ANTHOR
Education and Learning Journal
https://anthor.org/index.php/anthor
Email. anthor.Jurnal@gmail.com
E-ISSN: 2963 –198X P-ISSN: 2963–2498 | Halaman: 1621 - 1627
Volume 5 Nomor 5 Tahun 2026

Uji Validitas dan Reliabilitas Instrumen Penilaian Kinerja Wasit Bola Basket

Informasi Naskah

Tria Rahmawati[1]
Ilham Truly Fajri[2]
Alen Rismayadi[3]
Lukmannul Haqim Lubay[4]
Widi Kusumah[5]
Patriana Nurmansyah Awwaludin[6]

Program Studi Pendidikan Kepelatihan Olahraga, Fakultas Pendidikan Olahraga dan Kesehatan, Universitas Pendidikan Indonesia, Indonesia[1,5,6]
Program Studi Kepelatihan Fisik Olahraga, Fakultas Pendidikan Olahraga dan Kesehatan, Universitas Pendidikan Indonesia, Indonesia[2,3]
Program Studi Pendidikan Guru Sekolah Dasar Pendidikan Jasmani, Fakultas Pendidikan Olahraga dan Kesehatan, Universitas Pendidikan Indonesia, Indonesia[4]

ilhamtrulyfa@upi.edu

Abstrak
Penelitian ini menguji validitas butir dan konsistensi internal Match Assessment versi bahasa Indonesia untuk menilai kinerja wasit bola basket. Penelitian menggunakan pendekatan kuantitatif deskriptif dengan desain pengujian instrumen. Sampel terdiri atas 30 wasit level B2 yang bertugas pada Pekan Olahraga Pelajar Wilayah (POPWIL) Jawa Barat tahun 2026 di Bandung, Purwakarta, dan Cirebon serta dipilih melalui purposive sampling. Seorang asesor yang berlisensi A, berkualifikasi referee coach, dan menjabat Ketua Bidang Perwasitan DPD PERBASI Jawa Barat melakukan observasi. Instrumen hasil terjemahan memuat 13 butir dikotomis dengan skor 0 untuk Belum dan 1 untuk Sudah pada tujuh area kinerja. Validitas butir diuji menggunakan korelasi Pearson antara skor butir dan skor total; butir diterima apabila r hitung lebih besar dari 0,361 dan nilai signifikansi lebih kecil dari 0,05. Butir 1, 4, 5, 6, 9, 10, dan 11 memenuhi kedua kriteria, sedangkan butir 2, 3, 7, 8, 12, dan 13 belum memenuhi kriteria. Cronbach's Alpha seluruh butir sebesar 0,610 dan termasuk cukup reliabel berdasarkan kriteria penelitian. Instrumen dapat mendukung penilaian awal dan umpan balik, tetapi enam butir yang belum valid perlu diperjelas, dilengkapi petunjuk skor, dan diuji ulang dengan sampel lebih besar serta beberapa asesor sebelum digunakan untuk keputusan lisensi atau penugasan.

Kata Kunci:
Bola Basket, Instrument, Kinerja Wasit, Reliabilitas, Validitas.

Abstract
This study examined the item validity and internal consistency of the Indonesian version of Basketball Australia's Match Assessment for evaluating basketball referee performance. A descriptive quantitative instrument-testing design was used. Thirty B2-level referees officiating the 2026 West Java Student Regional Sports Week (POPWIL) in Bandung, Purwakarta, and Cirebon were selected through purposive sampling. One assessor who was an A-licensed referee, a referee coach, and the head of the West Java PERBASI refereeing division observed the referees. The translated instrument comprised 13 dichotomous items scored 0 for Not Yet and 1 for Achieved across communication, mechanics, game calling, game control, professionalism, game administration, and rules. Item validity was tested using Pearson item-total correlations; items were accepted when the calculated correlation exceeded 0.361 and p was below 0.05. Items 1, 4, 5, 6, 9, 10, and 11 met both criteria, whereas items 2, 3, 7, 8, 12, and 13 did not. Cronbach's Alpha for all 13 items was 0.610, indicating adequate reliability under the study criterion. The instrument can support preliminary referee feedback, but the six non-valid items require clearer wording, scoring guidance, and retesting with larger samples and multiple assessors before the instrument is used for consequential licensing or assignment decisions.

Keyword:
Basketball, Instrument, Referee Performance, Reliability, Validity.

