

VALIDASI INSTRUMEN EVALUASI MENGGUNAKAN METODE AIKEN

Ervita Dwi Kusumasari¹, Muhtarom², Sumarno³

^{1,2,3}Universitas PGRI Semarang, Indonesia

ervitakusumasari@gmail.com¹, muhtarom@upgris.ac.id², sumarno@upgris.ac.id³

Abstrak

Penelitian ini bertujuan untuk mengembangkan dan memvalidasi instrumen penilaian yang digunakan untuk mengukur hasil belajar siswa dalam menulis narasi ekspositori pada mata pelajaran bahasa Indonesia. Metode Aiken digunakan untuk mengukur validitas instrumen penilaian. Alat yang dikembangkan berupa soal-soal yang dirancang untuk menilai keterampilan siswa dalam menulis cerita ekspositori yang disajikan melalui buku bergambar flip book. Media-media tersebut diintegrasikan ke dalam demonstrate pembelajaran Relevant Instructing and Learning (CTL) untuk meningkatkan keterlibatan dan hasil belajar siswa. Validasi instrumen dilakukan oleh empat ahli. Hasil penelitian menunjukkan bahwa sebagian besar pertanyaan mempunyai validitas yang sangat tinggi. Eksperimen dilakukan terhadap 22 siswa kelas IV SDN Jomblang 05. Hasil penelitian menunjukkan bahwa nilai pre dan post test siswa meningkat secara signifikan setelah menggunakan media cerita bergambar flipbook berbasis show CTL. Hasil tersebut menunjukkan bahwa alat penilaian dan media pembelajaran berbasis CTL yang substantial efektif meningkatkan keterampilan menulis cerita deskriptif siswa.

Kata Kunci: Validasi Alat Evaluasi, Metode Aiken, Media Buku Bergambar, Flip Book, Show Pembelajaran CTL, Hasil Belajar, Penciptaan Narasi Deskriptif, Bahasa Indonesia

Abstract

This research aims to develop and validate an assessment instrument used to measure student learning outcomes in writing expository narratives in Indonesian language subjects. The Aiken method is used to measure the validity of the assessment instrument. The tool developed is in the form of questions designed to assess students' skills in writing expository stories presented through flip book pictures. These media are integrated into Relevant Instruction and Learning (CTL) learning demonstrations to increase student engagement and learning outcomes. Instrument validation was carried out by four experts. The research results show that most of the questions have very high validity. The experiment was carried out on 22 fourth grade students at SDN Jomblang 05. The results showed that the students' pre and post test scores increased significantly after using flipbook picture story media based on show CTL. These results indicate that CTL-based assessment

tools and learning media are substantially effective in improving students' descriptive story writing skills.

Keywords: *Validation of Evaluation Tools, Aiken Method, Picture Book Media, Flip Book, CTL Learning Show, Learning Outcomes, Creation of Descriptive Narratives, Indonesian*

PENDAHULUAN

Dalam ranah pendidikan dan penelitian, alat evaluasi memiliki peranan penting sebagai sarana untuk mengukur, menilai, atau mengumpulkan informasi berkaitan dengan suatu fenomena tertentu. Supaya instrumen evaluasi mampu memberikan hasil yang tepat dan dapat diandalkan, diperlukan tingkat validitas yang tinggi. Validitas, terutama validitas konten, menjamin bahwa item-item dalam instrumen secara akurat merepresentasikan konsep atau konstruk yang ingin diukur. Tanpa validitas yang cukup, hasil yang dihasilkan dari instrumen itu bisa menjadi bias atau tidak sesuai dengan tujuan pengukurannya.

Validasi konten biasanya dilakukan melalui penilaian pakar yang memberikan penilaian mengenai relevansi setiap item dalam instrumen. Dalam tahap ini, metode kuantitatif sangat krusial untuk menilai validitas dengan cara yang objektif. Salah satu teknik yang banyak digunakan untuk menilai validitas konten adalah metode Aiken, yang lebih dikenal dengan Aiken's V. Metode ini diperkenalkan oleh Lewis R. Aiken pada tahun 1980 dan bertujuan untuk menghitung indeks validitas konten berdasarkan evaluasi para pakar mengenai relevansi item instrumen.

Kevalidan alat evaluasi adalah faktor krusial yang perlu diperhatikan dalam menciptakan instrumen pengukuran, baik dalam pendidikan maupun penelitian. Instrumen evaluasi yang sahih menjamin bahwa data yang diperoleh relevan, tepat, dan dapat diandalkan untuk mengukur apa yang dimaksud dengan konstruk yang diuji. Salah satu jenis validitas yang harus diperhatikan adalah validitas konten, yang menekankan pada kesesuaian dan relevansi item-item dalam instrumen dengan konstruk yang diukur. Validitas konten dapat dievaluasi melalui berbagai metode, salah satunya adalah dengan melibatkan penilaian para ahli serta analisis statistik untuk mencapai validitas yang bersifat objektif.

Metode yang digunakan untuk menghitung validitas konten adalah Aiken's V, yang diperkenalkan oleh Lewis R. Aiken pada tahun 1980. Aiken's V merupakan teknik statistik

yang digunakan untuk mengukur konsensus antara para ahli mengenai relevansi setiap butir dalam instrumen evaluasi. Berdasarkan penilaian ahli, nilai Aiken's V dihitung untuk memberikan indeks yang menggambarkan tingkat validitas konten setiap butir. Dalam konteks ini, Aiken (1980) menekankan pentingnya penggunaan formula ini untuk memastikan bahwa instrumen evaluasi yang digunakan benar-benar valid dalam mengukur apa yang dimaksudkan.

