

Aligning Assessment and Curriculum: A Study of Summative Tests in Educational Settings

¹ Ella Aprilia Tiana, ²Nirwanto Maruf*

1,2 Universitas Muhammadiyah Gresik, Gresik, Indonesia*

*Corresponding Author

Email: nirwanto.maruf@umg.ac.id

Abstract

The study aims to explore the alignment between English summative tests and prescribed learning objectives within curriculum documents, highlighting areas of both conformity and inconformity. Employing meticulous analysis, the research identifies instances where test content aligns with specified learning objectives, while also pinpointing discrepancies or deviations. Through qualitative analysis, the study examines the degree of alignment between test content and prescribed objectives. The findings underscore numerous instances of alignment, particularly among test items correlating directly with prescribed topics and skills. Nevertheless, the research also uncovers areas of incongruity, indicating potential gaps in the assessment process. This study emphasizes the significance of rigorous test design and validation procedures to ensure that summative assessments accurately reflect intended curriculum standards and effectively measure student learning outcomes within the Merdeka curriculum context.

Keywords: *Content validity; Summative assessments; Merdeka curriculum*

INTRODUCTION

English holds a significant position as a foreign language in Indonesia, particularly within the realm of education. It is considered the primary foreign language that should be imparted to students at all levels of education in the country, especially among younger learners in elementary school. Both governmental bodies and educational institutions are actively engaged in enhancing the teaching and learning processes associated with English education in Indonesia. The particular importance is the aspect of evaluation. Evaluation serves the purpose of determining whether the learning objectives established by teachers have been effectively met by the students. Consequently, teachers and test developers must possess relevant information and employ evaluations, often through the construction of tests, as assessment tools. Tests are designed to gauge students' achievements and the overall efficacy of teaching and learning programs (Ali & Iqbal, 2013; de Vries et al., 2022; de Vries et al., 2023; Jarvis, 2015). Through such assessments, students can gauge their own knowledge and track their progress, while teachers can refine their instructional strategies on a daily basis based on the outcomes of these evaluations. Thus, assessments not only measure student learning but also inform and shape ongoing teaching practices (Fitriana et al., 2018; Boström & Palm, 2023).

To evaluate students' mastery of material covered over a semester, teachers, acting as test makers, must conduct assessments through testing. Assessment forms a crucial component of the curriculum, comprising a series of systematically and continuously conducted activities (Chen, 2021; Gupta, Shah, & Singh, 2021;

Luthfiyyahet al., 2022). Within the Merdeka curriculum framework, this assessment is termed as summative assessment, typically administered at the conclusion of a semester. The feedback gleaned from summative assessments serves as a metric for measuring student progress, thereby guiding teachers in devising and crafting activities for subsequent projects. Furthermore, the outcomes of these assessments are instrumental in enhancing the teaching and learning processes. As noted by Weeden, Winter, and Broadfoot (2002), along with Hughes (2003), summative tests represent the culmination or verification of what has been learned throughout the teaching and learning journey.

When crafting a robust test, teachers must prioritize its validity, which pertains to the extent to which the test accurately measures its intended objectives (Alderson, Clapham, & Wall, 1995). Validity encompasses two key aspects: external validity, comprising concurrent and predictive validity, and internal validity, which includes face, response, and content validity. A test is deemed valid when it satisfactorily addresses these dimensions. Among these, content validity holds particular significance. Content validity assesses whether the test items effectively capture and represent the content outlined for measurement (Heaton, 1998). Therefore, gauging the content validity of a test necessitates a thorough examination of the content standards outlined in the syllabus. Test developers should prioritize content validity during test construction, ensuring alignment with the curriculum and syllabus (Khodirin, 2013). Consequently, if a test accurately reflects the syllabus, it demonstrates strong content validity.