Information Production:
© 2026 *by the authors of this article. This work is licensed under an Attribution-ShareAlike* 4.0 *International* (CC BY-SA 4.0)

PENDAHULUAN

Bola basket berlangsung cepat, dinamis, dan dipenuhi perubahan penguasaan bola, kontak, serta perpindahan posisi dalam waktu singkat. Dalam kondisi tersebut, wasit harus mengikuti permainan, memperoleh sudut pandang yang tepat, dan mengambil keputusan yang dapat memengaruhi kelanjutan pertandingan, pemberian sanksi, perolehan angka, dan persepsi keadilan kedua tim. Tugas wasit tidak berhenti pada pengenalan foul dan violation. Wasit juga harus menggunakan sinyal, berkomunikasi dengan pemain dan petugas meja, mengelola pergantian serta time-out, menjaga konsentrasi, dan menyelesaikan administrasi pertandingan sesuai prosedur. Karena itu, kualitas pertandingan bergantung pula pada kinerja perangkat pertandingan yang mampu menerapkan peraturan secara konsisten dan profesional (FIBA, 2024).

Kinerja perwasitan bersifat multidimensional. Pengetahuan peraturan perlu didukung oleh kondisi fisik, kesiapan mental, kepemimpinan, kepercayaan diri, komunikasi, kemampuan mengambil keputusan, dan penempatan posisi. Penelitian terhadap wasit bola basket di Jawa Barat menunjukkan hubungan kebugaran jasmani dengan kepercayaan diri (Rosyadi et al., 2017), sedangkan kajian terhadap wasit bola basket

e-ISSN: 2963 – 198X
p-ISSN: 2963 – 2498

mahasiswa menegaskan pentingnya pemahaman foul dan violation (Rohmana & Herpandika, 2021). Pada futsal dan hoki, kondisi fisik, kesiapan mental, pemahaman aturan, pengetahuan, kepemimpinan, dan kepercayaan diri juga berkaitan dengan kinerja wasit (Arief & Nurcahya, 2020; Indriansah & Arief, 2020; Margiyan et al., 2022). Temuan pada wasit bola voli menunjukkan bahwa klasifikasi dan pengalaman perlu dipertimbangkan ketika kinerja dinilai (Sanusi et al., 2022).

Pembinaan wasit memerlukan evaluasi yang mampu mengubah pengamatan pertandingan menjadi umpan balik yang spesifik. Pelatihan dapat meningkatkan kemampuan wasit pemula (Prakasa et al., 2019), sementara media pembelajaran sinyal membantu standardisasi komunikasi nonverbal wasit bola basket (Pambudi & Suharjana, 2018). Namun, manfaat pembinaan tersebut sulit dipantau apabila evaluasi hanya didasarkan pada kesan umum seorang penilai. Penilaian tanpa indikator yang jelas dapat membuat satu kesalahan menonjol menutupi aspek lain atau menghasilkan standar yang berbeda antarpertandingan. Instrumen observasi menyediakan kerangka yang sama bagi asesor untuk mengenali kemampuan yang telah dikuasai dan bagian yang masih perlu diperbaiki.

Instrumen yang praktis belum tentu menghasilkan informasi yang sahih. Pengujian alat ukur olahraga di Indonesia memperlihatkan bahwa validitas dan reliabilitas perlu ditetapkan sebelum hasil digunakan secara luas, baik pada tes tenis, reactive agility tenis meja, kemampuan fisik bola basket, maupun pengelolaan sekolah sepak bola (Alim & Nurfadhila, 2021; Guntur et al., 2020; Setyawan & Sugiyanto, 2016; Sulistiyono et al., 2022; Tomoliyus & Sunardianta, 2020). Validitas butir menunjukkan seberapa kuat skor sebuah indikator bergerak bersama skor keseluruhan pada sampel penelitian, sedangkan reliabilitas internal memberi gambaran konsistensi kumpulan butir. Kedua hasil tersebut tetap harus dibaca bersama isi konstruk, variasi skor, dan kondisi pengamatan (Boateng et al., 2018).

Dalam konteks perwasitan bola basket di Indonesia, sumber yang digunakan dalam penelitian ini belum menunjukkan adanya instrumen baku nasional yang disertai bukti validitas dan reliabilitas. Salah satu sumber yang dapat diadaptasi adalah Match Assessment dalam Intermediate Referee Workbook Basketball Australia. Instrumen ini memuat 13 butir dengan pilihan Belum dan Sudah yang dikelompokkan ke dalam tujuh area: komunikasi, mekanika perwasitan, pemanggilan pertandingan, pengendalian pertandingan, profesionalisme, administrasi pertandingan, dan peraturan (Basketball Australia, n.d.). Cakupan tersebut sesuai dengan kebutuhan untuk mengamati wasit secara lebih utuh daripada hanya menghitung keputusan foul atau violation.