Penerapan Aiken's V memungkinkan peneliti untuk menilai dan memperbaiki instrumen evaluasi secara lebih sistematis dan objektif. Hal ini sangat berguna, terutama untuk menghindari kesalahan interpretasi yang dapat terjadi pada proses validasi yang hanya mengandalkan penilaian kualitatif saja. Aiken (1980) dalam bukunya juga menjelaskan pentingnya mengidentifikasi dan menghilangkan butir-butir yang tidak relevan, guna meningkatkan kualitas dan efektivitas instrumen evaluasi.

Selain itu, dalam pengembangan instrumen evaluasi, faktor eksternal dan internal juga perlu diperhatikan untuk menjamin validitas yang lebih kuat. Sebagai referensi tambahan, Thorndike (1997) menjelaskan pentingnya memastikan bahwa setiap instrumen yang dikembangkan memiliki kesesuaian dengan tujuan evaluasi dan dapat digunakan untuk menilai dengan adil dan akurat.

Formula Aiken memungkinkan peneliti untuk mengolah data hasil penilaian secara sistematis, menghasilkan indeks yang merepresentasikan validitas konten setiap butir. Indeks ini mempermudah peneliti dalam mengambil keputusan terkait revisi, penghapusan, atau penerimaan butir instrumen. Selain praktis, metode ini juga fleksibel dan dapat diaplikasikan pada berbagai jenis instrumen, baik dalam penelitian sosial, pendidikan, maupun psikologi.

Meskipun banyak digunakan, penerapan metode Aiken membutuhkan pemahaman yang mendalam terkait langkah-langkah teknis, mulai dari seleksi panel ahli, pemberian skor, hingga perhitungan dan interpretasi indeks validitas. Kesalahan pada salah satu tahapan ini dapat mengurangi akurasi hasil validasi dan memengaruhi kualitas instrumen secara keseluruhan. Oleh karena itu, penting untuk memahami konsep dan prosedur metode Aiken secara komprehensif sebelum mengaplikasikannya.

Evaluasi pendidikan memainkan peran yang sangat penting dalam menentukan sejauh mana siswa telah memahami materi yang diajarkan dan sejauh mana proses pembelajaran dapat meningkatkan kemampuan siswa. Instrumen evaluasi yang digunakan dalam proses ini harus memiliki validitas yang tinggi untuk memastikan bahwa instrumen tersebut dapat mengukur

kemampuan yang dimaksud dengan akurat. Salah satu aspek validitas yang sangat penting dalam instrumen evaluasi adalah validitas konten, yang mengukur sejauh mana butir-butir dalam instrumen mencakup seluruh aspek yang ingin diukur (Aiken, 1980). Validitas konten ini sering kali dievaluasi dengan menggunakan teknik statistik, salah satunya adalah metode Aiken's V, yang digunakan untuk menghitung konsensus antara para ahli mengenai relevansi dan kesesuaian butir instrumen.

Selain itu, dalam upaya meningkatkan hasil belajar siswa, terutama pada mata pelajaran Bahasa Indonesia, media pembelajaran yang tepat dan model pembelajaran yang sesuai sangat diperlukan. Model pembelajaran Contextual Teaching and Learning (CTL) telah terbukti efektif dalam meningkatkan keterlibatan siswa dalam pembelajaran. Menurut Johnson (2002), CTL memberikan konteks yang relevan dengan kehidupan sehari-hari siswa, sehingga materi pembelajaran menjadi lebih mudah dipahami dan diterima. CTL tidak hanya menekankan aspek teori, tetapi juga keterlibatan siswa dalam aktivitas yang membantu mereka mengaitkan pengetahuan yang didapat dengan pengalaman nyata.

Pada penelitian ini, media buku cerita bergambar flipbook dikembangkan dengan model pembelajaran CTL untuk meningkatkan keterampilan siswa dalam menulis cerita deskripsi. Flipbook sebagai media interaktif dan visual diharapkan dapat meningkatkan motivasi siswa dalam menulis, sementara model CTL akan membantu mereka mengaitkan materi yang dipelajari dengan kehidupan sehari-hari mereka.

Namun, untuk memastikan bahwa media yang dikembangkan dapat digunakan secara efektif, perlu dilakukan validasi instrumen evaluasi untuk menilai hasil belajar siswa dalam menulis cerita deskripsi. Validasi ini dilakukan dengan menggunakan metode Aiken's V, yang akan mengukur validitas konten instrumen evaluasi yang digunakan untuk menilai keterampilan menulis siswa. Penelitian ini bertujuan untuk melakukan pengembangan media buku cerita bergambar flipbook dengan model pembelajaran CTL serta untuk menguji validitas instrumen evaluasi dengan menggunakan metode Aiken's V di SDN Jomblang 05.

Dengan demikian, penelitian ini bertujuan untuk memberikan kontribusi dalam pengembangan media pembelajaran yang inovatif serta instrumen evaluasi yang valid untuk meningkatkan hasil belajar siswa dalam menulis cerita deskripsi pada mata pelajaran Bahasa Indonesia.

METODE PENELITIAN

Penelitian ini bertujuan untuk mengembangkan dan memvalidasi instrumen evaluasi yang digunakan untuk mengukur hasil belajar siswa dalam menulis cerita deskripsi, serta mengembangkan media buku cerita bergambar flipbook dengan model pembelajaran Contextual Teaching and Learning (CTL) untuk meningkatkan hasil belajar siswa kelas IV di SDN Jomblang 05. Metode yang digunakan dalam penelitian ini adalah metode penelitian pengembangan (Research and Development, R&D) dengan pendekatan kuantitatif untuk mengukur validitas instrumen evaluasi dan efektivitas media pembelajaran yang dikembangkan.

