

INSTRUMEN PENGUKURAN KEMAMPUAN KOMUNIKASI DALAM KONTEKS PENDIDIKAN: SEBUAH TINJAUAN LITERATUR SISTEMATIS

Stela Veduta Ketlin Manalu¹, Jodion Siburian², Lely Mardiyanti³

¹⁻³ Universitas Jambi, Indonesia

✉ stelaveduta0404@gmail.com

Abstract Communication competence is a multidimensional construct that is assessed differently across educational settings, making instrument selection and interpretation challenging. This study aimed to map communication measurement instruments used in educational contexts, examine the measurement properties reported for those instruments, and identify priorities for future instrument development. A systematic literature review was conducted and reported with reference to PRISMA 2020. Searches covered Google Scholar, Scopus, and ScienceDirect for publications from 2020 to 2024. The original screening record identified 87 records, 65 of which were retained after title and abstract screening and 14 after full-text eligibility assessment. After methodological reclassification, five primary studies that directly developed or evaluated communication instruments in educational or health-professions education contexts formed the core psychometric synthesis; the remaining studies were used only as contextual literature. The core instruments assessed mathematical communication, science communication, general clinical communication, and trauma-informed communication. Evidence was strongest for content validity and internal consistency, while structural validity, measurement error, inter-rater reliability, and responsiveness were reported less consistently. Internal consistency was generally adequate where reported, but inter-rater reliability varied substantially across performance-based assessments. The review indicates that no single instrument can be recommended across contexts. Future development should define the communication construct explicitly, report multiple sources of validity evidence, examine measurement error and responsiveness, and undertake careful cultural and linguistic adaptation before instruments are used in new populations.

Keywords: communication competence; measurement instrument; psychometric properties; systematic literature review; educational assessment

Abstrak Kemampuan komunikasi merupakan konstruk multidimensi yang diukur secara berbeda pada berbagai konteks pendidikan, sehingga pemilihan dan interpretasi instrumen menjadi tantangan. Penelitian ini bertujuan memetakan instrumen pengukuran kemampuan komunikasi dalam konteks pendidikan, menelaah properti pengukuran yang dilaporkan, dan mengidentifikasi prioritas pengembangan instrumen berikutnya. Tinjauan literatur sistematis dilakukan dan dilaporkan dengan mengacu pada PRISMA 2020. Pencarian mencakup Google Scholar, Scopus, dan ScienceDirect untuk publikasi periode 2020–2024. Catatan penyaringan awal memuat 87 rekaman; 65 dipertahankan setelah penyaringan judul dan abstrak dan 14 setelah penilaian teks lengkap. Setelah reklasifikasi metodologis, lima studi primer yang secara langsung mengembangkan atau mengevaluasi instrumen komunikasi dalam konteks pendidikan atau pendidikan profesi kesehatan digunakan sebagai sintesis psikometrik inti; studi lainnya digunakan hanya sebagai literatur kontekstual. Instrumen inti mencakup komunikasi matematis, komunikasi sains, komunikasi klinis, dan komunikasi berbasis trauma-informed care. Bukti yang paling konsisten dilaporkan

adalah validitas isi dan konsistensi internal, sedangkan struktur internal, measurement error, reliabilitas antarpemilai, dan responsivitas belum dilaporkan secara merata. Reliabilitas antarpemilai pada instrumen berbasis kinerja juga menunjukkan variasi yang besar. Temuan menunjukkan bahwa belum ada satu instrumen yang dapat direkomendasikan untuk seluruh konteks. Pengembangan berikutnya perlu mendefinisikan konstruk komunikasi secara eksplisit, melaporkan berbagai sumber bukti validitas, menguji measurement error dan responsivitas, serta melakukan adaptasi budaya dan bahasa secara cermat sebelum instrumen digunakan pada populasi baru.

