

**KESESUAIAN INSTRUMEN EVALUASI BAHASA JEPANG DENGAN KAIDAH
PENYUSUNAN SOAL PILIHAN GANDA**

Tri Rahayu Ning Megasari

S-1 Pendidikan Bahasa Jepang, Fakultas Bahasa dan Seni, Universitas Negeri Surabaya

tri.19045@mhs.unesa.ac.id

Amira Agustin Kocimaheni, S. Pd., M. Pd.

S-1 Pendidikan Bahasa Jepang, Fakultas Bahasa dan Seni, Universitas Negeri Surabaya

amiraagustin@unesa.ac.id

Abstract

This study aims to describe the level of suitability between Japanese language evaluation instruments in two high schools in the Sidoarjo region with the rules for preparing multiple choice questions in the aspects of material and construction. The research design used is descriptive qualitative with data collection techniques in the form of documentary studies, observation and interviews. The research subjects in this study were Japanese language educators in two high schools in Sidoarjo who had taught Japanese for >5 years. Data analysis activities using the Miles and Huberman model consisting of data reduction, data presentation and conclusion drawing. The results showed that: 1) the level of suitability between Japanese language evaluation instruments in SMA A with the rules for preparing multiple choice questions in the material aspect is quite good with a percentage of 74.5%, while the level of suitability between Japanese language evaluation instruments in SMA B with the rules for preparing multiple choice questions in the material aspect is good with a percentage of 82.5%; 2) the level of suitability between Japanese language evaluation instruments in SMA A with the rules for preparing multiple choice questions in the construction aspect is very good with a percentage of 95%, while the level of suitability between Japanese language evaluation instruments in SMA B with the rules for preparing multiple choice questions in the construction aspect is very good with a percentage of 95.3%.

Keywords: suitability, evaluation instruments, rules of preparation, multiple choice questions

要旨

この研究の目的は、シドアルジョ地域の2つの高等学校における日本語の評価方法と多肢選択問題の作成ルールとの適合度を、教材と構成という側面から明らかにすることである。使用した研究デザインは記述的質的研究であり、データ収集手法は文書調査、観察、インタビューである。この研究の対象者は、シドアルジョの2つの高校で5年以上日本語を教えている日本語教育者である。データ分析には、マイルズ&ヒューバーマンモデルを用い、データの削減、データの提示、結論の導出から構成されている。その結果、以下のことが明らかになった： 1) A高校の日本語評価方法と教材面における多肢選択問題作成ルールの適合度は74.5%と非常に良好であり、B高校の日本語評価方法と教材面における多肢選択問題作成ルールの適合度は82.5%と良好であった。 2) A高校の日本語評価方法と構成面における多肢選択問題作成ルールの適合度は95%と非常に良好であり、B高校の日本語評価方法と構成面における多肢選択問題作成ルールの適合度は95.3%と非常に良好であった。

キーワード: 適合性、評価方法、作成ルール、多肢選択問題

PENDAHULUAN

Evaluasi merupakan salah satu komponen penting dalam pembelajaran yang harus dilakukan oleh setiap pendidik untuk mengetahui tingkat keberhasilan suatu proses pembelajaran mengacu pada ketercapaian tujuan pembelajaran (Arifin, 2011:6). Melalui evaluasi pembelajaran pendidik dapat mengidentifikasi faktor-faktor yang menjadi penyebab keberhasilan dan ketidak

berhasilan peserta didik selama pembelajaran sehingga dapat ditemukan langkah yang tepat sebagai perbaikan untuk pembelajaran berikutnya (Sudijono, 2009:16). Dalam pelaksanaannya, proses evaluasi media yang dapat mempermudah pendidik memperoleh data atau mengumpulkan informasi yakni alat atau instrumen evaluasi. Menurut pendapat Arikunto (2013:40), instrumen evaluasi merupakan suatu alat yang digunakan

untuk melakukan kegiatan pengukuran terkait kemampuan dan pengetahuan peserta didik yang akan dievaluasi.

Jenis instrumen yang seringkali digunakan oleh pendidik adalah instrumen yang berbentuk tes khususnya tes objektif dengan bentuk soal pilihan ganda, karena dinilai efektif untuk tingkat kemampuan peserta didik dalam menguasai cakupan pengetahuan yang lebih kompleks dengan tingkatan yang lebih tinggi. Selain itu, proses pengoreksian dan penilaian pada soal pilihan ganda lebih mudah dan tidak membutuhkan waktu lama karena pendidik hanya perlu mencocokkan jawaban peserta didik dengan kunci jawaban yang telah disusun. Namun, dalam menyusun soal pilihan ganda dibutuhkan keterampilan dan ketelitian untuk menuliskan pengecoh pada pilihan jawaban. Selain itu, pilihan jawaban harus logis dan homogen agar pengecoh dapat berfungsi dengan baik, karena pengecoh yang berfungsi dengan baik menjadi salah satu indikator soal dikatakan berkualitas dari segi kuantitatif (Widiyanto, 136-138).

Tes yang berkualitas sangat dibutuhkan untuk memperoleh hasil pengukuran yang valid terkait dengan hasil belajar peserta didik. Sebuah tes dikatakan valid apabila tes tersebut telah memenuhi kriteria tes yang baik, dalam artian mampu mengukur apa yang ingin diukur. Terkait dengan hal tersebut, penyusunan soal tes harus didasarkan pada kompetensi dasar dan indikator yang telah disusun sebelumnya (Nurgiyantoro, 2013:149-155). Selain itu, pendidik juga harus memperhatikan aturan-aturan penyusunan dan penulisan soal tes atau yang disebut dengan kaidah yang meliputi tiga aspek yakni aspek materi, aspek konstruksi dan aspek bahasa (Pusat Penilaian Pendidikan, 2019). Kaidah penyusunan soal memegang peranan yang cukup penting yakni untuk menjaga kualitas soal dan memastikan agar soal mampu menjalankan fungsi pengukurannya dengan baik. Maksudnya adalah suatu soal harus mampu mengukur kemampuan peserta didik dengan sebenar-benarnya (Hayati & Mardapi, 2014:27).