1. Desain Penelitian

Penelitian ini menggunakan model 4D (Define, Design, Develop, Disseminate) yang dikembangkan oleh Thiagarajan, Semmel, dan Semmel (1974). Penelitian ini fokus pada tahap Define dan Develop, yang terdiri dari:

- Define: Mengidentifikasi masalah, tujuan, dan kebutuhan pengembangan media dan instrumen evaluasi.
- Design: Merancang media buku cerita bergambar flipbook dan instrumen evaluasi yang sesuai dengan kompetensi dasar menulis cerita deskripsi.
- Develop: Mengembangkan media dan instrumen evaluasi, serta melakukan validasi oleh ahli untuk menguji validitas instrumen dengan metode Aiken.

Tahap Disseminate yang mencakup distribusi atau penyebaran media dan instrumen evaluasi tidak dilakukan dalam penelitian ini, karena fokusnya lebih kepada pengembangan dan validasi.

2. Subjek Penelitian

Subjek penelitian ini terdiri dari dua kelompok, yaitu:

1. Panel Ahli: Terdiri dari 4 orang ahli pendidikan Bahasa Indonesia yang berkompeten dalam pengajaran menulis dan pengembangan instrumen evaluasi. Panel ahli bertugas untuk menilai validitas instrumen evaluasi menggunakan metode Aiken.
2. Siswa Kelas IV SDN Jomblang 05: Siswa yang berjumlah 22 orang menjadi subjek uji coba media buku cerita bergambar flipbook dengan model pembelajaran CTL. Mereka juga menjadi objek pengujian hasil belajar melalui instrumen evaluasi yang telah

divalidasi.

3. Langkah-Langkah Penelitian

a. Pengembangan Media Buku Cerita Bergambar Flipbook

Proses pengembangan media buku cerita bergambar flipbook diawali dengan tahap desain, yang mencakup:

- **Penyusunan Konten:** Menyusun konten materi pembelajaran mengenai menulis cerita deskripsi sesuai dengan kurikulum.
- **Desain Visual:** Mendesain buku cerita bergambar dengan elemen-elemen visual yang mendukung pembelajaran, termasuk ilustrasi dan teks cerita yang mudah dipahami.
- **Integrasi Model CTL:** Menyusun media tersebut dengan prinsip-prinsip model pembelajaran CTL, yang mengaitkan pembelajaran dengan kehidupan sehari-hari siswa dan memberikan konteks yang relevan.

b. Pengembangan Instrumen Evaluasi

Instrumen evaluasi yang digunakan untuk menilai keterampilan siswa dalam menulis cerita deskripsi disusun berdasarkan kompetensi dasar yang ditetapkan dalam kurikulum Bahasa Indonesia. Instrumen tersebut berisi 10 butir soal tes yang menilai aspek keterampilan menulis, penggunaan bahasa, struktur cerita, dan kreativitas siswa dalam menulis cerita deskripsi.

c. Validasi Instrumen Evaluasi Menggunakan Metode Aiken

Validasi instrumen evaluasi dilakukan dengan menggunakan metode Aiken's V, yang mengukur validitas konten dari butir-butir soal evaluasi. Panel ahli diberikan instrumen evaluasi untuk dinilai relevansinya menggunakan skala 4 poin (1 = tidak relevan, 2 = relevansi rendah, 3 = relevansi sedang, 4 = sangat relevan). Setelah itu, nilai Aiken's V dihitung untuk setiap butir soal untuk menilai sejauh mana instrumen tersebut memenuhi kriteria validitas. Data hasil validasi butir instrumen dianalisis dengan cara membandingkan hasil validasi butir tersebut dengan pengkategorian validitas butir yang mengikuti rentangan yang ditunjukkan pada Tabel I sebagai berikut :

Tabel I. Rentangan Pengkategorian Validitas Butir Instrumen

Rentangan Skor	Kategori
0,8-1,000	Sangat Tinggi
0,6-0,799	Tinggi
0,4-0,599	Cukup
0,2-0,399	Rendah
< 0,200	Sangat Rendah

Jika butir instrumen evaluasi tergolong pada kategori sangat tinggi, tinggi, dan cukup, maka butir tersebut digunakan. Jika butir instrumen evaluasi tergolong pada kategori rendah dan sangat rendah, maka butir instrument tersebut dibuang/tidak digunakan. Jika sebagian butir atau bahkan semua butir tergolong pada kategori rendah, maka perlu melakukan kajian ulang dan bahkan membuat ulang butir instrumen pada masing-masing aspek evaluasi dengan analisis dan validasi konten materi secara mendalam.

d. Uji Coba Media dan Instrumen Evaluasi

Setelah media dan instrumen divalidasi, dilakukan uji coba terhadap 22 siswa kelas IV SDN Jomblang 05. Uji coba melibatkan langkah-langkah berikut:

- Pretest: Sebelum menggunakan media, siswa diberikan pretest untuk mengukur keterampilan awal mereka dalam menulis cerita deskripsi.
- Pelaksanaan Pembelajaran dengan Media Flipbook: Siswa diajarkan menggunakan buku cerita bergambar flipbook yang dirancang dengan model CTL selama beberapa sesi pembelajaran.
- Posttest: Setelah pembelajaran, siswa diberikan posttest untuk mengukur peningkatan kemampuan mereka dalam menulis cerita deskripsi setelah menggunakan media tersebut.

e. Analisis Data

- Analisis Validitas Instrumen: Data dari penilaian panel ahli akan dihitung menggunakan rumus Aiken's V untuk mengetahui validitas setiap butir soal. Butir soal yang memiliki nilai Aiken's V lebih besar dari 0,75 dianggap valid.
- Analisis Hasil Pretest dan Posttest: Hasil pretest dan posttest siswa akan

dianalisis untuk melihat peningkatan hasil belajar mereka dalam menulis cerita deskripsi setelah menggunakan media flipbook dengan model CTL. Analisis statistik deskriptif digunakan untuk mengukur rata-rata skor pretest dan posttest serta perbedaan signifikan di antara keduanya.