Kata Kunci: kemampuan komunikasi; instrumen pengukuran; properti psikometrik; tinjauan literatur sistematis; asesmen pendidikan

PENDAHULUAN

Komunikasi merupakan kompetensi penting dalam proses pendidikan karena peserta didik dan pendidik perlu menyampaikan gagasan, memahami pesan, memberi respons, serta menyesuaikan cara berkomunikasi dengan tujuan dan situasi interaksi. Perkembangan teknologi juga memperluas bentuk komunikasi dari tatap muka ke berbagai media digital, sehingga kompetensi komunikasi tidak lagi dapat dipahami hanya sebagai kemampuan berbicara secara verbal. Studi pada mahasiswa menunjukkan bahwa pola penggunaan teknologi berkaitan dengan persepsi mereka terhadap kompetensi komunikasi, sementara kajian komunikasi instruksional menegaskan pentingnya perilaku komunikasi interpersonal guru terhadap keterlibatan dan hasil belajar peserta didik (Morreale et al., 2015; Xie & Derakhshan, 2021).

Secara konseptual, istilah kemampuan, keterampilan, kompetensi, dan perilaku komunikasi tidak selalu merujuk pada konstruk yang sama. Kompetensi komunikasi umumnya mencakup pengetahuan, keyakinan atau afeksi, serta perilaku yang sesuai dengan tuntutan konteks, sedangkan keterampilan komunikasi lebih sering mengacu pada performa tertentu yang dapat diamati. Perbedaan konseptual tersebut berimplikasi langsung pada pengembangan instrumen karena definisi konstruk menentukan domain, indikator, bentuk item, dan cara interpretasi skor. Pengembangan alat ukur yang baik karena itu perlu dimulai dari definisi domain yang jelas agar terhindar dari construct underrepresentation maupun masuknya aspek yang tidak relevan dengan konstruk (Fantini, 2020; Boateng et al., 2018).

Permasalahan pengukuran menjadi lebih kompleks ketika instrumen digunakan pada populasi atau budaya yang berbeda dari konteks pengembangannya. Adaptasi instrumen tidak cukup dilakukan melalui penerjemahan, tetapi perlu mempertimbangkan kesetaraan konstruk, bahasa, proses respons, administrasi, penskoran, dan interpretasi hasil. Pedoman International Test Commission menekankan bahwa adaptasi lintas bahasa dan budaya memerlukan bukti empirik untuk memastikan skor tetap dapat ditafsirkan secara tepat. Temuan Kabir dan Sponseller (2020), misalnya, menunjukkan bahwa instrumen self-efficacy dalam komunikasi antarbudaya perlu diuji kembali struktur dan reliabilitasnya setelah adaptasi bahasa pada kelompok guru dan mahasiswa. Prinsip ini relevan ketika instrumen komunikasi hendak digunakan dalam konteks pendidikan Indonesia (International Test Commission, 2017; Kabir & Sponseller, 2020).

Penelitian 2020–2024 memperlihatkan bahwa instrumen komunikasi dalam pendidikan berkembang pada domain yang sangat spesifik. Azmi dan Salam (2020) serta Sudrajat (2024) menelaah tes komunikasi matematis pada siswa SMP; Rahmawati et al. (2021) mengembangkan instrumen komunikasi

sains di pendidikan tinggi; sedangkan Thang et al. (2024) dan Fan et al. (2024) mengevaluasi instrumen komunikasi dalam pendidikan profesi kesehatan. Keragaman tersebut menunjukkan bahwa bukti psikometrik tidak dapat digeneralisasi begitu saja dari satu konteks ke konteks lain. Selain itu, studi terbaru di luar periode sintesis juga menunjukkan bahwa pelaporan bukti validitas instrumen komunikasi masih lebih sering menekankan konten dan struktur internal dibandingkan proses respons dan konsekuensi penggunaan skor (Dorrestein et al., 2025).

Berdasarkan kondisi tersebut, penelitian ini bertujuan untuk: (1) mengidentifikasi instrumen pengukuran kemampuan komunikasi yang digunakan dalam konteks pendidikan; (2) menganalisis bukti validitas, reliabilitas, measurement error, dan responsivitas yang dilaporkan; serta (3) mengidentifikasi kesenjangan metodologis yang perlu diperhatikan dalam pengembangan atau adaptasi instrumen komunikasi selanjutnya. Fokus pada konteks pendidikan dipilih agar sintesis lebih relevan dengan kebutuhan asesmen peserta didik, mahasiswa, calon profesional, dan pendidik.