Berdasarkan paparan tersebut, dapat diketahui bahwa penting untuk melakukan analisis untuk memastikan kualitas soal, namun realita yang ada di lapangan tidak seperti yang diharapkan. Seperti pada penelitian oleh Rahmawati (2019) yang menemukan bahwa masih terdapat ketidaksesuaian antara soal Penilaian Tengah Semester (PTS) IPA yang disusun oleh pendidik dengan kaidah penyusunan soal aspek bahasa. Menurut pendapat Nurgiyantoro (2013:108), soal yang disusun secara mandiri oleh pendidik seringkali digunakan tanpa dilakukan ujicoba terlebih dahulu karena berbagai faktor. Pendapat tersebut selaras dengan apa yang ditemukan di lapangan. Dari hasil observasi pra-penelitian

dan wawancara dengan pendidik bahasa Jepang di dua SMA di wilayah Sidoarjo, soal-soal yang digunakan sebagai instrumen evaluasi disusun secara mandiri oleh pendidik. Selain itu, soal yang digunakan cenderung tidak diujicobakan dan ditelaah terlebih dahulu sebelum resmi diujikan, sehingga belum diketahui tingkat kesesuaian soal dengan kaidah penyusunan soal. Berdasarkan hal tersebut, peneliti merasa perlu melakukan penelitian ini untuk mengetahui seberapa besar tingkat kesesuaian antara soal yang disusun pendidik dengan kaidah penyusunan soal pilihan ganda aspek materi dan konstruksi. Kesesuaian antara instrumen evaluasi bahasa Jepang dengan aspek bahasa tidak dilakukan karena soal yang digunakan tidak sepenuhnya ditulis dalam bahasa Indonesia tetapi juga dalam bahasa Jepang (menggunakan huruf Jepang).

METODE

Berdasarkan jenis data dan cara analisisnya, penelitian ini merupakan penelitian deskriptif kualitatif. Data yang dikaji dalam penelitian ini adalah data kualitatif yang diperoleh dari sumber data yang meliputi: 1) soal tes sumatif bahasa Jepang; 2) kunci jawaban tes sumatif; 3) kisi-kisi soal tes sumatif; 4) Rencana Pelaksanaan Pembelajaran (RPP) yang disusun oleh pendidik; 5) silabus bahasa Jepang yang digunakan di SMA A dan SMA B untuk tahun ajaran 2022/2023. Subjek penelitian dalam penelitian ini adalah narasumber 1 (N1) yang merupakan pendidik bahasa Jepang di SMA A dan narasumber 2 (N2) yang merupakan pendidik bahasa Jepang di SMA B. SMA A merupakan sekolah menengah atas yang berlokasi di wilayah kabupaten Sidoarjo bagian barat, sedangkan SMA B merupakan sekolah menengah atas yang berlokasi di wilayah kabupaten Sidoarjo bagian utara. N1 dan N2 berusia >25 tahun dan telah mengajar bahasa Jepang selama >5 tahun.

Dalam suatu penelitian, teknik pengumpulan data merupakan tahapan yang paling utama, karena tujuan dari dilakukannya penelitian adalah untuk mendapatkan data. Seorang peneliti harus mampu mengetahui teknik pengumpulan data yang akan digunakan, sehingga akan diperoleh data yang sesuai dengan kriteria yang telah ditetapkan (Sugiyono, 2016:308). Teknik pengumpulan data yang digunakan dalam penelitian ini adalah metode studi dokumenter, observasi dan wawancara semi terstruktur

Menurut pendapat Nasution (dalam Sugiyono, 2016:307), dalam penelitian kualitatif peneliti berkedudukan sebagai instrumen penelitian utama, karena hal-hal yang berkaitan dengan masalah maupun hasil penelitian tidak dapat ditentukan dengan pasti dan akan berkembang seiring berjalannya waktu. Seiring teori tersebut, instrumen penelitian dalam penelitian ini adalah peneliti sendiri dengan dibantu oleh instrumen pendukung

yang berupa: 1) *checklist* perangkat pembelajaran; 2) lembar observasi; dan 3) pedoman wawancara. Selain itu, selama proses penelitian dibutuhkan alat bantu tambahan yang meliputi alat tulis dan *handphone*/alat perekam untuk wawancara, serta aplikasi Anates V4 yang digunakan untuk menganalisis efektivitas pengecoh.

Proses analisis data dilakukan dengan menggunakan model Miles & Huberman yang meliputi tahapan reduksi data, penyajian data dan penarikan kesimpulan.

1. Data yang diperoleh dari *checklist* dan wawancara direduksi dan diobservasi dengan mengacu pada kriteria aspek materi dan konstruksi pada kaidah penyusunan soal pilihan ganda sebagaimana dalam tabel 1.

Tabel 1 Kriteria Aspek Materi dan Konstruksi Kaidah Penyusunan Soal Pilihan Ganda

Aspek	Kriteria
Materi	Butir soal sesuai dengan indikator
	Isi materi benar secara keilmuan
	Hanya ada satu kunci jawaban benar
	Isi materi sesuai dengan kelas/jenjang pendidikan
	Butir pengecoh berfungsi dengan baik
Konstruksi	Pokok soal dirumuskan dengan jelas
	Pokok soal tidak mengarah ke jawaban benar
	Pilihan jawaban dirumuskan dengan jelas
	Pilihan jawaban homogen
	Tidak ada bentuk negatif ganda
	Panjang pilihan jawaban kurang lebih sama
	Antarbutir soal tidak bergantung satu sama lain
	Pilihan dalam bentuk angka/waktu diurutkan

Adaptasi dari: Nurgiyantoro (2013)

2. Penyajian Data

Data hasil observasi terkait kesesuaian dan ketidaksesuaian diubah ke dalam bentuk persentase dan disajikan dalam tabulasi data. Data hasil analisis disajikan dalam bentuk narasi dan dikomparasikan dengan hasil wawancara

3. Penarikan kesimpulan

Membuat kesimpulan mengacu pada hasil analisis terkait tingkat kesesuaian antara soal tes sumatif bahasa Jepang dengan kaidah penyusunan soal pilihan ganda aspek materi dan konstruksi.