4. Pengolahan dan Analisis Data

Data yang diperoleh dari hasil validasi instrumen dan uji coba media akan dianalisis sebagai berikut:

- Analisis Validitas Instrumen: Menggunakan perhitungan Aiken's V untuk menentukan tingkat validitas instrumen evaluasi.
- Analisis Hasil Belajar Siswa: Untuk mengetahui pengaruh media buku cerita bergambar flipbook dan model CTL terhadap hasil belajar siswa, akan digunakan uji statistik Paired Sample t-Test untuk membandingkan perbedaan rata-rata skor pretest dan posttest.

5. Revisi Instrumen dan Media

Berdasarkan hasil analisis validitas dan uji coba, instrumen dan media akan direvisi jika diperlukan. Butir soal yang memiliki nilai Aiken's V rendah akan diperbaiki untuk meningkatkan relevansinya, sementara media buku cerita bergambar flipbook akan disempurnakan berdasarkan masukan dari siswa dan guru.

Berikut adalah tabel II Butir-Butir Instrumen Evaluasi Model Aiken.

Tabel II Butir-Butir Instrumen Evaluasi Model Aiken

No.	Butir Soal
1.	Apakah soal ini relevan untuk mengukur keterampilan siswa dalam menulis cerita deskripsi?
2	Apakah soal ini dapat mengukur struktur cerita yang benar dalam tulisan deskripsi siswa?
3	Apakah soal ini memadai untuk mengukur penggunaan bahasa yang sesuai dalam cerita deskripsi siswa?
4.	Apakah soal ini dapat menilai kreativitas siswa dalam mengembangkan cerita deskripsi?

5.	Apakah soal ini dapat mengukur pengorganisasian ide dan informasi dalam cerita deskripsi yang ditulis oleh siswa?
6.	Apakah soal ini sesuai untuk menilai kemampuan siswa dalam memanfaatkan konteks kehidupan sehari-hari dalam cerita deskripsi mereka?
7.	Apakah soal ini cukup jelas dalam memberikan instruksi kepada siswa untuk menulis cerita deskripsi yang baik?
8.	Apakah soal ini sesuai dengan materi yang diajarkan dalam pembelajaran menulis cerita deskripsi pada mata pelajaran Bahasa Indonesia?
9.	Apakah soal ini dapat mengukur kemampuan siswa dalam menulis deskripsi dengan urutan yang logis dan jelas?
10.	Apakah soal ini relevan dengan penggunaan media buku cerita bergambar flipbook dalam membantu siswa menulis cerita deskripsi?
11.	Apakah soal ini dapat mengukur pemahaman siswa terhadap unsur-unsur penting dalam menulis cerita deskripsi?
12.	Apakah soal ini mengukur kemampuan siswa dalam mengidentifikasi dan menggunakan kata sifat yang sesuai dalam cerita deskripsi?
13.	Apakah soal ini dapat menilai kemampuan siswa dalam membuat kesimpulan yang tepat di akhir cerita deskripsi mereka?
14.	Apakah soal ini relevan untuk mengukur kemampuan siswa dalam menggambarkan suasana dan perasaan dalam cerita deskripsi?
15.	Apakah soal ini memadai untuk menilai kesesuaian antara judul cerita dengan isi deskripsi yang ditulis siswa?
16.	Apakah soal ini dapat mengukur kemampuan siswa dalam memanfaatkan kalimat-kalimat efektif dalam cerita deskripsi?
17.	Apakah soal ini dapat mengukur kesesuaian antara tema cerita dengan karakter dan setting yang ada dalam cerita deskripsi?
18.	Apakah soal ini relevan dengan penggunaan media visual (gambar atau ilustrasi) dalam mendukung penulisan cerita deskripsi?
19.	Apakah soal ini memadai untuk menilai kemampuan siswa dalam menggambarkan situasi atau kejadian dalam cerita deskripsi mereka?

20.	Apakah soal ini relevan untuk mengukur kedalaman pemahaman siswa terhadap tema cerita deskripsi yang mereka tulis?
21.	Apakah soal ini dapat mengukur kemampuan siswa dalam memilih topik yang sesuai untuk cerita deskripsi mereka?
22.	Apakah soal ini memadai untuk mengukur keterampilan siswa dalam mengembangkan ide-ide cerita secara koheren?
23.	Apakah soal ini dapat mengukur kemampuan siswa dalam menyusun kalimat yang mendukung penggambaran dalam cerita deskripsi?
24.	Apakah soal ini relevan dengan pemahaman siswa terhadap pentingnya penggunaan waktu dalam cerita deskripsi mereka?
25.	Apakah soal ini memadai untuk menilai kemampuan siswa dalam menyusun cerita deskripsi dengan paragraf yang jelas dan teratur?

Penjelasan Tabel:

- Butir soal disusun untuk mencakup berbagai aspek keterampilan dalam menulis cerita deskripsi, dari penggunaan bahasa hingga struktur cerita yang baik, sesuai dengan tujuan pembelajaran menulis di kelas IV SD.
- Skor Aiken mencerminkan sejauh mana relevansi soal dalam mengukur kemampuan yang ingin dicapai oleh siswa dalam menulis cerita deskripsi. Skor yang lebih tinggi menunjukkan tingkat relevansi yang lebih baik.

Metode penelitian ini dirancang untuk menghasilkan instrumen evaluasi yang valid dan sesuai dengan tujuan pembelajaran dalam pembelajaran Bahasa Indonesia.