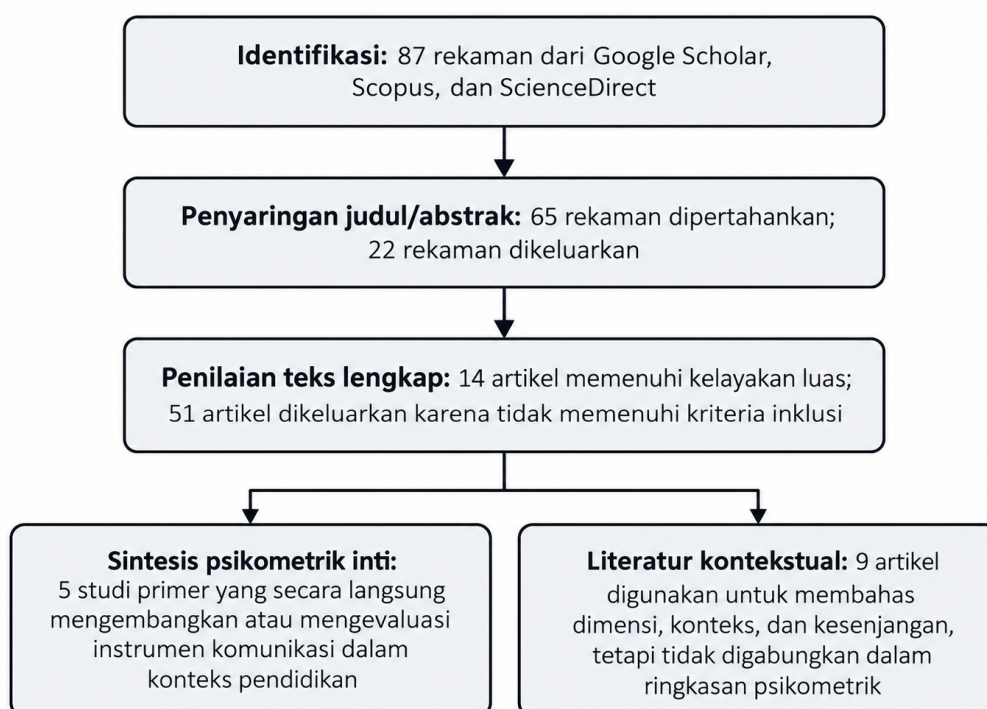
METODE

Penelitian ini menggunakan Systematic Literature Review (SLR). Pelaporan proses identifikasi dan seleksi literatur mengacu pada PRISMA 2020, sedangkan pelaporan strategi pencarian mempertimbangkan rekomendasi PRISMA-S agar sumber informasi, istilah pencarian, batas pencarian, dan tahapan seleksi dapat ditelusuri secara lebih jelas (Page et al., 2021; Rethlefsen et al., 2021). PRISMA digunakan sebagai pedoman pelaporan, bukan sebagai metode analisis.

Pencarian awal dilakukan pada Google Scholar, Scopus, dan ScienceDirect untuk publikasi tahun 2020–2024. Blok istilah pencarian yang tercatat adalah: ("communication competence" OR "communication skills" OR "mathematical communication" OR "science communication") AND (instrument OR scale OR questionnaire OR assessment) AND (validation OR reliability OR psychometric). Padanan bahasa Indonesia meliputi ("kemampuan komunikasi" OR "keterampilan komunikasi" OR "komunikasi matematis" OR "komunikasi sains") AND (instrumen OR skala OR kuesioner OR penilaian) AND (validitas OR reliabilitas OR psikometrik). Istilah tersebut disesuaikan dengan fitur pencarian pada masing-masing sumber. Metadata bibliografis dan DOI studi yang dipertahankan diverifikasi kembali melalui laman penerbit pada tahap revisi. Log penelitian awal tidak menyimpan tanggal pencarian dan sintaks spesifik tiap basis data secara terpisah; keterbatasan dokumentasi ini dinyatakan secara eksplisit pada bagian keterbatasan tinjauan.