HASIL DAN PEMBAHASAN

Dari hasil *checklist*, diperoleh data terkait perangkat pembelajaran yang disusun oleh pendidik bahasa Jepang di SMA A dan SMA B untuk tahun ajaran 2022/2023. Dapat diketahui bahwa pendidik di SMA A menyusun perangkat pembelajaran dengan lengkap, sedangkan pendidik di SMA B tidak menyusun kisi-kisi soal tes sumatif tahun ajaran 2022/2023 karena keterbatasan waktu. Data kemudian direduksi untuk memilah data penting yang sesuai dengan fokus penelitian. Dalam hal ini, data yang direduksi yakni soal tes sumatif bahasa Jepang tahun 2022/2023 SMA B yang terdiri dari 16 butir soal pilihan ganda, 5 butir soal pilihan ganda kompleks, 9 butir soal benar-salah, 2 butir soal menjodohkan, dan 3 butir soal uraian singkat. Data direduksi dengan cara memilih 16 butir soal pilihan ganda untuk dianalisis sesuai dengan fokus penelitian.

Dari kegiatan wawancara dengan menggunakan instrumen pedoman wawancara diperoleh data berupa informasi terkait penyusunan soal tes yang dilakukan oleh pendidik bahasa Jepang di SMA A dan SMA B. Data hasil wawancara kemudian direduksi dan dituliskan dalam bentuk transkrip hasil wawancara

Dari kegiatan observasi terhadap soal dengan menggunakan instrumen lembar observasi, diperoleh informasi mengenai tingkat kesesuaian dan ketidaksesuaian antara soal sebagai instrumen evaluasi bahasa Jepang yang disusun oleh pendidik di SMA A dan SMA B dengan kaidah penyusunan soal pilihan ganda pada aspek materi dan aspek konstruksi. Rincian mengenai hasil observasi dijelaskan dalam poin-poin berikut:

1. Kesesuaian Instrumen Evaluasi Bahasa Jepang SMA A dan SMA B dengan Aspek Materi

Menurut pendapat Nuriyantoro (2013:42-47), penyusunan soal tes harus didasarkan pada indikator dan juga materi pada bahan ajar, karena pada dasarnya indikator dan bahan ajar disusun dengan mengacu pada rumusan kompetensi dasar dalam kurikulum. Namun, dari hasil observasi diketahui bahwa terdapat ketidaksesuaian antara soal tes sumatif bahasa Jepang SMA A dan SMA B dengan kriteria-kriteria yang ada pada aspek materi

Tabel 2 Persentase Ketidaksesuaian antara Soal Tes Sumatif SMA A dengan Kaidah Penyusunan Soal Pilihan Ganda Aspek Materi

Kriteria	Butir Soal	Jumlah Butir Soal	Persentase (%)
Butir soal sesuai dengan indikator yang ada pada kisi-kisi dan RPP	8, 12, 21, 23, 25, 26, 31, 32, 37, 38	10	25%
Isi materi benar secara keilmuan bahasa Jepang	-	0	0%
Hanya ada satu kunci jawaban	31	1	2,50%
Isi materi sesuai dengan kelas/jenjang pendidikan	-	0	0%
Butir pengecoh berfungsi dengan baik	1 s/d 40	40	100%
Total Persentase Ketidaksesuaian			127,50%
Rata-rata Persentase Ketidaksesuaian			25,50%

Mengacu pada tabel 2, ketidaksesuaian antara soal tes sumatif bahasa Jepang yang digunakan di SMA A dengan aspek materi terletak pada kriteria pertama dengan

persentase sebesar 25%, kriteria ketiga sebesar 2,5% dan kriteria kelima sebesar 100%.

Tabel 3 Persentase Ketidaksesuaian antara Soal Tes Sumatif SMA B dengan Kaidah Penyusunan Soal Pilihan Ganda Aspek Materi

Kriteria	Butir Soal	Jumlah Butir Soal	Persentase (%)
Butir soal sesuai dengan indikator yang ada pada kisi-kisi dan RPP	11, 12, 15	3	18,75%
Isi materi benar secara keilmuan bahasa Jepang	-	0	0%
Hanya ada satu kunci jawaban	-	0	0%
Isi materi sesuai dengan kelas/jenjang pendidikan	-	0	0%
Butir pengecoh berfungsi dengan baik	1, 2, 4, 5, 7, 9, 10, 11, 14, 15, 16	11	68,75%
Total Persentase Ketidaksesuaian			87,50%
Rata-rata Persentase Ketidaksesuaian			17,50%

Mengacu pada tabel 3, ketidaksesuaian antara soal tes sumatif bahasa Jepang yang digunakan di SMA B dengan aspek materi terletak pada kriteria pertama dengan persentase sebesar 18,75% dan kriteria kelima sebesar 68,75%.

Kriteria pertama: butir soal sesuai dengan indikator pada kisi-kisi dan RPP

Mengacu pada tabel 2 dan 3, sebanyak 10 butir soal tes sumatif bahasa Jepang SMA A dan 3 butir soal tes sumatif bahasa Jepang SMA B terbukti tidak sesuai dengan kriteria pertama aspek materi. Ketidaksesuaian soal disebabkan oleh 3 faktor yakni: 1) butir soal tidak sesuai dengan indikator pada kisi-kisi; 2) butir soal tidak sesuai dengan indikator RPP; 3) butir soal tidak mewakili indikator RPP. Contoh beberapa soal yang tidak sesuai adalah soal nomor 26 dan 31 pada tes sumatif SMA A serta soal nomor 12 pada tes sumatif SMA B.