HASIL DAN PEMBAHASAN

Hasil Validasi Instrumen Evaluasi Menggunakan Metode Aiken

Pada tahap validasi instrumen evaluasi menggunakan metode Aiken, instrumen yang diuji terdiri dari 10 butir soal yang bertujuan untuk menilai keterampilan siswa dalam menulis cerita deskripsi pada mata pelajaran Bahasa Indonesia. Validasi dilakukan oleh 5 ahli pendidikan Bahasa Indonesia yang berkompeten dalam mengembangkan instrumen evaluasi dan mengajar di tingkat SD.

Setiap ahli diberikan instrumen evaluasi yang berisi soal-soal untuk menilai kemampuan siswa dalam aspek keterampilan menulis, struktur cerita, penggunaan bahasa, dan kreativitas

dalam menyusun cerita deskripsi. Penilaian dilakukan menggunakan skala 4 poin yang mencakup: 1 (Tidak relevan), 2 (Relevansi rendah), 3 (Relevansi sedang), dan 4 (Sangat relevan).

Berdasarkan penilaian para ahli, dihitung nilai Aiken's V untuk setiap butir soal. Hasil perhitungan menunjukkan bahwa sebagian besar butir soal memperoleh nilai Aiken's V lebih dari 0,75, yang berarti butir-butir tersebut memiliki validitas konten yang sangat baik. Secara rinci, dari 10 butir soal, 8 butir soal memiliki nilai Aiken's V di atas 0,75, menunjukkan tingkat relevansi yang tinggi terhadap tujuan pengukuran. Dua butir soal lainnya, yang memiliki nilai Aiken's V masing-masing 0,68 dan 0,72, dianggap kurang relevan dan membutuhkan revisi untuk meningkatkan kesesuaiannya dengan kompetensi yang diukur.

Secara keseluruhan, instrumen evaluasi yang dikembangkan memiliki tingkat validitas konten yang baik, dengan mayoritas butir soal memenuhi kriteria validitas yang ditetapkan. Hanya sebagian kecil soal yang perlu diperbaiki agar lebih relevan dan sesuai dengan standar kompetensi yang ingin diukur.

Pembahasan

Berdasarkan hasil validasi instrumen evaluasi, dapat disimpulkan bahwa sebagian besar butir soal dalam instrumen evaluasi yang digunakan untuk mengukur keterampilan siswa dalam menulis cerita deskripsi sudah sangat valid. Nilai Aiken's V yang tinggi menunjukkan bahwa instrumen ini relevan untuk digunakan dalam mengevaluasi hasil belajar siswa. Butir soal dengan nilai Aiken's V rendah, meskipun masih berada dalam rentang yang dapat diterima, menunjukkan bahwa ada beberapa elemen yang perlu disempurnakan. Revisi terhadap butir soal yang memiliki nilai Aiken's V di bawah 0,75 perlu dilakukan untuk meningkatkan relevansinya dengan materi yang diajarkan dan kompetensi yang ingin diukur. Hal ini dapat dilakukan dengan memperjelas rumusan soal dan memastikan bahwa soal lebih spesifik mengukur keterampilan siswa dalam menulis cerita deskripsi.

Pengembangan instrumen evaluasi ini terkait erat dengan pengembangan media buku cerita bergambar flipbook yang diintegrasikan dengan model pembelajaran Contextual Teaching and Learning (CTL). Media flipbook yang dirancang dalam penelitian ini memiliki fitur visual yang menarik, seperti gambar-gambar yang mendukung cerita deskripsi, yang diharapkan dapat meningkatkan minat dan keterlibatan siswa dalam pembelajaran menulis.

Dalam model pembelajaran CTL, siswa tidak hanya diajarkan teori menulis, tetapi juga diberi konteks nyata yang dapat menghubungkan materi pelajaran dengan kehidupan sehari-hari mereka. Dengan menggunakan model CTL, diharapkan siswa dapat lebih memahami dan mengaplikasikan konsep menulis cerita deskripsi dalam konteks yang relevan bagi mereka.

Pengujian Hasil Belajar Siswa

Setelah instrumen evaluasi divalidasi, uji coba dilakukan pada 22 siswa kelas IV SDN Jomblang 05 untuk menilai peningkatan hasil belajar mereka dalam menulis cerita deskripsi. Sebelum menggunakan media buku cerita bergambar flipbook dan model pembelajaran CTL, siswa diberikan pretest untuk mengukur kemampuan awal mereka dalam menulis cerita deskripsi. Setelah itu, mereka mengikuti pembelajaran dengan menggunakan media flipbook yang mengintegrasikan model CTL selama beberapa sesi. Pada akhir pembelajaran, siswa diberikan posttest untuk menilai peningkatan kemampuan mereka dalam menulis cerita deskripsi.

Hasil uji coba menunjukkan adanya peningkatan signifikan antara skor pretest dan posttest. Rata-rata skor pretest siswa adalah 65, sementara rata-rata skor posttest meningkat menjadi 80. Analisis statistik menggunakan uji t berpasangan menunjukkan adanya perbedaan yang signifikan antara pretest dan posttest ($p\text{-value} < 0,05$), yang mengindikasikan bahwa penggunaan media flipbook dengan model CTL berhasil meningkatkan hasil belajar siswa dalam menulis cerita deskripsi.

Revisi Instrumen dan Media

Berdasarkan hasil validasi dan uji coba, beberapa butir soal dalam instrumen evaluasi yang memiliki nilai Aiken's V rendah (0,68 dan 0,72) perlu direvisi untuk meningkatkan validitasnya. Revisi tersebut mencakup penyempurnaan rumusan soal agar lebih spesifik dan relevan dengan materi yang diajarkan. Selain itu, media buku cerita bergambar flipbook juga akan disempurnakan berdasarkan masukan dari siswa dan guru untuk meningkatkan kualitas visual dan konten yang ada dalam buku cerita tersebut.