Match Assessment dikembangkan dalam bahasa dan sistem pembinaan Australia. Dalam penelitian ini, dokumen diterjemahkan ke bahasa Indonesia oleh penerjemah tersumpah. Penerjemahan formal membantu menjaga keterbacaan, tetapi tidak otomatis memastikan kesetaraan penggunaan pada konteks baru. Adaptasi lintas bahasa perlu mempertahankan makna konsep, petunjuk, dan cara pengguna menafsirkan indikator (Beaton et al., 2000). Istilah seperti lead, trail, gap, off-ball, post play, dan preventative voice memiliki arti teknis dalam praktik perwasitan; padanan bahasa yang benar tetap membutuhkan contoh perilaku dan pemahaman mekanika yang sama di antara asesor.

Skala dikotomis pada Match Assessment membuat penilaian cepat dilakukan selama pertandingan, tetapi juga dapat menghasilkan skor homogen ketika hampir seluruh wasit telah menguasai kemampuan dasar. Selain itu, beberapa indikator hanya dapat diamati apabila situasi tertentu muncul. Keputusan violation, foul tertentu, respons terhadap cedera, atau penjelasan aturan mungkin tidak memperoleh kesempatan observasi yang sama pada setiap pertandingan. Karena itu, korelasi item-total yang rendah tidak serta-merta berarti area kompetensi tersebut tidak penting; temuan dapat mencerminkan redaksi yang terlalu luas, kesempatan observasi yang terbatas, atau variasi skor yang kecil.

Penelitian ini mengisi kebutuhan awal tersebut dengan menguji Match Assessment versi bahasa Indonesia pada 30 wasit level B2 yang bertugas dalam POPWIL Jawa Barat. Analisis difokuskan pada korelasi Pearson antara setiap skor butir dan skor total serta Cronbach's Alpha untuk seluruh 13 butir. Tujuan penelitian adalah mengetahui butir yang memenuhi kriteria validitas dan tingkat reliabilitas internal instrumen pada kondisi penelitian. Hasilnya diharapkan menjadi dasar untuk mempertahankan indikator yang berfungsi baik, memperbaiki indikator yang belum memenuhi kriteria, dan membatasi penggunaan instrumen agar tidak melampaui bukti yang tersedia.

METODE PENELITIAN

Penelitian menggunakan pendekatan kuantitatif deskriptif dengan desain pengujian instrumen. Penelitian tidak memberikan perlakuan kepada sampel; data berbentuk skor observasi kinerja wasit selama pertandingan. Pelaksanaan berlangsung pada kegiatan Pekan Olahraga Pelajar Wilayah (POPWIL) Jawa Barat di Bandung, Purwakarta, dan Cirebon pada 15 Juni sampai 10 Juli 2026. Populasi penelitian adalah wasit bola basket level B2 di Jawa Barat. Sampel dipilih dengan purposive sampling berdasarkan tiga kriteria: memiliki

e-ISSN: 2963 – 198X
p-ISSN: 2963 – 2498

lisensi B2, bertugas pada POPWIL Jawa Barat tahun 2026, dan memiliki data lengkap pada seluruh butir. Sebanyak 30 wasit memenuhi kriteria tersebut.

Instrumen penelitian adalah Match Assessment yang diadopsi dari Intermediate Referee Workbook Basketball Australia dan telah diterjemahkan oleh penerjemah tersumpah. Instrumen memuat 13 butir dalam tujuh area kinerja. Setiap butir diberi skor 0 untuk Belum dan 1 untuk Sudah sehingga skor total berada pada rentang 0-13. Penelitian ini tidak mengembangkan instrumen baru; pengujian diarahkan pada kualitas butir versi terjemahan ketika digunakan pada konteks wasit B2 di Jawa Barat. Ringkasan cakupan instrumen disajikan pada tabel 1.

Tabel 1. Area dan indikator Match Assessment

Area	Butir	Fokus pengamatan
Komunikasi	1-3	Peluit; sinyal/pelaporan; komunikasi pencegahan dan penjelasan
Mekanika perwasitan	4-6	Posisi lead; posisi trail; tanggung jawab area dan off-ball
Pemanggilan pertandingan	7-8	Keputusan violation; keputusan foul dan kontak ilegal
Pengendalian pertandingan	9	Komunikasi keputusan dan interaksi dengan pemain
Profesionalisme	10-11	Konsentrasi dan sikap proaktif
Administrasi pertandingan	12	Persiapan, meja, keamanan, dan penyelesaian
Peraturan	13	Pengetahuan aturan dan prinsip penempatan posisi

Penilaian dilakukan oleh satu asesor yang merupakan wasit berlisensi A, memiliki kualifikasi referee coach, dan menjabat Ketua Bidang Perwasitan DPD PERBASI Jawa Barat. Latar belakang tersebut dipilih karena asesor harus memahami peraturan, mekanika, standar kinerja, dan istilah teknis dalam instrumen. Penggunaan satu asesor menjaga sumber penilaian tetap sama untuk seluruh wasit, tetapi tidak memungkinkan pengujian kesepakatan antarpenilai.

