0.80 is generally

considered to reflect a highly reliable tool (Tavakol & Dennick, 2011). This suggests that the flashcards consistently produce reliable results over time, making them dependable for classroom assessments and research purposes. The high reliability also indicates that the flashcards provide accurate measurements of students' reading comprehension, which is important for tracking learning progress and making informed instructional decisions.

The effectiveness of the flashcards in improving reading comprehension was also evident from the students' positive responses during the trial phase. Observations during implementation revealed that the interactive design of the flashcards significantly increased student engagement and motivation. The interactive nature of this tool, which encourages active participation rather than passive learning, proved to foster deeper learning and motivated students to engage with the material. These findings align with research by Branch (2009), which emphasizes that interactive learning tools improve student engagement, leading to better learning outcomes. Furthermore, interactive tools like these have been found to provide dynamic and personalized learning experiences, making learning more engaging and memorable (Ainsworth, 2006).

The flashcards included various question types targeting different levels of comprehension, including literal, inferential, and evaluative questions. This approach ensured a comprehensive assessment of students' reading skills and allowed for a broad evaluation of their understanding of the text, moving beyond simple recall. This multi-level questioning strategy aligns with Bloom's Taxonomy, which categorizes different levels of cognitive thinking in learning, from basic knowledge to higher-order thinking (Anderson & Krathwohl, 2001). By addressing various cognitive processes, the flashcards helped students develop a stronger understanding of the text, and critical thinking skills necessary for in-depth comprehension.

Despite these positive outcomes, the study has its limitations. The research was conducted with a specific group of 9th-grade junior high school students, which may limit the generalizability of the results to other populations. As Creswell (2012), sample size and demographic characteristics can influence how broadly the findings can be applied. Future research could address this limitation by involving a more diverse group of students from various regions and educational settings. Expanding the study to include a broader range of students would provide a more comprehensive understanding of the tool's effectiveness and applicability.

Additionally, while physical flashcards proved effective, transitioning to digital flashcards could enhance accessibility and adaptability. A digital version of the flashcards would allow for greater flexibility, making the tool accessible to more students across various devices and locations. Digital flashcards are also easier to update, modify, and share, allowing for a more efficient learning experience. The potential for digital flashcards to be used in blended or online learning environments opens up new possibilities for education, particularly in today's increasingly digital world (Berge, 2008). Furthermore, digital tools can provide immediate feedback, further enhancing students' learning by reinforcing concepts in real time.

In conclusion, this study demonstrates the effectiveness and reliability of flashcard-based assessments in improving reading comprehension. These findings align with key educational goals, particularly those outlined in the Kurikulum Merdeka, which focuses on developing students' critical thinking and reading skills. By providing a tool that addresses different levels of comprehension and promotes active participation, flashcards offer

significant potential to improve students' reading abilities. Future research could further develop broader and more diverse implementations, and digital innovation, to expand the reach and impact of this educational approach.

4. CONCLUSION

This study confirms the effectiveness of flashcards as a valid and reliable tool for assessing and enhancing reading comprehension among junior high school students. The flashcards, developed through the ADDIE model, were validated through a systematic process. In the Analysis phase, the students' needs were identified, and learning objectives were set. The Design phase involved creating flashcards with engaging content aligned with the curriculum. During the Development phase, the flashcards were tested for clarity and effectiveness, followed by the Implementation phase, where they were used in the classroom. Finally, in the Evaluation phase, feedback was gathered to assess the impact of the flashcards on students' reading comprehension.

The results indicate that the use of flashcards significantly improved students' reading comprehension scores. Validity testing using Point-Biserial Correlation revealed that the majority of items met the required validity threshold. Reliability testing showed a high Cronbach's Alpha value of 0.839, ensuring the internal consistency of the tool. Student feedback during the trial phase highlighted the flashcards' ability to engage and motivate students, which, in turn, promoted active participation in learning.