Berikut adalah hasil dan pembahasan serta tabel hasil perhitungan validasi butir instrumen evaluasi menggunakan formula Aiken dengan 4 data.

Setelah proses validasi oleh 4 ahli dilakukan pada 25 butir soal, perhitungan skor Aiken dilakukan untuk mengevaluasi relevansi masing-masing butir soal. Berdasarkan skor Aiken

yang dihitung, berikut adalah hasil yang ditemukan:

1. Skor Aiken yang diperoleh menunjukkan bahwa sebagian besar butir soal memiliki tingkat relevansi yang tinggi dengan materi yang diajarkan. Sebagian besar butir soal memperoleh skor antara 3 dan 4, yang menunjukkan bahwa soal-soal tersebut sangat relevan atau relevan untuk mengukur kompetensi menulis cerita deskripsi siswa.
2. Butir soal dengan skor rendah (skor 2 atau 1) menunjukkan bahwa soal tersebut kurang relevan atau tidak cukup efektif dalam mengukur kompetensi yang diinginkan. Beberapa butir soal ini berkaitan dengan media yang digunakan dalam pembelajaran (misalnya buku cerita bergambar flipbook), dan perbaikan pada pertanyaan ini dapat dilakukan untuk meningkatkan relevansi dan efektivitas soal.
3. Distribusi skor Aiken menunjukkan bahwa sebagian besar ahli memberikan penilaian yang serupa terhadap relevansi soal-soal tersebut, yang menunjukkan konsistensi dalam penilaian.
4. Kesimpulan: Berdasarkan hasil perhitungan dan pembahasan, instrumen evaluasi ini dapat digunakan untuk mengukur keterampilan menulis cerita deskripsi pada siswa kelas IV SD, dengan beberapa perbaikan pada butir soal yang mendapat skor rendah untuk meningkatkan kesesuaian dan efektivitas soal dalam mengukur kompetensi yang diinginkan.

Tabel berikut menunjukkan hasil perhitungan validasi dari 25 butir soal dengan 4 data (validasi oleh 4 ahli). Skor Aiken dihitung berdasarkan penilaian ahli.

Tabel III Hasil Perhitungan Validasi Butir Instrumen Evaluasi Menggunakan Formula Aiken

No. Butir	Pakar				S1	S2	S3	S3	ΣS	V	Kategori	Berubah Menjadi
	1	2	3	4								
1	4	5	4	4	3	4	3	3	13	0,813	Sangat Tinggi	Butir 1
2	4	4	4	4	3	3	3	3	12	0,750	Sangat Tinggi	Butir 2

3	5	5	4	4	4	4	3	3	14	0,875	Sangat Tinggi	Butir 3
4	5	5	4	5	4	4	3	4	15	0,938	Sangat Tinggi	Butir 4
5	4	5	5	4	3	4	4	3	14	0,875	Sangat Tinggi	Butir 5
6	4	5	4	5	3	4	3	4	14	0,875	Sangat Tinggi	Butir 6
7	4	5	4	4	3	4	3	3	13	0,813	Sangat Tinggi	Butir 7
8	5	4	5	4	4	3	4	3	14	0,875	Sangat Tinggi	Butir 8
9	4	4	4	5	3	3	3	4	13	0,813	Sangat Tinggi	Butir 9
10	2	2	2	3	1	1	1	2	5	0,313	Rendah	Drop
11	5	4	5	4	4	3	4	3	14	0,875	Sangat Tinggi	Butir 10
12	3	4	5	4	2	3	4	3	12	0,750	Sangat Tinggi	Butir 11
13	4	5	5	4	3	4	4	3	14	0,875	Sangat Tinggi	Butir 12
14	3	4	4	4	2	3	3	3	11	0,688	Sangat Tinggi	Butir 13
15	4	4	3	4	3	3	2	3	11	0,688	Sangat Tinggi	Butir 14
16	4	5	4	5	3	4	3	4	14	0,875	Sangat Tinggi	Butir 15
17	4	5	5	4	3	4	4	3	14	0,875	Sangat Tinggi	Butir 16

18	3	4	4	4	3	3	2	3	11	0,688	Sangat Tinggi	Butir 17
19	4	5	4	4	3	4	3	3	13	0,813	Sangat Tinggi	Butir 18
20	4	5	4	5	3	4	3	4	14	0,875	Sangat Tinggi	Butir 19
21	4	5	4	4	3	4	3	3	13	0,813	Sangat Tinggi	Butir 20
22	3	2	2	3	2	1	1	2	6	0,373	Sangat Rendah	Drop
23	4	5	4	3	3	4	3	2	12	0,750	Sangat Tinggi	Butir 21
24	5	4	4	4	4	3	3	3	13	0,813	Sangat Tinggi	Butir 22
25	3	4	4	4	2	3	3	3	11	0,688	Sangat Tinggi	Butir 23

Untuk memperoleh nilai S1, S2, S3, dan S4 yang hasilnya ditunjukkan pada Tabel III di atas diperoleh menggunakan formula :

$$S = r - l_0 \quad (1)$$

Dimana:

r = skor rating kepentingan yang dipilih oleh pakar

l_0 = skor terendah dari semua skor rating kepentingan

Berdasarkan formula S tersebut, maka S1 untuk butir ke-1 = $4 - 1 = 3$, S2 untuk butir ke-1 = $5 - 1 = 4$, S3 untuk butir ke-1 = $4 - 1 = 3$, dan S4 untuk butir ke-1 = $4 - 1 = 3$. Penentuan S1, S2,

S3, dan S4 tersebut dapat dilanjutkan sampai dengan butir ke-25 dengan langkah perhitungan yang sama.

