



The Content Validity of the Anxiety Scale Using Aiken's V Approach

Yatny Mariam Lestari*¹, Leni Anjani², Hermansyah³, Retno Dwiyaniti⁴

* yatni.mariam@gmail.com

Fakultas Psikologi, Universitas Muhammadiyah Purwokerto, Banyumas, Indonesia

ABSTRACT

Anxiety is one of the most common mental disorders experienced by individuals across various populations. The availability of anxiety measurement instruments with strong content validity in the context of the general population in Indonesia remains limited. This study aims to examine the content validity of an anxiety scale using the Aiken's V approach. The scale consists of 30 items covering four main dimensions: cognitive, affective, physiological, and behavioral. Content validity assessment was conducted by four expert panelists. The analytical technique used was Aiken's V content validity coefficient, with a score range of 0 to 1. The results indicate that all 30 items have V values ranging from 0.813 to 1.00. All items (100%) obtained V values greater than 0.80, indicating very high validity. The average V value across all items is 0.94. Therefore, the anxiety scale is considered to have excellent content validity based on expert judgment. The implication of this study is that the instrument can be used for the early identification of anxiety in the general population in Indonesia.

Keywords: Aiken's V; anxiety; general population; content validity

PENDAHULUAN

Kecemasan merupakan salah satu gangguan mental yang paling umum ditemukan di berbagai populasi di seluruh dunia (Hang et al., 2022; Laranjeira et al., 2023). Generalized anxiety disorder (GAD) adalah bentuk kecemasan yang paling umum, yang ditandai dengan kekhawatiran yang berlebihan dan terus-menerus. Prevalensi gangguan kecemasan diperkirakan mencapai 20% pada populasi muda di berbagai negara (Hang et al., 2022). Selama pandemi COVID-19, prevalensi gejala psikopatologis meningkat tiga kali lipat di beberapa negara, dengan tingkat kecemasan yang cukup tinggi terdeteksi di kalangan mahasiswa pendidikan tinggi (Laranjeira et al., 2023).

Dampak negatif dari kecemasan sangat luas, mencakup penurunan kualitas hidup, gangguan fungsi sosial, penurunan produktivitas, serta meningkatnya risiko komorbiditas dengan gangguan mental lain seperti depresi. Selain itu, kecemasan yang tidak tertangani dengan baik dapat menyebabkan gangguan tidur kronis, peningkatan risiko penyalahgunaan alkohol, serta gangguan fungsi kognitif (Hang et al., 2022). Studi lain juga mengkonfirmasi bahwa kecemasan berkorelasi signifikan dengan gejala depresi, stres yang dirasakan, gangguan tidur, dan ketidakpuasan hidup (Zhang et al., 2021). Gejala kecemasan juga ditemukan lebih tinggi pada individu dengan gangguan psikiatri dibandingkan dengan populasi umum (Ali et al., 2021).

Tantangan utama dalam upaya deteksi dini kecemasan adalah ketersediaan instrumen ukur yang valid dan reliabel. Meskipun Depression Anxiety Stress Scales (DASS-21) dan Generalized Anxiety Disorder Scale (GAD-7) telah banyak digunakan dan menunjukkan properti psikometri yang baik di berbagai populasi, namun instrumen-instrumen tersebut dikembangkan dalam konteks budaya Barat. Perbedaan budaya dapat mempengaruhi cara individu memahami dan melaporkan gejala kecemasan, sehingga diperlukan upaya validasi lintas budaya atau pengembangan skala yang sesuai dengan konteks lokal, termasuk di Indonesia.

Beberapa penelitian terbaru telah mengembangkan dan memvalidasi berbagai instrumen psikologi untuk populasi dan konteks yang beragam. (Xu et al., 2025) mengembangkan Patient-Reported Outcome (PRO) scale untuk pasien dengan herniasi diskus lumbalis (LDH) yang mencakup lima dimensi, yaitu gejala fisik, disfungsi fisik, emosi psikologis, adaptasi sosial, dan efek terapeutik. Studi ini menggunakan analisis faktor eksploratori (EFA) dan paralel analysis untuk menentukan struktur faktor, serta menguji reliabilitas dengan Cronbach's alpha (0,95) dan validitas kriteria dengan korelasi terhadap Oswestry Disability Index (0,74) dan Self-Rating Anxiety Scale (0,62). Namun, studi ini memiliki keterbatasan karena sampel hanya diambil dari satu rumah sakit besar di perkotaan China, sehingga hasilnya tidak dapat digeneralisasi ke pasien di daerah pedesaan atau fasilitas kesehatan yang lebih kecil (Xu et al., 2025, hlm. 15).