Pengumpulan data diawali dengan izin kepada penyelenggara dan koordinasi dengan pihak perwasitan. Daftar wasit diperiksa berdasarkan kriteria sampel, kemudian asesor menerima instrumen dan petunjuk skor. Asesor mengamati setiap wasit dan memilih Belum atau Sudah pada 13 butir. Jawaban dikodekan 0 dan 1, sedangkan responden diberi kode R1-R30. Seluruh 390 sel penilaian terisi sehingga tidak diperlukan penanganan data hilang.

Analisis deskriptif mencakup rata-rata, simpangan baku, minimum, maksimum, frekuensi skor total, serta jumlah dan persentase skor Sudah untuk setiap butir. Validitas butir diuji dengan korelasi Pearson Product Moment antara skor butir dan skor total. Dengan n=30 dan df=28, r tabel pada taraf 0,05 adalah 0,361. Butir dinyatakan valid apabila r hitung lebih besar dari 0,361 dan nilai signifikansi lebih kecil dari 0,05. Butir yang tidak memenuhi salah satu syarat diklasifikasikan belum valid.

Reliabilitas internal seluruh 13 butir dihitung menggunakan Cronbach's Alpha. Kriteria yang ditetapkan dalam penelitian adalah alpha kurang dari 0,60 sebagai belum reliabel, 0,60-0,69 cukup reliabel, 0,70-0,79 reliabel, 0,80-0,89 sangat reliabel, dan 0,90 atau lebih sangat tinggi. Hipotesis nol menyatakan instrumen belum reliabel bila alpha kurang dari 0,60, sedangkan hipotesis alternatif diterima bila alpha sekurang-kurangnya 0,60.

HASIL DAN PEMBAHASAN

Skor total 30 wasit memiliki rata-rata 10,20 dari maksimum 13 dengan simpangan baku 2,14. Skor terendah adalah 6 dan skor tertinggi 13. Distribusi menunjukkan enam wasit (20,00%) memperoleh skor 13, tujuh wasit (23,33%) memperoleh skor 9, dan responden lain tersebar pada skor 6, 8, 10, 11, dan 12. Variasi ini menunjukkan bahwa asesor masih membedakan kinerja total di antara sampel meskipun beberapa butir memiliki proporsi pencapaian yang tinggi.

Tabel 2. Statistik deskriptif dan distribusi skor total

Ukuran	Nilai	Skor	Frekuensi	Persentase
N	30	6	3	10,00%
Jumlah butir	13	8	2	6,67%
Rata-rata	10,20	9	7	23,33%
Simpangan baku	2,14	10	4	13,33%
Minimum	6	11	5	16,67%
Maksimum	13	12	3	10,00%
		13	6	20,00%

Proporsi skor Sudah berbeda antarbutir. Butir 7 (keputusan violation) memiliki proporsi tertinggi, yaitu 28 dari 30 wasit (93,33%), sedangkan butir 4 (posisi lead) terendah, yaitu 12 wasit (40,00%). Butir 2, 11, dan

e-ISSN: 2963 – 198X
p-ISSN: 2963 – 2498

13 masing-masing mencapai 90,00%. Variasi yang sangat kecil pada butir berproporsi tinggi perlu diperhatikan karena dapat mengurangi kemampuan item membedakan kinerja antarwasit.

Tabel 3. Proporsi skor "Sudah" per butir

Butir	Indikator ringkas	Sudah (n)	Sudah (%)	SD
1	Peluit kuat/tegas	23	76,67%	0,430
2	Sinyal dan pelaporan	27	90,00%	0,305
3	Komunikasi pencegahan	20	66,67%	0,479
4	Posisi lead	12	40,00%	0,498
5	Posisi trail	20	66,67%	0,479
6	Tanggung jawab area	21	70,00%	0,466
7	Keputusan violation	28	93,33%	0,254
8	Keputusan foul	25	83,33%	0,379
9	Komunikasi keputusan	26	86,67%	0,346
10	Konsentrasi	26	86,67%	0,346
11	Sikap proaktif	27	90,00%	0,305
12	Administrasi pertandingan	24	80,00%	0,407
13	Pengetahuan peraturan	27	90,00%	0,305