Untuk memperoleh nilai V yang hasilnya telah ditunjukkan pada Tabel III di atas

diperoleh menggunakan formula :

$$V = \frac{\Sigma S}{n(c - 1)}$$

Dimana:

ΣS = Total keseluruhan dari selisih antara skor yang diberikan oleh pakar dengan skor terendah rating kepentingan

n = Jumlah pakar yang terlibat

c = Jumlah pilihan skor rating kepentingan

Berdasarkan formula V tersebut, maka nilai V untuk butir ke-1 dapat dihitung sebagai berikut:

$$V = \frac{\Sigma S}{n(c - 1)}$$

$$V = \frac{13}{4(5 - 1)} - \frac{13}{4 \times 4} = \frac{13}{16} = 0,813$$

Demikian seterusnya dengan langkah yang sama untuk memperoleh nilai V pada tiap butir instrumen selanjutnya.

Penentuan kategori tiap butir instrumen yang ditunjukkan pada Tabel III diatas dilakukan dengan membandingkan nilai V tiap butir instrumen dengan rentangan pengkategorian validitas instrumen yang ditunjukkan pada Tabel I. Kategori sangat tinggi akan muncul jika nilai V berada pada rentang skor 0,8-1,000. Kategori tinggi jika nilai V berada pada rentang skor 0,6-0,799. Kategori cukup jika nilai V berada pada rentang skor 0,4-0,599. Kategori rendah jika nilai V berada pada rentang skor 0,2-0,399. Kategori sangat rendah jika nilai V berada pada rentang di bawah 0,200.

Pengembangan Media Buku Cerita Bergambar Flipbook dengan Model Pembelajaran CTL (Contextual Teaching and Learning) untuk Meningkatkan Hasil Belajar pada Materi Menulis Cerita Deskripsi pada Mata Pelajaran Bahasa Indonesiaa dalah butir yang termasuk dalam kategori minimal yaitu cukup, sedangkan jika tidak memenuhi kategori minimal tersebut maka butir dibuang.

Berdasarkan ketentuan tersebut, maka butir-butir instrument yang masih digunakan adalah butir ke-1 dengan penomoran butir tetap menjadi butir ke-1, butir ke-2 dengan

penomoran butir tetap menjadi butir ke-2, butir ke-3 dengan penomoran butir tetap menjadi butir ke-3, butir ke-4 dengan penomoran butir tetap menjadi butir ke-4, butir ke-5 dengan penomoran butir tetap menjadi butir ke-5, butir ke-6 dengan penomoran butir tetap menjadi butir ke-6, butir ke-7 dengan penomoran

butir tetap menjadi butir ke-7, butir ke-8 dengan penomoran butir tetap menjadi butir ke-8, butir ke-9 dengan penomoran butir tetap menjadi butir ke-9, butir ke-11 dengan penomoran butir berubah menjadi butir ke-10, butir ke-12 dengan penomoran butir berubah menjadi butir ke-11, butir ke-13 dengan penomoran butir berubah menjadi butir ke-12, butir ke-14 dengan penomoran butir berubah menjadi butir ke-13, butir ke-15 dengan penomoran butir berubah menjadi butir ke-14, butir ke-16 dengan penomoran butir berubah menjadi butir ke-15, butir ke-17 dengan penomoran butir berubah menjadi butir ke-16, butir ke-18 dengan penomoran butir berubah menjadi butir ke-17, butir ke-19 dengan penomoran butir berubah menjadi butir ke-18, butir ke-20 dengan penomoran butir berubah menjadi butir ke-19, butir ke-21 dengan penomoran butir berubah menjadi butir ke-20, butir ke-23 dengan penomoran butir berubah menjadi butir ke-21, butir ke-24 dengan penomoran butir berubah menjadi butir ke-22, butir ke-25 dengan penomoran butir berubah menjadi butir ke-23.

Berdasarkan hasil yang diperoleh dalam penelitian ini, maka kendala yang ditemukan dalam penelitian yang pernah dilakukan oleh Sugiharni pada tahun 2017 telah mampu dijawab melalui hasil perhitungan validitas butir instrument menggunakan formula *Alkin* dengan melibatkan 4 pakar untuk melakukan uji validasi butir instrumen. Kendatipun secara umum hasil penelitian ini telah mampu memberikan solusi terhadap permasalahan yang ditemukan dalam penelitian lainnya, namun masih terdapat kendala yang ditemukan dalam penelitian ini yaitu keraguan untuk mengambil kepastian penggunaan butir instrumen yang validitasnya masih berada pada kategori “cukup”.

Pembahasan:

1. Skor Aiken Rata-rata: Sebagian besar butir soal memiliki rata-rata skor Aiken 4, yang berarti soal-soal ini sangat relevan dalam mengukur kompetensi yang diinginkan. Soal-soal ini sudah sangat sesuai untuk digunakan dalam evaluasi hasil belajar siswa dalam menulis cerita deskripsi.
2. Skor Rendah: Beberapa soal, seperti soal nomor 10 dan beberapa soal lainnya yang

relevansinya lebih rendah (skor rata-rata 2), perlu direvisi untuk meningkatkan kesesuaiannya dalam mengukur kemampuan siswa. Mungkin diperlukan perubahan dalam konteks soal atau menyesuaikan dengan materi pembelajaran.

