

## ABSTRAK

Putri, Eka Ayu Amelia (2025), Pengembangan Instrumen Tes Diagnostik Bentuk *Four-Tier* Pada Mata Pelajaran Bahasa Indonesia Tingkat SMP. Tesis, Penelitian dan Evaluasi Pendidikan, Program Pascasarjana, Universitas Pendidikan Ganesha.

Tesis ini sudah disetujui dan diperiksa oleh Pembimbing I: Dr. Ir. I Gede Ratnaya, S.T., M.Pd., MCE. dan Pembimbing II: Dr. Nyoman Trisna Herawati, S.E., Ak., M.Pd.

Kata-kata kunci: instrumen diagnostik, *four-tier*, pemahaman konsep, Bahasa Indonesia, SMP.

Penelitian ini bertujuan untuk mengembangkan instrumen tes diagnostik bentuk *four-tier* pada materi Buku Fiksi dan Nonfiksi dalam mata pelajaran Bahasa Indonesia kelas VII SMP. Pengembangan instrumen ini dimaksudkan untuk mengidentifikasi pemahaman konsep peserta didik secara lebih mendalam, termasuk membedakan peserta didik yang paham konsep, mengalami miskonsepsi, dan tidak paham konsep. Selain itu, penelitian ini bertujuan menghasilkan instrumen evaluasi yang valid dan reliabel sebagai dasar bagi guru dalam merancang tindak lanjut pembelajaran berupa remedial dan pengayaan.

Penelitian ini menggunakan metode *Research and Development* (R&D) dengan model pengembangan Dick and Carey yang telah dimodifikasi. Penelitian dilaksanakan di SMP Negeri 13 Malang. Teknik pengambilan sampel menggunakan *purposive sampling*, dengan subjek penelitian sebanyak 129 peserta didik kelas VII. Pengumpulan data menggunakan instrumen penelitian berupa tes diagnostik bentuk *four-tier* sebanyak 20 butir soal yang memuat jawaban, tingkat keyakinan, alasan, dan tingkat keyakinan terhadap alasan. Seluruh instrumen telah diuji validitasnya, meliputi validitas isi melalui *expert judgment* oleh ahli evaluasi instrumen dan ahli materi dengan menggunakan formula Gregory, validitas konsistensi butir melalui analisis korelasi point biserial, validitas konstruk konvergen melalui nilai *loading factor*, *Average Variance Extracted* (AVE), dan *Composite Reliability* (CR), serta validitas kriteria konkuren melalui analisis korelasi dengan tes formatif. Reliabilitas instrumen diuji menggunakan rumus KR-20, sedangkan teknik analisis data dilakukan secara kuantitatif deskriptif.

Hasil penelitian menunjukkan bahwa: (1) instrumen yang dikembangkan memiliki nilai validitas isi Gregory berada pada kategori sangat valid, dengan skor rata-rata dari ahli evaluasi instrumen sebesar 0.900 dan ahli materi sebesar 0.850; (2) Uji validitas konsistensi butir soal juga dilakukan dengan menganalisis butir soal menggunakan rumus korelasi point biserial yang sudah di uji cobakan kepada 60 peserta didik sebagai sampel. Pada uji validitas butir soal *tier* 1 dan *tier* 3 didapatkan nilai  $r_{hitung}$  dari setiap butir soal memiliki nilai  $r_{hitung} \geq r_{table}$  sehingga dapat dinyatakan bahwa setiap butir soal tersebut valid; (3) Validitas konstruk konvergen dianalisis melalui dua tahap pengujian menggunakan nilai *loading factor*  $\geq 0.7$ , *Average Variance Extracted* (AVE)  $\geq 0.5$ , dan *Composite Reliability* (CR)  $\geq 0.7$ , yang menunjukkan bahwa semua indikator valid secara konstruk. Validitas kriteria juga memenuhi syarat dengan nilai korelasi antara skor instrumen dengan tes formatif sebesar 0.526 (kategori cukup); (4) Reliabilitas instrumen diperoleh