Kriteria inklusi untuk sintesis psikometrik inti adalah: (1) artikel penelitian primer peer-reviewed; (2) terbit tahun 2020–2024 dalam bahasa Indonesia atau Inggris; (3) berfokus pada konteks pendidikan, termasuk pendidikan profesi; (4) secara langsung mengembangkan, memvalidasi, atau mengevaluasi reliabilitas instrumen komunikasi; dan (5) melaporkan sekurang-kurangnya satu properti pengukuran. Artikel konseptual, studi intervensi yang tidak mengevaluasi instrumen, artikel klinis yang tidak berkaitan dengan pendidikan, serta tinjauan sekunder tidak digabungkan dalam sintesis psikometrik. Literatur tersebut tetap dapat digunakan sebagai sumber kontekstual pada pembahasan. Artikel tahun 2025 yang relevan juga diperlakukan sebagai literatur pendukung di luar periode sintesis, bukan sebagai studi inti.

Berdasarkan catatan penyaringan yang tersedia, sebanyak 87 rekaman teridentifikasi pada tahap awal. Setelah penyaringan judul dan abstrak, 65 rekaman dipertahankan sehingga 22 rekaman dikeluarkan pada tahap ini. Penilaian teks lengkap kemudian menghasilkan 14 artikel yang memenuhi kelayakan awal, sedangkan 51 artikel dikeluarkan karena tidak memenuhi fokus atau kriteria yang ditetapkan. Setelah kriteria konteks pendidikan dan properti psikometrik diterapkan secara konsisten, lima studi primer ditetapkan sebagai unit sintesis psikometrik inti, sedangkan sembilan artikel lain digunakan sebagai literatur kontekstual dan tidak digabungkan dalam perhitungan rentang koefisien psikometrik. Alur seleksi disajikan pada Gambar 1.



Gambar 1. Alur seleksi literatur dan penetapan studi sintesis psikometrik inti

Data yang diekstraksi dari studi inti meliputi nama instrumen, konstruk yang diukur, populasi dan ukuran sampel, bentuk/format instrumen, metode validasi, reliabilitas, measurement error, serta informasi responsivitas jika tersedia. Karena instrumen yang ditinjau heterogen dan tidak seluruhnya merupakan patient-reported outcome measures, prinsip COSMIN digunakan secara adaptif sebagai kerangka pemetaan kelengkapan bukti, bukan sebagai penilaian formal risiko bias. Properti yang diperiksa meliputi validitas isi, struktur/internal structure, reliabilitas, measurement error, validitas konstruk, dan responsivitas; pengembangan skala juga dipertimbangkan berdasarkan prinsip umum validasi instrumen (Prinsen et al., 2018; Boateng et al., 2018). Tidak tersedianya dokumentasi tentang seleksi dan ekstraksi oleh dua penelaah independen pada penelitian awal diakui sebagai salah satu keterbatasan.

HASIL DAN PEMBAHASAN

Dari 14 artikel yang lolos penilaian teks lengkap secara luas, lima studi memenuhi kriteria sebagai studi inti karena secara langsung mengembangkan atau mengevaluasi instrumen komunikasi dalam konteks pendidikan dan melaporkan bukti psikometrik. Studi tersebut mewakili pendidikan matematika, pendidikan sains, pendidikan kedokteran, dan pendidikan terapi okupasi. Sembilan artikel lainnya digunakan untuk memperkaya pembahasan mengenai dimensi komunikasi, konteks digital, komunikasi antarbudaya, dan penggunaan komunikasi dalam profesi, tetapi tidak digunakan untuk menghitung rentang koefisien psikometrik.