3. **Konsistensi Penilaian:** Hasil perhitungan menunjukkan konsistensi yang baik antara penilaian ahli dalam hal relevansi soal. Hal ini menunjukkan bahwa instrumen yang digunakan sudah cukup solid dalam mengukur kompetensi yang diing

KESIMPULAN

Berdasarkan hasil perhitungan validasi butir instrumen evaluasi menggunakan metode Aiken pada 25 butir soal yang diajukan, dapat disimpulkan bahwa:

1. **Relevansi Soal:** Sebagian besar butir soal memiliki relevansi yang sangat baik dalam mengukur kompetensi menulis cerita deskripsi pada siswa kelas IV SD. Skor Aiken yang diperoleh menunjukkan bahwa mayoritas soal sangat relevan (skor 4) atau relevan (skor 3) dalam mengukur keterampilan menulis, penggunaan bahasa yang tepat, kreativitas, serta pengorganisasian ide dalam cerita deskripsi.
2. **Perbaikan Soal dengan Skor Rendah:** Beberapa soal, terutama soal yang terkait dengan penggunaan media buku cerita bergambar flipbook, memperoleh skor yang lebih rendah (skor 2). Hal ini menunjukkan bahwa ada beberapa aspek yang perlu diperbaiki untuk meningkatkan relevansi soal terhadap tujuan pembelajaran dan materi yang diajarkan. Perbaikan dapat mencakup penyesuaian soal agar lebih sesuai dengan konteks pembelajaran yang dijalankan.
3. **Konsistensi Penilaian:** Penilaian dari keempat ahli menunjukkan tingkat konsistensi yang tinggi dalam menilai relevansi soal-soal tersebut. Ini menunjukkan bahwa instrumen yang digunakan sudah cukup solid dan dapat diandalkan untuk menilai keterampilan menulis cerita deskripsi siswa.
4. **Instrumen yang Valid:** Instrumen evaluasi yang telah divalidasi menggunakan metode Aiken ini dapat digunakan untuk mengukur kompetensi siswa dalam menulis cerita deskripsi, dengan catatan beberapa soal perlu direvisi agar lebih relevan dan efektif dalam mengukur keterampilan yang diinginkan.

Secara keseluruhan, instrumen evaluasi ini memiliki validitas yang baik dalam mengukur kompetensi menulis cerita deskripsi, dan hanya membutuhkan beberapa perbaikan pada butir

soal tertentu untuk mencapai kesempurnaan.

Berdasarkan hasil dan pembahasan, maka dapat ditarik kesimpulan bahwa pada mulanya terdapat 25 instrumen evaluasi model *Alkin* yang digunakan untuk mengevaluasi pengembangan media buku cerita bergambar flipbook dengan model pembelajaran CTL (Contextual Teaching and Learning) untuk meningkatkan hasil belajar pada materi menulis cerita deskripsi pada mata pelajaran bahasa Indonesia siswa kelas IV SDN Jomblang 05. Setelah dilakukan uji validitas menggunakan formula *Aiken*, maka diperoleh 23 butir yang siap digunakan dan diperoleh 2 butir yang harus dibuang karena nilai validitasnya rendah. Adapun hal yang dapat dilakukan untuk mengatasi kendala dalam penelitian ini adalah melakukan kombinasi penggunaan formula *Aiken* dengan metode backward chaining yang melakukan penelusuran terhadap aspek yang menjadi penyebab mengapa butir tersebut berada pada kategori cukup.

DAFTAR PUSTAKA

- Aiken, L. R. (1985). *Three Coefficients for Analyzing the Reliability and Validity of Ratings*. Educational and Psychological Measurement, 45(1), 131–142.
- Arikunto, S. (2016). *Prosedur Penelitian: Suatu Pendekatan Praktik*. Jakarta: Rineka Cipta.
- Borg, W. R., & Gall, M. D. (1983). *Educational Research: An Introduction*. New York: Longman.
- Creswell, J. W. (2012). *Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research*. Boston: Pearson.
- Daryanto. (2016). *Media Pembelajaran Peranannya Sangat Penting dalam Mencapai Tujuan Pembelajaran*. Yogyakarta: Gava Media.
- Kemendikbud. (2016). *Standar Kompetensi dan Kompetensi Dasar Bahasa Indonesia SD/MI*. Jakarta: Kemendikbud.
- Nurgiyantoro, B. (2013). *Penilaian dalam Pengajaran Bahasa dan Sastra*. Yogyakarta: BPFE.
- Peraturan Menteri Pendidikan dan Kebudayaan Nomor 20 Tahun 2016 tentang Standar Kompetensi Lulusan Pendidikan Dasar dan Menengah.
- Purwanto, M. (2010). *Validasi Instrumen Evaluasi Pendidikan*. Jurnal Pendidikan dan Pembelajaran, 17(2), 124–132.
- Setiawan, D., & Hadi, W. (2022). *Model Pembelajaran Kontekstual untuk Meningkatkan*

- Keterampilan Menulis Cerita*. Jurnal Pendidikan Dasar, 10(3), 45–57.
- Sugiyono. (2018). *Metode Penelitian Kuantitatif, Kualitatif, dan R&D*. Bandung: Alfabeta.
- Widodo, A. (2014). *Strategi Pembelajaran Menulis pada Pendidikan Dasar*. Jakarta: Gramedia.
- Zaini, H., Munthe, B., & Aryani, S. (2016). *Strategi Pembelajaran Aktif*. Yogyakarta: Pustaka Insan Madani.
- Prasetyo, T., & Jannah, M. (2011). *Metode Penelitian Kuantitatif*. Jakarta: Rajawali Pers.
- Harmer, J. (2004). *How to Teach Writing*. Essex: Pearson Education Limited.
- Lestari, P. (2020). *Pengembangan Media Pembelajaran Interaktif Berbasis Flipbook*. Jurnal Teknologi Pendidikan, 8(2), 45–56.
- Mayer, R. E. (2009). *Multimedia Learning*. New York: Cambridge University Press.
- Robinson, P., & Ellis, N. C. (Eds.). (2008). *Handbook of Cognitive Linguistics and Second Language Acquisition*. New York: Routledge.