(Wu et al., 2023) mengembangkan dan memvalidasi Climate Change Anxiety Scale-Short Form (CCAS-S) untuk remaja Kanada. Studi ini melibatkan konsultasi dengan remaja sebagai ahli berdasarkan pengalaman hidup (*lived experience experts*) untuk mengadaptasi skala original, serta menggunakan confirmatory factor analysis (CFA) yang mendukung struktur satu faktor dengan internal consistency yang sangat baik (Cronbach's alpha = 0,95). Meskipun demikian, studi ini hanya menggunakan laporan diri (*self-report*) yang rentan terhadap bias respons, dan desain potong lintang (*cross-sectional*) tidak memungkinkan penilaian temporalitas hubungan antar variabel.

(Kurniawati et al., 2025) memvalidasi Pregnancy-Related Anxiety Scale (PrAS) untuk ibu hamil di Indonesia. Studi ini mengevaluasi validitas isi menggunakan Content Validity Index (CVI) dan Content Validity Ratio (CVR) yang melibatkan enam ahli, serta menguji validitas konstruk dengan CFA yang menghasilkan model tujuh dimensi (setelah menghilangkan satu dimensi yang tidak sesuai). Namun, penelitian ini hanya dilakukan di Jakarta (daerah perkotaan) dengan ukuran sampel yang relatif kecil (n=92), sehingga temuan belum dapat digeneralisasi ke seluruh wilayah Indonesia, khususnya daerah pedesaan.

(Nguyen et al., 2025) melakukan adaptasi dan validasi konten toolkit untuk mengukur literasi kesehatan mental guru di Vietnam menggunakan metode *expert panel*. Studi ini menggunakan CVI dan CVR dengan melibatkan enam ahli, dan menghasilkan 82 item final dari 103 item awal setelah proses validasi. Namun, studi ini baru mengevaluasi validitas isi dan belum memberikan indikator lain seperti reliabilitas atau validitas konstruk, sehingga diperlukan penelitian lebih lanjut untuk konfirmasi validitas dan reliabilitas secara penuh.

(Jeldres & Costa, 2023) melakukan tinjauan terhadap metode Lawshe untuk menghitung validitas isi di ilmu sosial. Studi ini membandingkan berbagai pendekatan statistik untuk menentukan nilai kritis Content Validity Ratio (CVR), termasuk metode *exact binomial test*, *chi-squared test*, dan *Bayesian statistics*. Hasilnya menunjukkan bahwa metode Bayesian dapat mengurangi jumlah panelis yang dibutuhkan sambil tetap menghasilkan nilai CVR kritis yang dapat diterima. Namun, kritik utama terhadap metode Lawshe adalah tidak adanya uji statistik yang jelas dalam proposal aslinya, serta ketergantungan nilai kritis pada jumlah panelis yang tidak konsisten antar peneliti

Meskipun berbagai studi terbaru telah mengembangkan dan memvalidasi instrumen

kecemasan pada beragam populasi, masih terdapat kesenjangan yang signifikan baik secara teoretis maupun empiris. Secara teoretis, (Jeldres & Costa, 2023) mengkritisi bahwa metode Lawshe untuk menghitung validitas isi tidak mencantumkan uji statistik yang jelas dalam proposal aslinya, serta nilai kritis Content Validity Ratio (CVR) sangat bergantung pada jumlah panelis yang tidak konsisten antar peneliti. Beberapa pendekatan alternatif seperti *exact binomial test*, *chi-squared test*, dan *Bayesian statistics* telah diusulkan oleh berbagai peneliti sebagaimana dirangkum oleh (Jeldres & Costa, 2023), namun masing-masing memiliki kelemahan tersendiri. Kesenjangan teoretis ini menunjukkan bahwa belum ada konsensus mengenai metode baku untuk menilai validitas isi, khususnya untuk skala kecemasan yang akan digunakan di populasi umum Indonesia.