In fact, some teachers do not consider the aspect of validity in making summative test, especially content validity. The most important and the main concern of test used in classroom teaching is content validity. (Mary & Sake, 1983) stated that content validity is related to the material that students have learned. The test must include samples of the teaching materials given. To fulfil this, teachers should refer their considerations to the teaching syllabus, or we can say in Merdeka curriculum is ATP or flow of learning objectives. In addition, there are other reasons why the test does not consider the procedures in the guidebook (Sumardi, 2008 quoted in (Nofiyanti, 2011) First, the time given to design is adequate. Second, teacher's competence is still limited in designing good tests. Third, teachers may not have sufficient experience in designing and constructing English tests. In line with the previous study said that the English exam test material is not appropriate to the English syllabus recommended. However, in planning tests, some teachers are quite familiar with the content taken during teaching and its major aspects, but some tests constructed by teachers have low content validity, because it is not easy for teachers to compose good tests (Fitriyanti, 2014). Therefore, the test needs to be improved.

This research aims to investigate and delineate the areas of both conformity and inconformity existing between English summative tests and the prescribed learning objectives outlined in curriculum documents. Through meticulous analysis, the study seeks to identify instances where the content of the summative tests aligns with the specified learning objectives, as well as areas where discrepancies or deviations are evident.

METHOD

Participants

The participants in this qualitative research study consist of English teachers and fifth-grade students from multiple elementary schools. They are UPT SDN 51 Gresik, UPT SDN 31 Gresik and UPT SDN 16 Gresik. UPT SDN 51 Gresik and UPT SDN 31 Gresik, these schools that have implemented the Merdeka curriculum in Indonesia.

Instruments

The primary instruments utilized in this study are the English summative tests administered to fifth-grade students during the first semester. Additionally, the English flow of learning objectives, provided by the Ministry of Education and Culture, serves as a crucial document for data collection. These documents are analyzed to ascertain the alignment between the content of the summative tests and the prescribed learning objectives.

Procedures

The research methodology employs a content analysis design to examine the conformity and inconformity between the English summative tests and the learning objectives. Firstly, the researchers collected the English summative tests administered to fifth-grade students during the first semester from multiple elementary schools. Simultaneously, the English flow of learning objectives, as outlined by the Ministry of Education and Culture, is obtained. Subsequently, a systematic comparison is conducted between the content of the summative tests and the learning objectives. The researchers identified instances of alignment and discrepancies to evaluate the content validity of the tests. Finally, findings are synthesized to provide insights into the effectiveness of the summative assessment in measuring students' achievement of the prescribed learning objectives.

Data Analysis

The data analysis process in this qualitative research study involves a meticulous examination of the collected documents: The English summative tests for fifth grade in the first semester and the English flow of learning objectives provided by the Ministry of Education and Culture. The overarching objective is to assess the alignment between the content of the summative tests and the specified learning objectives. Initially, each English summative test is scrutinized to identify the test items and the corresponding topics or skills being assessed. Simultaneously, the English flow of learning objectives is thoroughly reviewed to understand the scope and depth of the intended curriculum standards.

Subsequently, a comparative analysis is conducted between the content of the summative tests and the learning objectives outlined in the curriculum documents. Each test item is assessed for its relevance to the learning objectives, with particular attention paid to the extent to which the test adequately covers the prescribed content and skills. Through this process, instances of conformity and inconformity between the summative tests and the learning objectives are identified and documented. Any discrepancies or deviations from the intended alignment are carefully noted. Furthermore, qualitative coding techniques may be employed to

categorize the identified themes or patterns within the data. This may involve coding for themes such as alignment, relevance, and coverage of learning objectives. Ultimately, the findings from the data analysis are synthesized to provide insights into the content validity of the English summative tests. These insights contribute to a comprehensive understanding of the extent to which the tests accurately measure students' achievement of the prescribed learning objectives.

RESULTS AND DISCUSSION

The results of the data analysis reveal both areas of conformity and inconformity between the English summative tests and the prescribed learning objectives outlined in the curriculum documents.

Conformity

Upon thorough analysis, numerous instances of alignment emerge between the content featured in the summative tests and the stipulated learning objectives. Notably, test items that directly correlate with the prescribed topics and skills outlined in the curriculum standards exhibit robust alignment. This alignment serves as a testament to the efficacy of the summative tests in effectively assessing specific facets of the curriculum. It signifies that the assessment instruments accurately target and evaluate the intended learning outcomes as delineated in the curriculum. This alignment instills confidence in the validity of the assessment process, suggesting that students are being evaluated on the essential content and skills deemed critical for their academic progression.