Soal nomor 26 tes sumatif SMA A

26. Pola kalimat ~ o ~ kudasai digunakan untuk ...

- Meminta barang atau pesan makanan
- Menyebutkan barang
- Menanyakan barang
- Menitipkan barang
- Memberitahu sesuatu

Jawaban benar : A

Pada soal nomor 26 tes sumatif SMA A, peserta didik diminta untuk menentukan fungsi yang tepat dari pola kalimat “~ o ~kudasai”, sedangkan pada RPP tidak disebutkan indikator yang sesuai dengan soal, sehingga soal tersebut tidak sesuai dengan kriteria karena tidak mewakili indikator pada RPP.

Soal nomor 31 tes sumatif SMA A

31. Pola kalimat ~ ni ~ to ~ ga arimasu digunakan untuk ...

- a) Menyatakan benda
- b) Menyatakan seluruh benda
- c) Memesan makanan
- d) Menanyakan ada atau tidak benda
- e) Menyatakan adanya seseorang

Jawaban benar : B

Pada soal nomor 31 tes sumatif SMA A, peserta didik diminta untuk menentukan fungsi dari pola kalimat “~ ni ~ to ~ga arimasu”, sedangkan pada kisi-kisi soal tertulis indikator “disajikan pertanyaan langsung, peserta didik dapat menentukan pola kalimat yang sesuai untuk memesan makanan/minuman”. Soal tersebut tidak sesuai dengan kriteria karena tidak sesuai dengan indikator pada kisi-kisi.

Soal nomor 12 tes sumatif SMA B

12. 8じはんから 3じまで がっこうで べんきょうします。

Arti kata yang bergaris bawah adalah...

- A. Sekolah
- B. Rumah
- C. Naik
- D. Teman
- E. Tidur

Jawaban benar : A

Pada soal nomor 12 tes sumatif SMA B, peserta didik diminta untuk menyebutkan arti dari kosakata “がっこう” yang termasuk ke dalam kosakata nama tempat, sedangkan pada indikator peserta didik diminta untuk menyebutkan kosakata jangka waktu/alat transportasi. Soal tersebut tidak sesuai dengan indikator “peserta didik dapat menyebutkan kosakata jangka waktu/alat transportasi dalam bahasa Jepang” pada RPP.

Kriteria kedua: isi materi benar secara keilmuan bahasa Jepang

Soal yang disusun harus memuat materi yang sesuai dengan bahan ajar/buku ajar yang digunakan. dari hasil observasi, keseluruhan soal tes sumatif SMA A dan SMA B telah memuat materi yang sesuai dengan buku ajar.

Kriteria ketiga: hanya ada satu kunci jawaban

Pada soal pilihan ganda hanya ada satu jawaban yang paling tepat yang merupakan kunci jawaban (kecuali pilihan ganda kompleks). Dari hasil observasi, diketahui bahwa terdapat 1 butir soal tes sumatif SMA A yang tidak sesuai dengan kriteria ketiga pada aspek materi, sedangkan soal tes sumatif SMA B telah terbukti sesuai dengan kriteria ketiga pada aspek materi. Soal yang tidak sesuai adalah soal nomor 31.

31. Pola kalimat ~ ni ~ to ~ ga arimasu digunakan untuk ...

- a) Menyatakan benda

- b) Menyatakan seluruh benda
- c) Memesan makanan
- d) Menanyakan ada atau tidak benda
- e) Menyatakan adanya seseorang

Jawaban benar : B

Pada soal nomor 31 di atas, terdapat dua kemungkinan jawaban benar yakni pilihan A dan pilihan B, sehingga soal tersebut tidak sesuai dengan kriteria ketiga pada aspek materi.

Kriteria keempat: isi materi sesuai dengan kelas/jenjang pendidikan

Penyusunan materi ajar didasarkan pada kompetensi dasar yang telah dirumuskan sesuai dengan jenjang pendidikan seperti yang tertera dalam kurikulum. Dari hasil observasi, terbukti bahwa materi dalam keseluruhan soal tes sumatif SMA A dan SMA B telah dirumuskan sesuai dengan kelas/jenjang pendidikan.

Kriteria kelima: butir pengecoh berfungsi dengan baik

Menurut pendapat Nurgiyantoro (2013:200) pada soal pilihan ganda semua pilihan jawaban harus berfungsi dengan efektif. Artinya, meskipun pilihan jawaban tersebut bukan kunci jawaban tetap harus ada sejumlah peserta didik yang memilih. Mengacu pada tabel 2 dan 3, 40 butir soal tes sumatif bahasa Jepang SMA A pengecoh tidak berfungsi dengan baik, sedangkan pada soal tes sumatif bahasa Jepang SMA B terdapat 11 butir soal yang pengecohnya tidak berfungsi dengan baik. Ketidaksesuaian antara soal dengan kriteria kelima disebabkan oleh pilihan jawaban yang tidak sama atau tidak homogen sehingga memungkinkan peserta didik untuk menebak-nebak jawaban yang tepat. Contoh soal yang tidak sesuai dengan kriteria kelima adalah soal nomor 3 pada tes sumatif SMA A.

3. A : (... ?)

B : Senpuuki to hondana ga arimasu.

Kalimat pertanyaan yang cocok dengan jawaban tersebut adalah ...

- a) Hai, arimasu.
- b) Doko desuka.
- c) Nani desuka.
- d) Uchi ni nani ga arimasuka.
- e) Nani ni arimasuka.

Jawaban benar : D

Pada soal nomor 3 tes sumatif SMA A tersebut, pilihan jawaban D lebih panjang dibandingkan pilihan jawaban yang lain dan merupakan kunci jawaban.

Dari paparan sebelumnya dapat diketahui kesesuaian antara soal tes sumatif bahasa Jepang SMA A dan SMA B dengan aspek materi sebagaimana berikut.