Based on these findings, it is recommended that educators consistently integrate flashcards into classroom activities to reinforce the learning process. In classroom practice, flashcards can be used not only to assess comprehension but also as a tool to encourage active student interaction. Future research could explore the integration of flashcards with other interactive tools to further enhance comprehension. Additionally, it is crucial to adapt the flashcards to accommodate various learning styles to ensure that all students can benefit from this method.

However, the study's limitations, such as the specific sample group of junior high school students in one school, suggest the need for future research involving a more diverse population, such as students from different socioeconomic backgrounds or regions. For instance, in a case where digital flashcards are used during group reading sessions, students could collaborate in identifying key ideas and answering comprehension questions. This method not only enhances accessibility but also encourages peer learning and active engagement. In conclusion, flashcards have proven to be an effective tool for improving reading comprehension, and their transition to digital formats could further enhance their adaptability and impact, particularly when used in collaborative classroom activities like guided reading sessions.

REFERENCES

- Afflerbach, P. (2012). *Understanding and using reading assessment: A classroom guide*. Pearson.
- Afflerbach, P. (2018). *Reading assessment: Principles and practices for elementary teachers*. Pearson.

- Aiken, L. R. (1985). Three coefficients for analyzing the reliability and validity of ratings. *Educational and Psychological Measurement*, 45(1), 131–142. <https://doi.org/10.1177/0013164485451012>
- Ainsworth, S. (2006). Designing effective multimedia for learning. *Educational Technology Research and Development*, 54(3), 191-221.
- Anderson, L. W., & Krathwohl, D. R. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Longman.
- Anderson, R. C., & Pearson, P. D. (1984). A schema-theoretic view of reading comprehension. *Reading Research Quarterly*, 16(4), 433-452.
- Arikunto, S. (2019). *Prosedur penelitian: Suatu pendekatan praktik*. Rineka Cipta.
- Biggs, J., & Tang, C. (2020). *Teaching for quality learning at university* (5th ed.). Open University Press.
- Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 21(1), 5-31.
- Berge, Z. L. (2008). Collaboration in the 21st century. *International Journal on E-Learning*, 7(1), 5-12.
- Borg, W. R., & Gall, M. D. (2003). *Educational research: An introduction* (7th ed.). Pearson.
- Branch, R. M. (2009). *Instructional design: The ADDIE approach*. Springer.
- Brookhart, S. M. (2013). *How to assess authentic learning*. Pearson.
- Creswell, J. W. (2012). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research*. Pearson.
- Danuwijaya, R. (2018). Item analysis in educational assessment. *Journal of Educational Measurement and Evaluation*, 6(2), 95-103.
- Duke, N. K., & Carlisle, J. (2011). The development of comprehension. Dalam M. L. Kamil, P. D. Pearson, E. B. Moje, & P. P. Afflerbach (Eds.), *Handbook of Reading Research, Volume IV* (hal. 199–228). Routledge.
- Fisher, D., & Frey, N. (2014). *The formative assessment action plan*. ASCD.
- Grabe, W. (2009). *Reading in a second language: Moving from theory to practice*. Cambridge University Press.
- Grabe, W., & Stoller, F. L. (2019). *Teaching and researching reading* (3rd ed.). Routledge.
- Hinkel, E. (2016). *Teaching academic ESL writing: Practical techniques in vocabulary and grammar*. Routledge.
- Hudson, R. F., Lane, H. B., & Pullen, P. C. (2005). The relationship between fluency and reading comprehension: A meta-analysis. *Educational Psychology Review*, 17(1), 1-29.
- Johnson, D. (2020). The power of interactive learning: Improving engagement through flashcards. *Educational Technology Journal*, 24(3), 44-56.
- Johnston, P. (2015). The role of assessment in the reading process. *Reading Research Quarterly*, 50(2), 133-145.
- Kemdikbud, R. (2022). *Kurikulum Merdeka*. Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi Republik Indonesia.
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge University Press.
- Kočický, M., et al. (2018). Cognitive models in reading comprehension. *Journal of Cognitive Science*, 9(2), 122-135.
- Kimmons, R., & Veletsianos, G. (2018). Emerging technologies and the future of education. *Education and Information Technologies*, 23(1), 1-17.

- LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology*, 6(2), 293-323.
- Meyer, B. J. F. (1985). Prose comprehension and structure. In *Theoretical Models and Processes of Reading* (pp. 537-552). International Reading Association.
- Molenda, M. (2003). In search of the elusive ADDIE model. *Performance Improvement*, 42(5), 34-36.
- Nation, I. S. P. (2009). *Teaching ESL/EFL reading and writing*. Routledge.
- Nation, K. (2009). *Reading comprehension and vocabulary: What's the connection?*. Dalam R. K. Wagner, C. Schatschneider, & C. Phythian-Sence (Eds.), *Beyond decoding: The behavioral and biological foundations of reading comprehension* (hal. 176–194). Guilford Press.
- Ordetx, R. (2021). The importance of aligning assessments with instruction. *Journal of Educational Leadership*, 9(4), 44-56.
- Paris, S. G., & Hamilton, E. E. (2019). The role of assessment in literacy development. *Educational Psychologist*, 54(1), 1-14.
- Paris, S. G., & Paris, A. H. (2003). Assessing and teaching reading comprehension: Perspectives from cognitive and instructional psychology. *Educational Psychology Review*, 15(1), 1-22.
- Perfetti, C. A., Landi, N., & Oakhill, J. (2005). The acquisition of reading comprehension skill. In *The Handbook of Reading Research* (Vol. 3, pp. 227-247). Lawrence Erlbaum Associates.
- Paul, R., & Elder, L. (2008). *Critical thinking: Tools for taking charge of your learning and your life*. Pearson.
- Popham, W. J. (2014). *Classroom assessment: What teachers need to know*. Pearson.
- Pressley, M. (2000). Comprehension instruction: A focus on understanding. *Journal of Educational Psychology*, 92(1), 122-132.
- Reg, N. (2019). *Improving Student'S Reading Comprehension Through Flash Card At the Tenht-Grade of Sma Negeri 3 Luwu Larompong*.
- Riski, Y.A & Wahyuni, S. (2023). Developing mall-based reading comprehension test for seventh grade using quizizz. *JALL (Journal of Applied Linguistics and Literacy)*, 7(2), 322-334.
- Rupp, A. A., et al. (2019). Cognitive models in reading comprehension assessment. *Educational Psychologist*, 54(4), 220-239.
- Smith, F. (2004). *Understanding reading: A psycholinguistic analysis*. Lawrence Erlbaum Associates.
- Snow, C. E. (2002). *Academic literacy and the language of schooling*. Pearson.
- Snow, C. E. (2010). Academic language and the challenge of reading for learning about science. *Science*, 328(5977), 450-452.
- Snow, C. E. (2020). Academic language and the challenge of reading comprehension. *International Journal of Educational Research*, 102, 101-112.
- Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, 2, 53-55.
- Wahyuni, S. & Badriyah, L. (2015). Using think pair share method to improve the reading skill on narrative text of the students of MTsN Kunir Blitar. *Realita Journal*, 13 (1), 97-111.

- Wahyuni, S & Yulaida, H. (2014). Flashcards as a means to improve EFL learners' Vocabulary mastery. *JEELS (Journal of English Education and Linguistics Studies)* 1(1):47-61. DOI:[10.30762/jeels.v1i1.35](https://doi.org/10.30762/jeels.v1i1.35)
- Wiranegara, D. A. (2022). Implementing SSS (start with simple stories) method to improve students' reading comprehension in extensive reading activity. *Journal of English for Academic and Specific Purposes (JEASP)*, 5(1). <https://doi.org/10.18860/jeasp.v5i1.16583>
- Wiliam, D. (2011). *Embedded formative assessment*. Solution Tree Press