Uji validitas menghasilkan tujuh butir yang memenuhi kedua kriteria, yaitu butir 1, 4, 5, 6, 9, 10, dan 11. Enam butir lain - 2, 3, 7, 8, 12, dan 13 - memiliki r hitung tidak lebih besar dari 0,361 dan/atau nilai signifikansi tidak lebih kecil dari 0,05. Korelasi tertinggi terdapat pada butir 5 (r=0,739; p<0,001), sedangkan korelasi terendah pada butir 2 (r=0,190; p=0,314). Butir 10 hanya sedikit melampaui batas r tabel, tetapi tetap memenuhi kriteria yang telah ditetapkan (r=0,364; p=0,048).

Tabel 4. Hasil uji validitas Pearson Product Moment

Butir	r hitung	Sig.	Status
1	0,390	0,033	Valid
2	0,190	0,314	Belum valid
3	0,302	0,104	Belum valid
4	0,569	0,001	Valid
5	0,739	<0,001	Valid
6	0,443	0,014	Valid
7	0,343	0,064	Belum valid
8	0,340	0,066	Belum valid
9	0,597	0,001	Valid
10	0,364	0,048	Valid
11	0,507	0,004	Valid
12	0,285	0,127	Belum valid
13	0,349	0,059	Belum valid

Catatan: Butir valid jika r hitung >0,361 dan Sig. <0,05 (n=30; df=28).

Cronbach's Alpha seluruh 13 butir adalah 0,610. Nilai tersebut lebih besar daripada batas 0,60 dan berada pada rentang 0,60-0,69, sehingga instrumen diklasifikasikan cukup reliabel berdasarkan pedoman penelitian. Dengan demikian, hipotesis nol ditolak dan hipotesis alternatif diterima. Hasil ini menunjukkan konsistensi internal yang cukup untuk penilaian awal, bukan bukti bahwa semua butir telah berfungsi sama baiknya.

Tabel 5. Hasil uji reliabilitas Cronbach's Alpha

Alpha	Jumlah butir	Batas	Tingkat
0,610	13	>=0,60	Cukup reliabel

Sumber: Data penelitian, diolah peneliti.

Pembahasan

Tujuh butir yang valid tidak terkumpul secara acak. Butir 4, 5, dan 6 mewakili mekanika perwasitan: posisi lead, posisi trail, serta tanggung jawab utama, sekunder, dan off-ball. Ketiga perilaku tersebut muncul berulang kali ketika penguasaan bola berpindah, sehingga asesor memperoleh kesempatan observasi pada berbagai situasi. Frekuensi kemunculan yang tinggi memungkinkan skor lebih mencerminkan perbedaan cara wasit bergerak, membuka sudut pandang, mempertahankan gap, dan menjaga area tanggung jawab. Temuan ini sejalan dengan gagasan bahwa keputusan yang tepat memerlukan posisi yang memungkinkan kejadian terlihat secara jelas; pengetahuan aturan saja tidak cukup ketika sudut pandang tertutup.

Butir 1, 9, 10, dan 11 juga dapat diamati langsung selama pertandingan. Kekuatan dan ketegasan peluit pada butir 1 terdengar setiap kali wasit menghentikan permainan. Komunikasi keputusan dan interaksi dengan

e-ISSN: 2963 – 198X
p-ISSN: 2963 – 2498

pemain pada butir 9 tampak melalui sinyal, bahasa tubuh, penjelasan singkat, dan respons terhadap reaksi pemain. Konsentrasi dan sikap proaktif pada butir 10 dan 11 berlangsung sepanjang pertandingan, termasuk saat pergantian, time-out, cedera, atau administrasi lain perlu ditangani. Kemunculan yang berulang memberi dasar penilaian lebih stabil daripada indikator yang hanya dapat terlihat pada situasi khusus. Pentingnya komunikasi, kesiapan mental, dan profesionalisme juga konsisten dengan temuan penelitian perwasitan sebelumnya (Arief & Nurcahya, 2020; Indriansah & Arief, 2020; Pambudi & Suharjana, 2018).