Tabel 4 Persentase Kesesuaian antara Soal Tes Sumatif SMA A dengan Kaidah Penyusunan Soal Pilihan Ganda Aspek Materi

Kriteria	Butir Soal	Jumlah Butir Soal	Persentase (%)
Butir soal sesuai dengan indikator yang ada pada kisi-kisi dan RPP	1, 2, 3, 4, 5, 6, 7, 9, 10, 11, 13, 14, 15, 16, 17, 18, 19, 20, 22, 24, 27, 28, 29, 30, 33, 34, 35, 36, 39, 40	30	75%
Isi materi benar secara keilmuan bahasa Jepang	1 s/d 40	40	100%
Hanya ada satu kunci jawaban	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 32, 33, 34, 35, 36, 37, 38, 39, 40.	39	97,50%
Isi materi sesuai dengan kelas/jenjang pendidikan	1 s/d 40	40	100%
Butir pengecoh berfungsi dengan baik	-	0	0%
Total Persentase Kesesuaian			372,50%
Rata-rata Presentase Kesesuaian			74,50%

Dari tabel 4 di atas, dapat diketahui rata-rata kesesuaian soal tes sumatif SMA A dengan aspek materi sebesar 74,5% yang menunjukkan bahwa soal tes sumatif

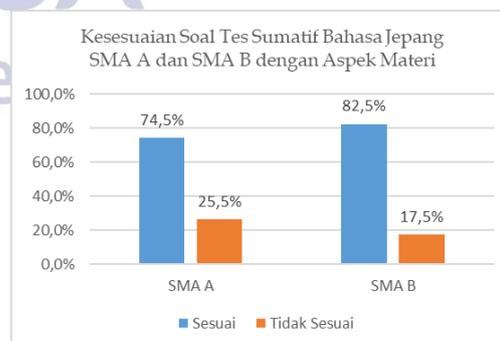
SMA A sudah cukup sesuai dengan kaidah penyusunan soal pilihan ganda aspek materi.

Tabel 5 Persentase Kesesuaian antara Soal Tes Sumatif SMA B dengan Kaidah Penyusunan Soal Pilihan Ganda Aspek Materi

Kriteria	Butir Soal	Jumlah Butir Soal	Persentase (%)
Butir soal sesuai dengan indikator yang ada pada kisi-kisi dan RPP	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 13, 14, 16	13	81,25%
Isi materi benar secara keilmuan bahasa Jepang	1 – 16	16	100%
Hanya ada satu kunci jawaban	1 – 16	16	100%
Isi materi sesuai dengan kelas/jenjang pendidikan	1 – 16	16	100%
Butir pengecoh berfungsi dengan baik	3, 6, 8, 12, 13	5	31,25%
Total Persentase Kesesuaian			412,50%
Rata-rata Presentase Kesesuaian			82,50%

Dari tabel 5 di atas, dapat diketahui rata-rata kesesuaian soal tes sumatif SMA B dengan aspek materi sebesar 82,5% yang menunjukkan bahwa tingkat kesesuaian soal tes sumatif SMA B dengan kaidah penyusunan soal pilihan ganda aspek materi tergolong baik.

Hasil analisis terkait tingkat kesesuaian antara soal tes sumatif bahasa Jepang SMA A dan SMA B dengan kaidah penyusunan soal pilihan ganda aspek materi dapat dilihat dalam diagram berikut:



Gambar 1 Diagram Persentase Kesesuaian Soal Tes Sumatif SMA A dan SMA B dengan Aspek Materi

2. Kesesuaian Instrumen Evaluasi Bahasa Jepang SMA A dan SMA B dengan Aspek Konstruksi

Menurut pendapat Nurgiyantoro (2013:22), sebuah tes dapat dikatakan sebagai tes yang baik apabila telah sesuai dengan tuntutan kriteria tes yang

baik meliputi aspek materi, konstruksi dan bahasa. Dari hasil observasi diketahui bahwa terdapat ketidaksesuaian antara soal tes sumatif bahasa Jepang SMA A dan SMA B dengan kriteria-kriteria yang ada pada aspek konstruksi.

Tabel 6 Persentase Ketidaksesuaian antara Soal Tes Sumatif SMA A dengan Kaidah Penyusunan Soal Pilihan Ganda Aspek Konstruksi

Kriteria	Butir Soal	Jumlah Butir Soal	Persentase (%)
Pokok soal dirumuskan dengan jelas	13, 14, 15, 18, 19, 37	6	15%
Pokok soal tidak mengarah ke jawaban benar	-	0	0%
Pilihan jawaban dirumuskan dengan jelas	15, 31	2	5%
Pilihan jawaban homogen	2, 3, 8	3	7,50%
Tidak ada bentuk negatif ganda	-	0	0%
Panjang pilihan jawaban kurang lebih sama	3, 8, 10, 23, 26	5	12,50%
Antarbutir soal tidak bergantung satu sama lain	-	0	0%
Pilihan dalam bentuk angka/waktu diurutkan	-	0	0%
Total Persentase Ketidaksesuaian			40%
Rata-rata Persentase Ketidaksesuaian			5%

Mengacu pada tabel 6, ketidaksesuaian antara soal tes sumatif bahasa Jepang yang digunakan di SMA A dengan aspek konstruksi terletak pada: kriteria pertama

dengan persentase sebesar 15%; kriteria ketiga sebesar 5%; kriteria keempat sebesar 7,5%; dan kriteria keenam sebesar 12,5%.

Tabel 7 Persentase Ketidaksesuaian antara Soal Tes Sumatif SMA B dengan Kaidah Penyusunan Soal Pilihan Ganda Aspek Konstruksi

Kriteria	Butir Soal	Jumlah Butir Soal	Persentase (%)
Pokok soal dirumuskan dengan jelas	-	0	0%
Pokok soal tidak mengarah ke jawaban benar	-	0	0%
Pilihan jawaban dirumuskan dengan jelas	-	0	0%
Pilihan jawaban homogen	2, 4, 12	3	18,75%
Tidak ada bentuk negatif ganda	-	0	0%
Panjang pilihan jawaban kurang lebih sama	3, 10, 14	3	18,75%
Antarbutir soal tidak bergantung satu sama lain	-	0	0%
Pilihan dalam bentuk angka/waktu diurutkan	-	0	0%
Total Persentase Ketidaksesuaian			37,50%
Rata-rata Presentase Ketidaksesuaian			4,70%

Mengacu pada tabel 7, ketidaksesuaian antara soal tes sumatif bahasa Jepang yang digunakan di SMA B dengan aspek konstruksi terletak pada kriteria keempat dengan persentase sebesar 18,75% dan kriteria keenam sebesar 18,57%.