Tabel 1. Sintesis studi inti dan bukti properti pengukuran

Studi/instrumen	Karakteristik dan bukti pengukuran
<p>Azmi & Salam (2020)</p> <p><i>Tes komunikasi matematis</i></p>	<p>Konteks dan sampel: Siswa SMP; ukuran sampel tidak dinyatakan secara eksplisit pada laporan metode</p> <p>Bentuk instrumen: 3 soal uraian dengan 4 komponen skor; rubrik 0–4</p> <p>Bukti validitas: Korelasi item–total $r = 0,636–0,782$; seluruh komponen dinyatakan valid</p> <p>Reliabilitas / measurement error: $\alpha = 0,822$; daya pembeda $0,42–0,72$</p> <p>Catatan mutu bukti: Bukti terbatas pada analisis butir klasik dan konsistensi internal; struktur internal dan measurement error tidak dilaporkan.</p>
<p>Rahmawati et al. (2021)</p> <p><i>Instrumen komunikasi sains</i></p>	<p>Konteks dan sampel: Mahasiswa pendidikan fisika; $n = 94$ (validasi empiris melibatkan 90 mahasiswa)</p> <p>Bentuk instrumen: 2 soal uraian pada dua indikator komunikasi sains</p> <p>Bukti validitas: Validasi ahli oleh dua ahli; analisis empiris menggunakan Quest, fit item $0,7–1,3$</p> <p>Reliabilitas / measurement error: Instrumen dilaporkan reliabel, tetapi koefisien reliabilitas tidak disajikan secara rinci dalam pelaporan utama</p> <p>Catatan mutu bukti: Pengembangan konstruk dan uji empirik tersedia; pelaporan koefisien reliabilitas masih perlu lebih transparan.</p>
<p>Thang et al. (2024)</p> <p><i>GKCSAF-TTC</i></p>	<p>Konteks dan sampel: 33 pediatric interns; 57 standardized-patient encounters</p> <p>Bentuk instrumen: 8 domain trauma-informed communication; skala penilaian 5 poin</p> <p>Bukti validitas: Bukti konten, proses respons, struktur internal, dan hubungan dengan variabel lain berdasarkan kerangka Messick</p> <p>Reliabilitas / measurement error: $\alpha = 0,93–0,96$; ICC antarpemilai = $0,80–0,83$; skor sensitif terhadap perubahan pra–pasca pelatihan</p> <p>Catatan mutu bukti: Bukti psikometrik paling lengkap di antara studi inti, meskipun masih bersifat preliminary pada satu program pendidikan.</p>
<p>Fan et al. (2024)</p> <p><i>GKCSAF</i></p>	<p>Konteks dan sampel: 30 pasien; 22 assessor (7 terapis, 15 mahasiswa); 5 penilai</p> <p>Bentuk instrumen: Penilaian langsung dan coded rating terhadap komunikasi terapi okupasi</p> <p>Bukti validitas: Fokus studi pada reliabilitas dan kesalahan pengukuran; bukti validitas lain tidak menjadi tujuan utama</p> <p>Reliabilitas / measurement error: Direct rating: ICC intra = $0,76$; inter = $0,31$; MDC% = $17,8\%$ dan $38,3\%$. Coded rating: ICC intra = $0,69$; inter = $0,22$</p> <p>Catatan mutu bukti: Menunjukkan bahwa konsistensi antarpemilai dapat menjadi masalah serius; pelatihan dan penggunaan beberapa penilai diperlukan.</p>

Studi/instrumen	Karakteristik dan bukti pengukuran
Sudrajat (2024) <i>Tes komunikasi matematis</i>	<p>Konteks dan sampel: 26 siswa kelas VIII SMP Negeri 7 Yogyakarta</p> <p>Bentuk instrumen: 5 soal uraian</p> <p>Bukti validitas: Korelasi item-total $r = 0,680-0,957$; seluruh butir dinyatakan valid</p> <p>Reliabilitas / measurement error: $\alpha = 0,870$; 3 butir daya pembeda cukup dan 2 baik; 3 butir kesukaran sedang dan 2 mudah</p> <p>Catatan mutu bukti: Pelaporan kualitas butir cukup rinci, tetapi validitas masih berfokus pada korelasi item-total dan belum mencakup struktur konstruk.</p>

Karakteristik Instrumen dan Konstruk yang Diukur

Kelima studi inti memperlihatkan bahwa istilah kemampuan komunikasi mencakup konstruk yang berbeda. Instrumen komunikasi matematis menekankan kemampuan menyatakan ide atau situasi matematika melalui representasi, penjelasan tertulis, model, serta evaluasi pemikiran matematis (Azmi & Salam, 2020; Sudrajat, 2024). Instrumen komunikasi sains menilai pemahaman gagasan kunci ujaran dan kemampuan menyajikan atau membaca data dalam tabel/grafik (Rahmawati et al., 2021). Pada pendidikan profesi kesehatan, GKCSAF dan adaptasinya menilai perilaku komunikasi yang ditampilkan dalam interaksi klinis atau simulasi, sehingga skor sangat dipengaruhi oleh konteks performa dan konsistensi penilai (Thang et al., 2024; Fan et al., 2024).