Enam butir yang belum valid tetap merepresentasikan kompetensi penting. Korelasi item-total yang rendah lebih tepat dibaca sebagai tanda bahwa bentuk pengukuran belum cukup membedakan sampel pada kondisi ini. Butir 7 memperoleh skor Sudah pada 93,33% wasit, sedangkan butir 2 dan 13 mencapai 90,00%. Ketika hampir semua subjek mendapat skor yang sama, variasi statistik mengecil dan hubungan dengan skor total menjadi lebih sulit terlihat. Keadaan ini mungkin dipengaruhi karakteristik wasit B2 yang telah menguasai prosedur dasar, bukan karena sinyal, violation, atau pengetahuan aturan tidak relevan. Penelitian pada lisensi atau pengalaman yang berbeda memang menunjukkan bahwa karakteristik wasit dapat memengaruhi variasi kinerja (Sanusi et al., 2022).

Butir 2 menilai sinyal dan pelaporan kepada meja. Keterampilan tersebut dilakukan rutin dan kemungkinan telah dikuasai sebagian besar sampel. Butir 3 menggabungkan preventative voice, penjelasan kepada pemain atau pelatih, dan pertukaran informasi dengan rekan wasit dalam satu skor. Seorang wasit dapat memperlihatkan salah satu perilaku, tetapi tidak selalu memperoleh kesempatan menunjukkan perilaku lain dalam pertandingan yang sama. Penggabungan beberapa indikator mengaburkan dasar keputusan Belum atau Sudah. Perbaikan dapat dilakukan dengan memisahkan indikator, menyediakan contoh perilaku, dan menetapkan aturan skor ketika hanya sebagian perilaku muncul.

Butir 7 dan 8 berhubungan dengan keputusan violation dan foul. Meskipun pengambilan keputusan adalah inti tugas wasit dan pemahaman kedua jenis pelanggaran perlu dikuasai (Rohmana & Herpandika, 2021), frekuensi situasi tidak sama pada setiap pertandingan. Wasit dapat tidak menjumpai jenis kejadian tertentu selama periode pengamatan. Penilaian satu pertandingan dengan skor dikotomis karena itu belum mampu memisahkan ketiadaan kesempatan dari ketidakmampuan. Penelitian selanjutnya dapat menambah jumlah pertandingan, mencatat kesempatan observasi, atau menggunakan klip video terstandar untuk menilai interpretasi keputusan.

Butir 12 merangkum persiapan, interaksi dengan petugas meja, keamanan, dan penyelesaian pertandingan. Rentang perilaku dari sebelum hingga sesudah pertandingan terlalu luas untuk direpresentasikan oleh satu skor 0 atau 1 tanpa kriteria tambahan. Butir 13 menggabungkan pengetahuan peraturan dengan penjelasan prinsip penempatan posisi. Pengetahuan tidak seluruhnya terlihat dalam observasi karena tidak semua aturan muncul di lapangan. Butir ini lebih kuat apabila observasi dilengkapi tes aturan tertulis atau analisis video. Dengan demikian, rekomendasi yang tepat adalah menelaah isi dan redaksi, bukan langsung menghapus dua area yang penting bagi tanggung jawab wasit.

Alpha 0,610 berada tepat di atas batas minimal yang ditetapkan. Nilai ini mendukung penggunaan instrumen sebagai penilaian awal, tetapi belum menunjukkan konsistensi yang kuat. Cronbach (1951) menjelaskan koefisien alpha sebagai ukuran konsistensi internal; besarnya dipengaruhi jumlah butir dan korelasi antarbutir. Match Assessment mencakup tujuh area yang secara konseptual berhubungan sebagai kinerja wasit, tetapi komunikasi, mekanika, pemanggilan, pengendalian, profesionalisme, administrasi, dan pengetahuan aturan bukan perilaku yang identik. Instrumen multidimensional dapat menghasilkan hubungan antarbutir yang moderat. Karena itu, alpha perlu ditafsirkan bersama validitas butir dan cakupan konstruk, bukan sebagai satu-satunya bukti kualitas (Boateng et al., 2018).

Skala Belum-Sudah memberi kepraktisan ketika asesor mengikuti pertandingan yang cepat, tetapi mengorbankan sensitivitas. Dua wasit yang sama-sama memperoleh skor Sudah dapat berbeda dalam frekuensi, kualitas, atau konsistensi perilakunya. Penggunaan kategori bertahap dengan deskriptor yang jelas dapat meningkatkan variasi, tetapi perubahan skala harus diuji kembali dan tidak boleh dilakukan hanya untuk menaikkan alpha. Pengembangan alat ukur olahraga sebelumnya menunjukkan perlunya prosedur pengujian berulang dan bukti yang sesuai dengan tujuan penggunaan (Alim & Nurfadhila, 2021; Guntur et al., 2020; Setyawan & Sugiyanto, 2016; Sulistiyono et al., 2022; Tomoliyus & Sunardianta, 2020).