Kriteria pertama: pokok soal dirumuskan dengan jelas

Mengacu pada tabel 6 dan 7, terdapat 6 butir soal tes sumatif bahasa Jepang SMA A yang terbukti tidak

sesuai dan 16 butir soal tes sumatif SMA B terbukti telah sesuai dengan kriteria pertama pada aspek konstruksi. Ketidaksesuaian soal disebabkan oleh kesalahan penulisan dan instruksi soal yang kurang sesuai. Contoh soal tes sumatif SMA A yang tidak sesuai adalah soal nomor 13 dan 16.

13. やきすき は とても おいしい です。

Gambar yang sesuai dengan kalimat bercetak tebal ialah ...



Jawaban benar : C

Gambar 2 Soal Tes Sumatif SMA A Nomor 13

Pada soal nomor 13, terdapat kesalahan penulisan pada kosakata nama makanan yang bercetak tebal. Penulisan kata “やきすき” yang benar adalah “すきやき”.

Soal nomor 16 tes sumatif SMA A

16. Mie ayam ini **tidak begitu pedas**.

Kata yang bercetak tebal mempunyai arti yaitu ...

- a) Amari karai desu
- b) Amari kara dewa arimasen
- c) Karai dewa arimasen
- d) Karakunai desu
- e) Amari karakunai desu

Jawaban benar : E

Pada soal nomor 16, instruksi “kata yang bercetak tebal mempunyai arti yaitu...” pada pokok soal kurang sesuai dengan stimulus kalimat pernyataan, karena apabila dilihat dari stimulus yang terdapat kata bercetak tebal dalam bahasa Indonesia instruksi yang tepat untuk digunakan adalah menrjemahkan ke dalam bahasa Jepang.

Kriteria kedua: pokok soal tidak mengarah ke jawaban benar

Pokok soal tidak mengandung kata-kata yang memungkinkan untuk merujuk pada kunci jawaban. Mengacu pada tabel 6 dan 7, keseluruhan soal tes sumatif SMA A dan SMA B telah sesuai dengan kriteria kedua aspek konstruksi.

Kriteria ketiga: pilihan jawaban dirumuskan dengan jelas

Mengacu pada tabel 6, terdapat 2 butir soal tes sumatif bahasa Jepang SMA A, yakni soal nomor 15 dan

nomor 31 yang tidak sesuai dengan kriteria ketiga pada aspek materi yang disebabkan oleh kesalahan dalam penulisan pilihan jawaban dan ambiguitas pada pilihan jawaban. Seluruh soal tes sumatif SMA B terbukti telah sesuai dengan kriteria ketiga aspek konstruksi sebagaimana yang tertulis pada tabel 7.

Soal nomor 15 tes sumatif SMA A

15. A : Bandung wa donna tokoro desuka.

B : (Terkenal) tokoro desu.

Untuk menjawab yang rumpang tersebut harusnya diisi dengan ...

- a) Kirei
- b) Yumei na
- c) Suzushii
- d) Yumei
- e) Atsui

Jawaban benar : B

Pada soal nomor 15 tes sumatif SMA A tersebut, terdapat kesalahan dalam penulisan pilihan jawaban yakni pada pilihan jawaban B dan D. Kata “yumei” seharusnya ditulis dengan “yuumei”.

Soal nomor 31 tes sumatif SMA A

31. Pola kalimat ~ ni ~ to ~ ga arimasu digunakan untuk

...

- a) Menyatakan benda
- b) Menyatakan seluruh benda
- c) Memesan makanan
- d) Menanyakan ada atau tidak benda
- e) Menyatakan adanya seseorang

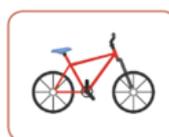
Jawaban benar : B

Pada soal nomor 31 tersebut, pilihan jawaban A dan B dinilai ambigu karena kedua pilihan tersebut merupakan kemungkinan jawaban benar.

Kriteria keempat: pilihan jawaban homogen

Mengacu pada tabel 6 dan 7, terdapat 3 butir soal tes sumatif SMA A dan 4 butir soal tes sumatif SMA B yang tidak sesuai dengan kriteria keempat aspek konstruksi. Salah satu contoh soal yang tidak sesuai adalah soal nomor 4 pada soal tes sumatif SMA B.

4. これは なんですか。



- A. じてんしゃ
- B. でんしゃ
- C. くるま
- D. あるいて
- E. ひとり

Jawaban benar : A

Gambar 3 Soal Tes Sumatif SMA B Nomor 4

Pada soal nomor 4 tes sumatif SMA B diatas, peserta didik diminta untuk menyebutkan kosakata kata benda sesuai gambar. Dari kelima pilihan jawaban, pilihan jawaban D dan E tidak homogen, karena pilihan “D. ある

いて” termasuk kata kerja dan pilihan “E. ひとり” termasuk cara penyebutan jumlah orang.

Kriteria kelima: tidak ada bentuk negatif ganda

Pokok soal tidak mengandung kata-kata yang sifatnya negasi (bukan, kecuali) lebih dari satu. Mengacu pada tabel 6 dan 7, seluruh soal tes sumatif SMA dan SMA B tidak terdapat kata-kata dengan makna negasi ganda, sehingga telah sesuai dengan kriteria kelima aspek konstruksi.