Moreover, it reflects positively on the efforts invested in crafting assessment tools that faithfully mirror the educational objectives set forth by the curriculum. Overall, the observed alignment underscores the importance of meticulous test design and validation procedures in ensuring that summative assessments accurately reflect the intended curriculum standards and effectively gauge student learning outcomes.

The concept of alignment, as articulated in Biggs and Tang's (2011) theory of constructive alignment, emphasizes the importance of ensuring that assessment tasks closely align with learning outcomes and instructional strategies. Our study's results, which reveal numerous instances of alignment between summative tests and curriculum objectives, support the principles of constructive alignment. This alignment serves to validate the efficacy of summative assessments in accurately measuring specific facets of the curriculum and evaluating intended learning outcomes.

Furthermore, the concept of validity, as elucidated by Messick, (1995), is highly relevant to our discussion. Messick's framework emphasizes the importance of ensuring that assessment instruments accurately measure the intended constructs. The observed alignment between summative tests and curriculum objectives in our study reflects positively on the content validity of the assessment instruments. This alignment suggests that the assessment tools effectively target and evaluate the essential content and skills specified in the curriculum, thereby instilling confidence in the validity of the assessment process.

Moreover, our findings align with previous research by Campbel, (2020); Pakpahan et al., (2023); Shafi et al., (2020), which underscore the importance of

alignment between assessment tasks and curriculum standards. Consistent with these studies, our findings highlight the critical role of meticulous test design and validation procedures in ensuring that summative assessments accurately reflect the intended curriculum standards and effectively gauge student learning outcomes.

Overall, the observed alignment between summative tests and curriculum objectives underscores the importance of rigorous assessment practices in educational settings (Puspitasari, 2022; Putri, 2018; Suek, 2021). By aligning assessment tasks with learning objectives, educators can enhance the validity and reliability of summative assessment processes, thereby facilitating meaningful evaluation of student learning and informing instructional decision-making (Maruf, 2023; Maruf & Helingo, 2022).

Inconformity

While our analysis revealed significant alignment between the content of the summative tests and the stipulated learning objectives, it also brought to light certain areas of incongruity. These instances of inconformity point to potential shortcomings within the assessment process, where certain test items may deviate from the intended curriculum content or fail to adequately assess specific prescribed skills.

These disparities between the content of the summative tests and the learning objectives raise concerns regarding the comprehensiveness and accuracy of the assessment instruments. When test items stray from the prescribed curriculum content, they may overlook crucial topics or skills deemed essential for student mastery. Similarly, when certain prescribed skills are not adequately assessed, there is a risk that students may not receive a comprehensive evaluation of their learning. Moreover, the identification of incongruence underscores the importance of ongoing evaluation and refinement of assessment practices. By recognizing areas of misalignment, educators can take proactive measures to address these discrepancies and enhance the effectiveness of summative assessments.

Furthermore, the existence of incongruity highlights the need for greater attention to detail in the design and validation of assessment instruments. Educators and test developers must meticulously align test items with the intended curriculum content and learning objectives to ensure the validity and reliability of summative assessments (Maruf & Anjely, 2020; Zuhri et al., 2021). Overall, the identification of inconformity between summative tests and learning objectives underscores the complexity of the assessment process and the importance of continuous improvement. By addressing these areas of incongruence, educators can strive to create more robust assessment tools that accurately reflect the intended curriculum standards and provide meaningful insights into student learning outcomes.

CONCLUSION

This study has provided valuable insights into the alignment between summative tests and curriculum objectives in the context of educational assessment. Through a meticulous analysis of test content and curriculum standards, we have identified both areas of alignment and inconformity. These findings reveal that numerous instances of alignment exist between the content of the summative tests and the stipulated learning objectives. Test items that directly correlate with

prescribed topics and skills demonstrate robust alignment, serving as a testament to the efficacy of the assessment process in evaluating specific facets of the curriculum. This alignment instills confidence in the validity of summative assessments, indicating that students are being evaluated on essential content and skills critical for their academic progression.