melalui pengujian KR-20 dengan nilai koefisien reliabilitas *tier* 1 sebesar 0.878 dan *tier* 3 sebesar 0.847 yang menurut pedoman Guilford termasuk dalam kategori sangat tinggi; dan (5) Implementasi instrumen menunjukkan bahwa dari 129 peserta didik, sebanyak 12% mengalami miskonsepsi, 75% paham konsep, dan 13% tidak paham konsep. Berdasarkan temuan tersebut dapat disimpulkan bahwa instrumen tes diagnostik bentuk *four-tier* yang dikembangkan efektif untuk mengidentifikasi pemahaman konsep peserta didik.



## ABSTRACT

Putri, Eka Ayu Amelia (2025). Development of a Four-Tier Diagnostic Test Instrument for Indonesian Language Subjects at the Junior High School Level. Thesis, Educational Research and Evaluation, Postgraduate Program, Ganesha University of Education.

This thesis has been approved and examined by Supervisor I: Dr. Ir. I Gede Ratnaya, S.T., M.Pd., MCE., and Supervisor II: Dr. Nyoman Trisna Herawati, S.E., Ak., M.Pd.

*Key words:* iagnostic instrument, four-tier, conceptual understanding, Indonesian Language, junior high school.

This study aimed to develop a four-tier diagnostic test instrument on Fiction and Nonfiction Texts in the Indonesian Language subject for seventh-grade junior high school students. The development of this instrument was intended to identify students' conceptual understanding in greater depth, including distinguishing students who understand the concepts, those who hold misconceptions, and those who do not understand the concepts. In addition, this study sought to produce a valid and reliable evaluation instrument to serve as a basis for teachers in designing follow-up instructional actions in the form of remedial and enrichment programs.

This study employed a Research and Development (R&D) method using a modified Dick and Carey development model. The research was conducted at SMP Negeri 13 Malang. The sampling technique used was purposive sampling, involving 129 seventh-grade students as research subjects. Data were collected using a four-tier diagnostic test instrument consisting of 20 items, each comprising an answer choice, a confidence level, a reason, and a confidence level for the reason. The instrument validity was tested through several procedures, including content validity assessed through expert judgment by instrument evaluation experts and subject matter experts using the Gregory formula, item consistency validity analyzed using point biserial correlation, convergent construct validity analyzed using loading factor values, Average Variance Extracted (AVE), and Composite Reliability (CR), as well as concurrent criterion validity analyzed through correlation with formative test scores. Instrument reliability was tested using the KR-20 formula, while data analysis techniques were conducted using quantitative descriptive analysis.

The results of the study showed that: (1) the developed instrument demonstrated very high content validity based on the Gregory index, with an average score of 0.900 from the instrument evaluation expert and 0.850 from the subject matter expert; (2) item consistency validity was tested using point biserial correlation on 60 students as a trial sample, and the results indicated that all items in tier 1 and tier 3 had r-count values greater than or equal to the r-table value, confirming that all items were valid; (3) convergent construct validity was analyzed in two stages using loading factor values  $\geq 0.7$ , Average Variance Extracted (AVE)  $\geq 0.5$ , and Composite Reliability (CR)  $\geq 0.7$ , indicating that all indicators were construct-valid. Criterion validity also met the required criteria, with a correlation coefficient between the instrument scores and formative test scores of 0.526 (moderate

category); (4) instrument reliability, tested using the KR-20 formula, yielded reliability coefficients of 0.878 for tier 1 and 0.847 for tier 3, which, according to Guilford's criteria, fall into the very high reliability category; and (5) the implementation of the instrument revealed that among the 129 students, 12% experienced misconceptions, 75% demonstrated conceptual understanding, and 13% did not understand the concepts. Based on these findings, it can be concluded that the developed four-tier diagnostic test instrument is effective in identifying students' conceptual understanding.