Kriteria keenam: panjang pilihan jawaban kurang lebih sama

Mengacu pada tabel 6 dan 7, terdapat 5 butir soal tes sumatif SMA A dan 3 butir soal tes sumatif SMA B tidak sesuai dengan kriteria keenam pada aspek konstruksi. Salah satu contoh soal yang tidak sesuai adalah soal nomor 23 pada tes sumatif SMA A dan soal nomor 10 pada tes sumatif SMA B.

Soal nomor 23 tes sumatif SMA A

23. Jakarta wa shizuka (tidak).

Bentuk negatif yang mengikuti kata sifat Na adalah

...

- a) Kunai
- b) Dewa ariamasen
- c) Arimasen
- d) Arimasu
- e) Desu

Jawaban benar : B

Pada soal nomor 23 tes sumatif SMA A tersebut, pilihan jawaban B lebih panjang dibandingkan pilihan jawaban yang lain dan merupakan kunci jawaban. Panjang pilihan jawaban yang tidak sama menyebabkan adanya kemungkinan peserta didik menjawab dengan benar dengan cara menebak-nebak dari jawaban yang paling panjang.

10. Pernyataan yang sesuai dengan gambar berikut adalah...



- A. 1じから 2じまで します
- B. あさ 1じに おきます。
- C. 1じに ねます。
- D. 10じから 5じまで ねました
- E. 2じかん べんきょうしませんでした

Jawaban benar : A

Gambar 4 Soal Tes Sumatif SMA B Nomor 10

Pada soal nomor 10 tes sumatif SMA B tersebut, peserta didik diminta untuk memilih pernyataan yang paling tepat sesuai dengan gambar. Pada pilihan jawaban, pilihan E lebih pendek dibandingkan pilihan jawaban yang lain. Hal ini menyebabkan adanya kemungkinan pilihan tersebut tidak dipilih oleh peserta didik karena dianggap bukan jawaban yang tepat.

Kriteria ketujuh: antarbutir soal tidak bergantung satu sama lain

Pokok soal tidak bergantung pada pokok soal/jawaban dari soal sebelumnya. Mengacu pada tabel 6 dan 7, dapat diketahui bahwa baik soal tes sumatif SMA A maupun SMA B terbukti sesuai dengan kriteria ketujuh pada aspek konstruksi. Kesesuaian terletak pada tiap-tiap soal tidak mengandung kata-kata yang merujuk pada pokok soal/jawaban dari soal-soal sebelumnya.

Kriteria kedelapan: pilihan dalam bentuk angka/waktu diurutkan

Mengacu pada tabel 13 dan 14, dapat diketahui bahwa baik soal tes sumatif SMA A maupun SMA B terbukti sesuai dengan kriteria kedelapan pada aspek konstruksi. Pilihan jawaban yang berbentuk angka atau keterangan waktu pada soal telah disusun secara berurutan.

Dari paparan sebelumnya dapat diketahui kesesuaian antara soal tes sumatif bahasa Jepang SMA A dan SMA B dengan aspek konstruksi sebagaimana berikut.

Tabel 8 Persentase Kesesuaian antara Soal Tes Sumatif SMA A dengan Kaidah Penyusunan Soal Pilihan Ganda Aspek Konstruksi

Kriteria	Butir Soal	Jumlah Butir Soal	Persentase (%)
Pokok soal dirumuskan dengan jelas	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 16, 17, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 38, 39, 40	34	85%
Pokok soal tidak mengarah ke jawaban benar	1 – 40	40	100%
Pilihan jawaban dirumuskan dengan jelas	1, 2, 3, 4, 5, 6, 7, 8, 9, 11, 10, 12, 13, 14, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 32, 33, 34, 35, 36, 37, 38, 39, 40	38	95%
Pilihan jawaban homogen	1, 4, 5, 6, 7, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40	37	92,50%
Tidak ada bentuk negatif ganda	1 – 40	40	100%
Panjang pilihan jawaban kurang lebih sama	1, 2, 4, 5, 6, 7, 9, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 24, 25, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40	35	87,50%
Antarbutir soal tidak bergantung satu sama lain	1 – 40	40	100%
Pilihan dalam bentuk angka/waktu diurutkan	1 – 40	40	100%
Total Persentase Kesesuaian			760%
Rata-rata Presentase Kesesuaian			95%

Dari tabel 8 di atas, dapat diketahui rata-rata kesesuaian soal tes sumatif SMA A dengan aspek konstruksi sebesar 95% yang tergolong sangat baik dan menunjukkan bahwa soal tes sumatif SMA A sudah sesuai

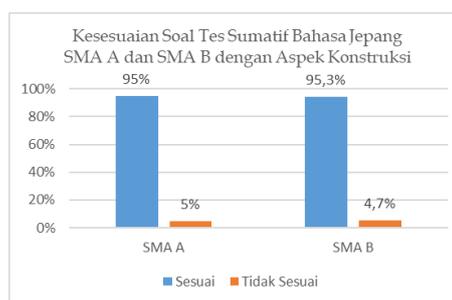
dengan kaidah penyusunan soal pilihan ganda aspek konstruksi.

Tabel 9 Persentase Kesesuaian antara Soal Tes Sumatif SMA B dengan Kaidah Penyusunan Soal Pilihan Ganda Aspek Konstruksi

Kriteria	Butir Soal	Jumlah Butir Soal	Persentase (%)
Pokok soal dirumuskan dengan jelas	1 – 16	16	100%
Pokok soal tidak mengarah ke jawaban benar	1 – 16	16	100%
Pilihan jawaban dirumuskan dengan jelas	1 – 16	16	100%
Pilihan jawaban homogen	1, 3, 5, 6, 7, 8, 9, 10, 11, 13, 14, 15, 16	13	81,25%
Tidak ada bentuk negatif ganda	1 – 16	16	100%
Panjang pilihan jawaban kurang lebih sama	1, 2, 4, 5, 6, 7, 8, 9, 11, 12, 13, 15, 16	13	81,25%
Antarbutir soal tidak bergantung satu sama lain	1 – 16	16	100%
Pilihan dalam bentuk angka/waktu diurutkan	1 – 16	16	100%
Total Persentase Kesesuaian			762,50%
Rata-rata Presentase Kesesuaian			95,30%

Dari tabel 9 di atas, dapat diketahui rata-rata kesesuaian soal tes sumatif SMA B dengan aspek konstruksi sebesar 95,3% yang tergolong sangat baik dan menunjukkan bahwa soal tes sumatif SMA B sudah sesuai dengan kaidah penyusunan soal pilihan ganda aspek konstruksi.