However, the analysis also identified areas of incongruity between summative tests and learning objectives. Some test items may stray from the intended curriculum content or fail to adequately assess certain prescribed skills. These disparities highlight potential gaps in the assessment process, raising concerns about the comprehensiveness and accuracy of the evaluation of student learning.

Moving forward, it is imperative for educators and curriculum developers to address these areas of incongruence and strive for greater alignment between summative tests and curriculum objectives. By adopting rigorous assessment practices and refining test design and validation procedures, educators can enhance the validity and reliability of summative assessments, thereby providing meaningful insights into student learning outcomes.

Ultimately, the findings of this study underscore the importance of ongoing evaluation and refinement of assessment practices in educational settings. By ensuring alignment between summative tests and curriculum objectives, educators can effectively measure student achievement and inform instructional decision-making to support student success.

REFERENCES

- Adha, K. (2014). an Analysis on the Content Validity
- Adiredja, R. K. (2012). Content validity analysis on achievement test at a private islamic junior high school.
- Afni Fitriana, N., Slamet, I., & Sujadi, I. (2018). *Classroom Assessment: Student Achievement in Mathematics through Formative Assessment*. <https://doi.org/10.2991/incomed-17.2018.44>
- Alderson, J., Clapham, C., & Wall, D. (1995). *Language test construction and evaluation*. Cambridge: Cambridge University Press.
- Ali, I., & Muhammad Iqbal, H. (2013). Effect of formative assessment on students' achievement in science. *World Applied Sciences Journal*, 26(5). <https://doi.org/10.5829/idosi.wasj.2013.26.05.1622>
- Areta, :, & Dari, W. (2014). an Analysis on the Content Validity of English Summative Test Items At the Even Semester of the Second Grade of Junior High School.
- Boström, E., & Palm, T. (2023). The effect of a formative assessment practice on student achievement in mathematics. *Frontiers in Education*, 8. <https://doi.org/10.3389/educ.2023.1101192>
- Campbel, S. (2020). Education and Curriculum Reform: The Impact They Have On Learning. *Budapest International Research and Critics in Linguistics and Education (BirLE) Journal*, 3(2). <https://doi.org/10.33258/birle.v3i2.1036>
- Chen, Y. (2021). Exploring Secondary School EFL Teachers' Assessment Literacy in Practice: A Case Study in China. *English Language Teaching*, 14(12).