Perbedaan tersebut menunjukkan bahwa istilah communication competence tidak dapat diperlakukan sebagai konstruk tunggal yang selalu memiliki indikator sama. Implikasi praktisnya, peneliti perlu memilih instrumen berdasarkan definisi konstruk dan tujuan penggunaan skor, bukan hanya berdasarkan nilai reliabilitas yang tinggi. Prinsip pengembangan skala menekankan bahwa domain konstruk, proses pengembangan item, pengujian dimensionalitas, reliabilitas, dan validitas harus saling konsisten agar skor memiliki interpretasi yang dapat dipertanggungjawabkan (Boateng et al., 2018).

Analisis Properti Psikometri Instrumen

Sintesis menunjukkan bahwa bukti konsistensi internal paling sering dilaporkan, tetapi pelaporannya tidak seragam. Azmi dan Salam (2020) melaporkan $\alpha = 0,822$ dan Sudrajat (2024) $\alpha = 0,870$ pada tes komunikasi matematis. Thang et al. (2024) melaporkan $\alpha = 0,93-0,96$ pada GKCSAF-TIC. Koefisien tersebut menunjukkan konsistensi internal yang memadai pada konteks masing-masing studi, tetapi tidak tepat digabungkan sebagai satu rentang kualitas universal karena bentuk instrumen, konstruk, populasi, serta model penskorannya berbeda. Reliabilitas juga tidak dapat direduksi hanya menjadi alfa Cronbach, terutama pada asesmen berbasis performa yang melibatkan penilai.

Temuan Fan et al. (2024) memperlihatkan pentingnya membedakan reliabilitas intra- dan antarpemilai serta melaporkan measurement error. Pada direct rating GKCSAF, ICC intra-pemilai berada pada 0,76, tetapi ICC antarpemilai hanya 0,31; MDC% juga meningkat dari 17,8% untuk penilai yang sama menjadi 38,3% untuk penilai berbeda. Artinya, skor dapat cukup stabil ketika penilai yang sama melakukan penilaian ulang, tetapi kurang konsisten jika penilai berganti.

Sebaliknya, [Thang et al. \(2024\)](#) melaporkan ICC antarpemilai 0,80–0,83 setelah pengembangan rubrik dan pelatihan pemilai. Perbandingan ini menunjukkan bahwa standarisasi rubrik dan pelatihan pemilai merupakan bagian penting dari kualitas instrumen komunikasi berbasis observasi.

Dari perspektif validitas, studi di bidang pendidikan matematika dan sains masih banyak menggunakan validasi ahli dan korelasi item–total, sedangkan bukti struktur internal, hubungan dengan variabel lain, proses respons, dan konsekuensi penggunaan skor belum dilaporkan secara konsisten. Kerangka validitas modern dan prinsip COSMIN menekankan bahwa kualitas alat ukur sebaiknya dinilai dari berbagai sumber bukti, tidak hanya dari satu koefisien statistik ([Prinsen et al., 2018](#)). Temuan ini juga sejalan dengan scoping review terbaru di pendidikan profesi kesehatan yang menunjukkan bahwa bukti proses respons dan konsekuensi jauh lebih jarang dilaporkan dibandingkan struktur internal dan bukti terkait konten ([Dorrestein et al., 2025](#)).

Kesenjangan dan Implikasi Pengembangan Instrumen

Kesenjangan pertama adalah terbatasnya instrumen pendidikan yang secara eksplisit memasukkan komunikasi digital sebagai bagian dari konstruk yang dinilai. Padahal peserta didik dan mahasiswa berkomunikasi melalui tatap muka, pesan tertulis, konferensi video, dan platform kolaborasi. [Morreale et al. \(2015\)](#) menunjukkan bahwa teknologi telah menjadi bagian dari pilihan dan pengalaman komunikasi mahasiswa. Oleh karena itu, instrumen yang dikembangkan untuk konteks pendidikan kontemporer perlu menentukan dengan jelas apakah komunikasi digital merupakan bagian dari konstruk utama, konteks performa, atau sekadar media penyampaian.