Hasil analisis terkait tingkat kesesuaian antara soal tes sumatif bahasa Jepang SMA A dan SMA B dengan kaidah penyusunan soal pilihan ganda aspek konstruksi dapat dilihat dalam diagram berikut:



Gambar 5 Diagram Persentase Kesesuaian Soal Tes Sumatif SMA A dan SMA B dengan Aspek Konstruksi

PENUTUP

Simpulan

Berdasarkan hasil analisis data yang telah dilakukan, dapat ditarik kesimpulan yang menjawab permasalahan penelitian bahwa:

1. Tingkat kesesuaian antara instrumen evaluasi bahasa Jepang di SMA A dengan kaidah penyusunan soal pilihan ganda aspek materi tergolong cukup baik dengan persentase sebesar 74,5%, sedangkan tingkat kesesuaian antara instrumen evaluasi bahasa Jepang di SMA B dengan kaidah penyusunan soal pilihan ganda aspek materi tergolong baik dengan persentase sebesar 82,5%.
2. Tingkat kesesuaian antara instrumen evaluasi bahasa Jepang di SMA A dengan kaidah penyusunan soal pilihan ganda aspek konstruksi tergolong sangat baik dengan persentase sebesar 95%, sedangkan tingkat kesesuaian antara instrumen evaluasi bahasa Jepang di SMA B dengan kaidah penyusunan soal pilihan ganda aspek konstruksi tergolong sangat baik dengan persentase sebesar 94,5%.
3. Ketidaksesuaian antara soal tes sumatif bahasa Jepang SMA A dan SMA B dengan kaidah penyusunan soal pilihan ganda aspek materi disebabkan oleh: 1) Indikator soal tidak sesuai dengan kisi-kisi dan RPP; 2) Soal tidak mewakili indikator pada RPP; 3) Adanya kemungkinan jawaban benar lebih dari 1; 4) Pengecoh yang tidak berfungsi karena pilihan jawaban tidak homogen dan panjangnya tidak sama. Ketidaksesuaian antara soal tes sumatif bahasa Jepang SMA A dan SMA B dengan kriteria aspek konstruksi disebabkan oleh: 1) Kesalahan penulisan pada soal ataupun pilihan jawaban; 2) Instruksi soal yang kurang sesuai dengan stimulus; 3) Panjang pilihan jawaban yang tidak sama dan tidak homogen; 4) Adanya ambiguitas pada pilihan jawaban.

Saran

Bersumber dari hasil penelitian dan kesimpulan pada penelitian ini, terdapat beberapa saran yang dapat diberikan sebagai bahan pertimbangan untuk memperbaiki kegiatan evaluasi pembelajaran bahasa Jepang. Adapun saran yang dapat diberikan adalah sebagai berikut:

- a. Bagi pendidik
 1. Pendidik hendaknya lebih memperhatikan relevansi antara soal dengan kompetensi dasar dan indikator dalam menyusun soal, agar soal dapat digunakan untuk mengukur kompetensi peserta didik secara optimal

2. Pendidik hendaknya lebih cermat dalam menyusun pilihan jawaban pada soal pilihan ganda terkait dengan pengecoh, agar soal dapat mengukur hasil belajar peserta didik dengan lebih akurat
- b. Bagi peneliti selanjutnya
Penelitian ini masih memiliki banyak kekurangan dan jauh dari kata sempurna. Analisis dalam penelitian ini hanya dilakukan dengan mengacu pada kaidah penyusunan soal pilihan ganda aspek materi dan konstruksi, sedangkan analisis aspek bahasa masih belum dilakukan. Hal tersebut dapat dijadikan sebagai masukan bagi peneliti selanjutnya untuk melakukan analisis kesesuaian antara instrumen evaluasi dengan kaidah penyusunan soal aspek bahasa.

DAFTAR PUSTAKA

- Arifin, Z. (2012). *Evaluasi Pembelajaran*. Jakarta: Direktorat Jenderal Pendidikan Islam Kementerian Agama.
- Arikunto, S. (2013). *Dasar-Dasar Evaluasi Pendidikan*. Jakarta: Bumi Aksara.
- Hayati, N. & Mardapi. (2014). Pengembangan Butir Soal Matematika SD di Kabupaten Lombok Timur sebagai Upaya dalam Pengadaan Bank Soal. *Jurnal Kependidikan*, 44(1), 26-38.
- Nurgiyantoro, B. (2013). *Penilaian Pembelajaran Bahasa Berbasis Kompetensi*. Yogyakarta: BPFE-YOGYAKARTA.
- Rahmawati, dkk. (2021). Analisis Kesesuaian Soal Penilaian Tengah Semester IPA dengan Kaidah Penyusunan Soal pada Aspek Bahasa di Sekolah Dasar. *Edukatif: Jurnal Ilmu Pendidikan*, 3(6), 3636-3646.
- Sudijono, A. (2009). *Pengantar Evaluasi Pendidikan*. Jakarta: Rajawali Pers.
- Sugiyono. (2016). *Metode Penelitian Pendidikan (Pendekatan Kuantitatif, Kualitatif, dan R&D)*. Bandung: Penerbit Alfabeta.
- Tim Pusat Penilaian Pendidikan. (2019). *Panduan Penilaian Tes Tertulis*. Jakarta: Pusat Penilaian Pendidikan.
- Widiyanto, J. (2018). *Evaluasi Pembelajaran*. Madiun: UNIPMA PRESS.