<https://doi.org/10.5539/elt.v14n12p1>

- de Vries, J. A., Dimosthenous, A., Schildkamp, K., & Visscher, A. J. (2022). The impact on student achievement of an assessment for learning teacher professional development program. *Studies in Educational Evaluation*, 74. <https://doi.org/10.1016/j.stueduc.2022.101184>
- de Vries, Jitske A., Dimosthenous, A., Schildkamp, K., & Visscher, A. J. (2023). The impact of an assessment for learning teacher professional development program on students' metacognition. *School Effectiveness and School Improvement*, 34(1). <https://doi.org/10.1080/09243453.2022.2116461>
- Gupta, P., Shah, D., & Singh, T. (2021). Competency-Based Assessment in Pediatrics for the New Undergraduate Curriculum. *Indian Pediatrics*, 58(8). <https://doi.org/10.1007/s13312-021-2290-7>
- Haynes, S. N., Richard, D. C., & Kubany, E. S. (1995). Content validity in psychological assesment: A functional approach to concepts and methods. American Psychological Association, 238-247.
- Heaton, J. B. (1998). Writing english language test. London: Longman Group, New Edition.
- Hughes, A. (2003). Testing for English Teacher. New York: Cambridge University Press.
- Jarvis, J. (2015). Using Formative Assessment Practices To Lift Student Achievement. *Journwal of Professional Learning*, (Sadler 1989).
- Luthfiyyah, R., Cahyono, B. Y., Ivone, F. M., & Suryati, N. (2022). EFL Secondary School Teachers' Conceptions of Online Assessment in Emergency Remote Teaching: A Phenomenographic Study. *Studies in English Language and Education*, 9(3). <https://doi.org/10.24815/siele.v9i3.23459>
- Maruf, N. (2023). The Interplay Of Teachers' Beliefs, Attitudes, And The Implementation Of Differentiated Instruction In Indonesian Efl Contexts. *English Review: Journal of English Education*, 11(2). <https://doi.org/10.25134/erjee.v11i2.7251>
- Maruf, N., & Anjely, A. M. R. (2020). Utilizing Cooperative Integrated Reading and Composition (CIRC) with mobile Learning to Enhance Students' Reading Comprehension. *British (Jurnal Bahasa Dan Sastra Inggris)*, 9(2), 10–19.
- Maruf, N., & Helingo, A. (2022). Assessment Strategy to Rectify EFL Students' Performance: A Need Analysis. *JET ADI BUANA*, 7(02). <https://doi.org/10.36456/jet.v7.n02.2022.6191>
- Mary, F., & Sake, S. (1983). Foreign language testing. New York, NY: Regents Publishing.
- Messick, S. (1995). Messick 1995 Validity of Psychological Assessment. *American Psychologist*, 50(9).
- Nofiyanti. (2011). An Analysis On The Content Validity Of The Summative Test For The First Year Students Of Department Of English Education Faculty Of Tarbiyah And Teahers ' Training State Islamic University Syarif Hidayatullah An Analysis On The Content Validity Of The Sum.
- Nugrahanto, A., Winarsih, D., & Farikah. (2018). The Content Validity Of The

- Summative Test Items of English For The Tenth Graders Of Sma Negeri 1 Magelang In The School Year 2015/2016. *Journal of Research on Applied Linguistics Language and Language Teaching*, 1(1), 7–14
- Pakpahan, H. M., Suherni, S., Pujiati, L., & Girsang, R. (2023). Effectiveness of Indonesian Education Curriculum Reform on the Quality of Processes in Learning. *Jurnal Penelitian Pendidikan IPA*, 9(1). <https://doi.org/10.29303/jppipa.v9i1.3930>
- Puspitasari, E. (2022). AN ANALYSIS OF CONTENT VALIDITY ON SUMMATIVE TEST FOR THE 12th GRADE STUDENTS. *JEPAL (Journal of English Pedagogy and Applied Linguistics)*, 2(2). <https://doi.org/10.32627/jepal.v2i2.420>
- Putri, B. D. T. (2018). THE VALIDITY ANALYSIS OF ENGLISH SUMMATIVE TEST OF JUNIOR HIGH SCHOOL. *Journal of Languages and Language Teaching*, 5(1). <https://doi.org/10.33394/jollt.v5i1.327>
- Shafi, M. M., Neyestani, M. R., Jafari, S. E. M., & Taghvaei, V. (2020). The Quality Improvement Indicators of the Curriculum at the Technical and Vocational Higher Education. *International Journal of Instruction*, 14(1). <https://doi.org/10.29333/IJI.2021.1415A>
- Siddiek, D. A. (2010). The impact of test content validity on language Teaching and learning. 2010.
- Suek, L. A. (2021). ITEM ANALYSIS OF AN ENGLISH SUMMATIVE TEST. *PEJLaC: Pattimura Excellence Journal of Language and Culture*, 1(1). <https://doi.org/10.30598/pejlac.v1.i1.pp9-18>
- Sultana, R., Shamim, K. M., Nahar, L., & Hasan, F. (2009). Content validity of written examinations in undergraduate anatomy. Bangladesh: *Journal of Anatomy*.
- Weeden, P., Winter, J., & Broadfoot, P. (2002). *Assessment what's in it for school?* London and New York: Routledge Falmer.
- Zuhri, S., Anwar, K., & Maruf, N. (2021). The correlation between extensive reading , critical reading , and self- esteem in students ' reading abilities. *Budapest International Research and Critics Institute-Journal (BIRCI-Journal)*, 4(3).