Kesenjangan kedua berkaitan dengan adaptasi budaya dan bahasa. Penggunaan instrumen yang dikembangkan di negara atau bahasa lain perlu mempertimbangkan kesetaraan konstruk, item, proses respons, dan interpretasi skor. [International Test Commission \(2017\)](#) menegaskan bahwa adaptasi tes memerlukan tahapan pengembangan, konfirmasi empiris, administrasi, penskoran, dan dokumentasi yang sistematis. [Kabir dan Sponseller \(2020\)](#) juga menunjukkan bahwa adaptasi instrumen komunikasi antarbudaya memerlukan back-translation, pengujian struktur faktor, dan pemeriksaan reliabilitas pada kelompok sasaran. Bagi konteks Indonesia, pendekatan ini lebih kuat daripada sekadar mengasumsikan bahwa karakter budaya tertentu otomatis berlaku pada seluruh responden.

Dengan demikian, rekomendasi pengembangan instrumen komunikasi di bidang pendidikan adalah: definisikan konstruk dan konteks penggunaan skor sejak awal; gunakan bentuk instrumen yang sesuai dengan performa yang hendak diukur; sertakan validitas isi dan proses respons; uji struktur internal ketika konstruk bersifat laten; laporkan reliabilitas yang sesuai dengan desain penskoran; laporkan measurement error untuk instrumen kinerja; serta uji responsivitas jika instrumen digunakan untuk mengevaluasi perubahan setelah intervensi. Rekomendasi ini perlu diterapkan secara proporsional sesuai jenis instrumen dan tujuan asesmen.

Keterbatasan Tinjauan

Tinjauan ini memiliki beberapa keterbatasan. Pertama, pencarian hanya menggunakan tiga sumber utama dan tanggal pencarian awal tidak terdokumentasi secara terpisah, sehingga reproduksibilitas pencarian belum optimal. Kedua, konstruk komunikasi yang ditemukan sangat heterogen, mulai dari komunikasi matematis hingga komunikasi klinis, sehingga hasil psikometrik tidak dapat digeneralisasi sebagai satu ukuran kualitas. Ketiga, catatan penelitian awal tidak mendokumentasikan seleksi dan ekstraksi oleh dua penelaah independen, sehingga potensi bias seleksi tetap ada. Keempat, penggunaan prinsip COSMIN dalam tinjauan ini bersifat adaptif karena sebagian instrumen berupa tes pendidikan dan performance assessment, bukan patient-reported outcome measures. Keterbatasan tersebut membuat rekomendasi perlu dipahami sebagai pemetaan bukti, bukan peringkat definitif terhadap instrumen.

SIMPULAN DAN SARAN

Tinjauan ini menunjukkan bahwa instrumen pengukuran kemampuan komunikasi dalam pendidikan sangat bergantung pada definisi konstruk dan konteks penggunaan. Pada lima studi inti, validitas isi/ahli dan konsistensi internal merupakan bukti yang paling sering dilaporkan, sedangkan struktur internal, measurement error, reliabilitas antarpemilai, dan responsivitas belum tersedia secara merata. Studi berbasis observasi memperlihatkan bahwa reliabilitas antarpemilai dapat menjadi kelemahan utama apabila rubrik dan pelatihan pemilai tidak cukup terstandarisasi. Karena itu, tidak ada satu instrumen yang dapat dinyatakan paling tepat untuk semua konteks pendidikan. Pemilihan atau pengembangan instrumen perlu didasarkan pada definisi konstruk yang jelas, kesesuaian populasi, tujuan interpretasi skor, serta kelengkapan bukti validitas dan reliabilitas. Untuk konteks Indonesia, adaptasi bahasa dan budaya juga perlu disertai pengujian empirik sebelum skor digunakan untuk keputusan pendidikan.

Penelitian selanjutnya disarankan mengembangkan dan menguji instrumen komunikasi pada populasi pendidikan Indonesia dengan definisi konstruk yang jelas serta melaporkan bukti validitas, reliabilitas, dan kesalahan pengukuran secara memadai. Untuk instrumen berbasis performa, standarisasi rubrik dan pelatihan pemilai perlu diperhatikan.