Penggunaan satu asesor yang kompeten membuat semua wasit dinilai dengan sumber standar yang sama, tetapi reliabilitas antarasesor tidak dapat dihitung. Pengalaman, posisi pengamatan, dan interpretasi terhadap kategori Belum dan Sudah dapat menghasilkan perbedaan ketika penilai berganti. Hasil penelitian karena itu tidak menunjukkan bahwa instrumen akan menghasilkan skor identik pada penilai lain. Pengujian berikutnya perlu melibatkan dua atau lebih asesor, pelatihan kalibrasi, dan pedoman contoh perilaku. Rekaman video yang sama dapat digunakan untuk memeriksa kesepakatan sebelum asesor menilai pertandingan secara langsung.

e-ISSN: 2963 – 198X
p-ISSN: 2963 – 2498

Proses terjemahan oleh penerjemah tersumpah memperkuat ketepatan formal bahasa, tetapi adaptasi instrumen juga menuntut kesetaraan konsep dan konteks (Beaton et al., 2000). Istilah teknis perwasitan dapat tetap dipertahankan dalam bahasa Inggris apabila telah menjadi bahasa kerja komunitas, disertai definisi operasional dan contoh. Telaah panel ahli dari referee coach, asesor, pengurus perwasitan, dan ahli pengukuran dapat menilai apakah redaksi versi Indonesia tetap mencakup kompetensi asal. Sesudah revisi, uji empiris perlu dilakukan pada sampel lebih besar, beberapa daerah, beberapa tingkat lisensi, lebih dari satu pertandingan, dan lebih dari satu asesor.

Secara praktis, tujuh butir valid dapat dipertahankan pada tahap berikutnya, sedangkan enam butir lain diperbaiki dan diuji ulang. Seluruh tujuh area tetap dapat digunakan sebagai kerangka diskusi antara wasit dan mentor, tetapi skor 0-13 tidak sebaiknya menjadi satu-satunya dasar kelulusan, promosi lisensi, atau penugasan. Hasil observasi perlu disertai catatan kejadian dan umpan balik, karena tujuan yang paling sesuai dengan bukti saat ini adalah diagnosis awal untuk pembinaan. Pembatasan ini mencegah keputusan berisiko tinggi dibuat dari instrumen yang reliabilitasnya masih cukup dan hampir separuh butirnya belum memenuhi kriteria validitas.

Penelitian ini terbatas pada 30 wasit B2 dari satu kegiatan, satu asesor, dan data yang belum mencatat jumlah pertandingan serta lama pengamatan setiap wasit. Pengujian juga hanya mencakup korelasi item-total dan Cronbach's Alpha. Oleh sebab itu, simpulan berlaku pada kondisi POPWIL Jawa Barat tahun 2026 dan belum dapat digeneralisasi ke semua wasit Indonesia. Tahap selanjutnya perlu memasukkan validitas isi oleh ahli, struktur dimensi, kesepakatan antarasesor, stabilitas lintas pertandingan, serta hubungan dengan kriteria kinerja lain sebelum Match Assessment digunakan secara luas.

KESIMPULAN

Pada 30 wasit level B2 yang bertugas dalam POPWIL Jawa Barat, tujuh dari 13 butir Match Assessment versi bahasa Indonesia memenuhi kriteria validitas Pearson. Butir valid adalah 1, 4, 5, 6, 9, 10, dan 11, sedangkan butir 2, 3, 7, 8, 12, dan 13 belum memenuhi kriteria r hitung >0,361 dan Sig. <0,05. Cronbach's Alpha sebesar 0,610 menempatkan instrumen pada tingkat cukup reliabel berdasarkan pedoman penelitian dan mendukung penerimaan hipotesis alternatif.

Instrumen dapat digunakan sebagai bahan penilaian awal dan kerangka umpan balik, tetapi belum sebaiknya menjadi satu-satunya dasar kelulusan, promosi lisensi, atau penugasan. Enam butir yang belum valid perlu diperjelas, diberi contoh perilaku dan petunjuk skor, lalu diuji kembali. Pengujian lanjutan memerlukan sampel yang lebih besar dan beragam, beberapa pertandingan, dua atau lebih asesor, serta pemeriksaan validitas isi dan reliabilitas antarpenilai. Simpulan ini dibatasi oleh sampel 30 wasit B2, satu kegiatan, dan penilaian oleh satu asesor.

Ucapan terima kasih

Penulis mengucapkan terima kasih kepada penyelenggara POPWIL Jawa Barat, DPD PERBASI Jawa Barat, asesor penelitian, dan seluruh wasit B2 yang berpartisipasi. Dukungan koordinasi, penilaian, dan ketersediaan data memungkinkan penelitian ini diselesaikan.