DAFTAR PUSTAKA

- Azmi, M. P., & Salam, A. (2020). Pengembangan instrumen tes kemampuan komunikasi matematis pada materi segi empat. *JURING (Journal for Research in Mathematics Learning)*, 3(2), 181--192. <https://doi.org/10.24014/juring.v3i2.10029>
- Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quinonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: A primer. *Frontiers in Public Health*, 6, 149. <https://doi.org/10.3389/fpubh.2018.00149>

- Dorrestein, L., Ritter, C., De Mol, Z., Wichtel, M., Cary, J., Vengrin, C., Artemiou, E., Adams, C. L., Ganshorn, H., Coe, J. B., Barkema, H., & Hecker, K. G. (2025). Validity evidence for communication skills assessment in health professions education: A scoping review. *BMJ Open*, 15, e096799. <https://doi.org/10.1136/bmjopen-2024-096799>
- Fan, S. C., Tsai, S. T., Wang, Y. C., Lee, M. L., Huang, S. L., & Hsieh, C. L. (2024). Reliability of the Gap-Kalamazoo communication skills assessment form in occupational therapy. *British Journal of Occupational Therapy*, 87(7), 424--433. <https://doi.org/10.1177/03080226241239574>
- Fantini, A. E. (2020). Reconceptualizing intercultural communicative competence: A multinational perspective. *Research in Comparative and International Education*, 15(1), 52--61. <https://doi.org/10.1177/1745499920901948>
- International Test Commission. (2017). *The ITC guidelines for translating and adapting tests (2nd ed.)*. https://www.intestcom.org/files/guideline_test_adaptation_2ed.pdf
- Kabir, R. S., & Sponseller, A. C. (2020). Interacting with competence: A validation study of the Self-Efficacy in Intercultural Communication Scale-Short Form. *Frontiers in Psychology*, 11, 2086. <https://doi.org/10.3389/fpsyg.2020.02086>
- Morreale, S., Staley, C., Stavrositu, C., & Krakowiak, M. (2015). First-year college students' attitudes toward communication technologies and their perceptions of communication competence in the 21st century. *Communication Education*, 64(1), 107--131. <https://doi.org/10.1080/03634523.2014.978799>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., et al. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372, n71. <https://doi.org/10.1136/bmj.n71>
- Prinsen, C. A. C., Mokkink, L. B., Bouter, L. M., Alonso, J., Patrick, D. L., de Vet, H. C. W., & Terwee, C. B. (2018). COSMIN guideline for systematic reviews of patient-reported outcome measures. *Quality of Life Research*, 27(5), 1147--1157. <https://doi.org/10.1007/s11136-018-1798-3>
- Rahmawati, L., Jumadi, & Ikhsan, J. (2021). Pengembangan instrumen penilaian kemampuan komunikasi sains sebagai bagian dari keterampilan abad 21. *Prima Magistra: Jurnal Ilmiah Kependidikan*, 2(2), 163--171. <https://doi.org/10.37478/jpm.v2i2.933>
- Rethlefsen, M. L., Kirtley, S., Waffenschmidt, S., Ayala, A. P., Moher, D., Page, M. J., Koffel, J. B., & PRISMA-S Group. (2021). PRISMA-S: An extension to the PRISMA Statement for reporting literature searches in systematic reviews. *Systematic Reviews*, 10, 39. <https://doi.org/10.1186/s13643-020-01542-z>
- Sudrajat, S. (2024). Analisis kualitas instrumen tes untuk mengukur kemampuan komunikasi matematika siswa SMP. *ELIPS: Jurnal Pendidikan Matematika*, 5(2), 210--221. <https://doi.org/10.47650/elips.v5i2.1398>

- Thang, C. K., Kucaj, S., Garell, C. L., Masood, K. M., Calhoun, A. W., Lay, K., Lee, J., Wilhalme, H., & Szilagyi, M. A. (2024). Development and validation of a trauma-informed care communication skills assessment tool. *Academic Pediatrics*, 24(8), 1333--1342. <https://doi.org/10.1016/j.acap.2024.07.008>
- Xie, F., & Derakhshan, A. (2021). A conceptual review of positive teacher interpersonal communication behaviors in the instructional context. *Frontiers in Psychology*, 12, 708490. <https://doi.org/10.3389/fpsyg.2021.708490>