REFERENSI

Alim, A., & Nurfadhila, R. (2021). Reliabilitas test-retest Dyer Tennis Test revision. Jurnal Keolahragaan, 9(1), 43–50. https://doi.org/10.21831/jk.v9i1.39122

Arief, R., & Nurcahya, Y. (2020). Hubungan antara kebugaran, kecerdasan intelektual, dan kecerdasan emosional dengan kinerja wasit futsal. Jurnal Kepelatihan Olahraga, 12(2), 67–73.

Basketball Australia. (n.d.). Intermediate referee workbook: Match assessment 1.

Beaton, D. E., Bombardier, C., Guillemin, F., & Ferraz, M. B. (2000). Guidelines for the process of cross-cultural adaptation of self-report measures. Spine, 25(24), 3186–3191. https://doi.org/10.1097/00007632-200012150-00014

Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, 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, Article 149. https://doi.org/10.3389/fpubh.2018.00149

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. Psychometrika, 16, 297–334. https://doi.org/10.1007/BF02310555

FIBA. (2024). Official basketball rules 2024 (Version 1.0a).

Guntur, G., Ngatman, N., Sridadi, S., Broto, D. P., & Pambudi, D. K. (2020). Pengujian validitas, reliabilitas, dan relevansi norma penilaian Dyer Tennis Test terhadap tingkat keterampilan bermain tenis. Jurnal

e-ISSN: 2963 – 198X
p-ISSN: 2963 – 2498

Keolahragaan, 8(2), 195–203. https://doi.org/10.21831/jk.v8i2.35266

Indriansah, D., & Arief, R. (2020). Hubungan antara kondisi fisik, kesiapan mental, dan pemahaman peraturan dengan kinerja wasit. Jurnal Kepelatihan Olahraga, 12(2), 89–94. https://doi.org/10.17509/jko-upi.v12i2.27037

Margiyan, M. E., Nuraini, S., & Hernawan. (2022). The effect of referee knowledge, leadership style and confidence on the performance of hockey referee. Gladi: Jurnal Ilmu Keolahragaan, 13(2), 161–177.

Pambudi, D. K., & Suharjana. (2018). Pengembangan video pembelajaran sinyal-sinyal wasit bola basket untuk guru penjasorkes sekolah menengah atas. Jurnal Keolahragaan, 6(1), 48–59. https://doi.org/10.21831/jk.v6i1.11880

Prakasa, D. A., Ma'mun, A., & Budiana, D. (2019). Program pelatihan wasit bola basket tingkat pemula: Sebuah studi eksperimen. Jurnal Penelitian Pendidikan, 19(2), 190–200. https://doi.org/10.17509/jpp.v19i2.19762

Rohmana, V. D., & Herpandika, R. P. (2021). Pemahaman peraturan foul and violation wasit bola basket Universitas Nusantara PGRI Kediri. Jurnal Kejaora, 6(2), 259–264. https://doi.org/10.36526/kejaora.v6i2.1540

Rosyadi, H., Mulyana, M., & Mulyana, D. (2017). Hubungan tingkat kebugaran jasmani dengan rasa percaya diri wasit bola basket. Jurnal Kepelatihan Olahraga, 10(2), 57–73. https://doi.org/10.17509/jko-upi.v10i2.16142

Sanusi, A., Sobarna, A., & Rizal, R. M. (2022). Hubungan klasifikasi wasit dan pengalaman menjadi wasit dengan kinerja wasit bola voli di Jawa Barat. Jurnal Master Penjas & Olahraga, 3(1), 178–185. https://doi.org/10.37742/jmpo.v3i1.51

Setyawan, T. A., & Sugiyanto, F. (2016). Pengembangan tes kemampuan fisik bolabasket siswa SMA putra di Provinsi Daerah Istimewa Yogyakarta. Jurnal Keolahragaan, 4(1), 85–97. https://doi.org/10.21831/jk.v4i1.6415

Sulistiyono, S., Suherman, W. S., Pambudi, D. K., & Martono, M. (2022). Validitas dan reliabilitas instrumen penilaian pengelolaan sekolah sepakbola berbasis sistem manajemen mutu. Medikora, 21(1), 61–70. https://doi.org/10.21831/medikora.v21i1.48829

Tomoliyus, & Sunardianta, R. (2020). Validitas dan reliabilitas instrumen tes reaktif agility tenis meja. Jurnal Keolahragaan, 8(2), 148–157. https://doi.org/10.21831/jk.v8i2.32492