Secara empiris, studi-studi validasi instrumen kecemasan yang ada masih memiliki keterbatasan dalam cakupan populasi dan metodologi. (Xu et al., 2025) mengembangkan PRO scale untuk pasien LDH tetapi hanya berfokus pada populasi klinis dengan kondisi medis spesifik, sehingga tidak dapat digeneralisasi ke populasi umum yang mengalami kecemasan tanpa kondisi medis penyerta. (Wu et al., 2023) memvalidasi skala kecemasan iklim untuk remaja Kanada, namun instrumen ini secara spesifik mengukur kecemasan yang terkait dengan perubahan iklim, bukan kecemasan umum. (Kurniawati et al., 2025) memvalidasi Pregnancy-Related Anxiety Scale (PrAS) di Indonesia, tetapi hanya berfokus pada ibu hamil di daerah perkotaan dengan sampel yang relatif kecil. (Nguyen et al., 2025) hanya menguji validitas isi toolkit literasi kesehatan mental guru tanpa menguji reliabilitas atau validitas konstruk, sehingga belum dapat memastikan kualitas psikometri instrumen secara menyeluruh. Dengan demikian, belum ada studi yang secara khusus mengembangkan dan memvalidasi skala kecemasan umum untuk populasi dewasa di Indonesia dengan menggunakan pendekatan validitas isi Aiken's V yang melibatkan panel ahli multidisiplin.

Berdasarkan kesenjangan teoretis dan empiris yang telah diuraikan, penelitian ini bertujuan untuk menjawab pertanyaan penelitian: Bagaimana validitas isi skala kecemasan untuk populasi umum di Indonesia berdasarkan penilaian panel ahli menggunakan pendekatan Aiken's V? Secara lebih spesifik, penelitian ini bertujuan untuk: (1) mengevaluasi tingkat kesepakatan antar panel ahli terhadap relevansi setiap butir skala kecemasan; (2) menghitung nilai koefisien Aiken's V untuk masing-masing butir skala; serta (3) menentukan butir-butir mana saja yang memenuhi kriteria validitas isi yang dapat diterima ($V \geq 0,80$) dan layak dipertahankan dalam skala kecemasan final. Kebaruan (*novelty*) dari penelitian ini terletak pada beberapa hal. Pertama, penelitian ini mengembangkan skala kecemasan yang secara spesifik ditujukan untuk populasi umum di Indonesia, berbeda dengan studi sebelumnya yang lebih berfokus pada populasi klinis (Xu et al., 2025), populasi remaja dengan konteks spesifik (Wu et al., 2023), atau populasi ibu hamil (Kurniawati et al., 2025). Kedua, penelitian ini menggunakan pendekatan Aiken's V yang memiliki keunggulan dalam menilai kesepakatan antar penilai ahli untuk skala dengan kategori respons lebih dari dua pilihan, berbeda dengan metode CVR/CVI yang banyak digunakan dalam studi validasi sebelumnya (Nguyen et al., 2025; Kurniawati et al., 2025). Ketiga, penelitian ini melibatkan panel ahli dari berbagai latar belakang keilmuan untuk memastikan skala yang dikembangkan komprehensif dalam mencakup seluruh dimensi kecemasan, yaitu dimensi kognitif, afektif, fisiologis, dan perilaku. Hasil dari penelitian ini diharapkan dapat memberikan kontribusi bagi pengembangan instrumen psikologi di Indonesia, khususnya dalam pengukuran kecemasan pada populasi umum, serta menjadi landasan bagi uji validitas dan reliabilitas lebih lanjut.

METODE PENELITIAN

Penelitian ini menggunakan desain metodologis (*methodological research*) yang bertujuan untuk mengembangkan dan menguji validitas isi skala kecemasan. Desain ini dipilih

karena penelitian berfokus pada evaluasi sifat psikometrik instrumen, khususnya validitas isi, sebelum instrumen digunakan secara luas pada populasi target (Boateng et al., 2018; Kurniawati et al., 2025; Xu et al., 2025). Menurut (Boateng et al., 2018), pengembangan skala terdiri dari tiga fase utama: pengembangan item, pengembangan skala, dan evaluasi skala, sementara (Gupta et al., 2025) menekankan bahwa validitas isi merupakan langkah fundamental yang menghubungkan konsep abstrak dengan matriks yang terukur.

Penelitian ini dilaksanakan dalam empat tahap. Tahap pertama adalah pengembangan konseptual skala kecemasan berdasarkan kajian literatur dan definisi operasional konstruk kecemasan yang mencakup empat dimensi: kognitif, afektif, fisiologis, dan perilaku. Tahap kedua adalah penyusunan kisi-kisi dan 30 butir pernyataan skala kecemasan, yang terbagi menjadi pernyataan *favorable* dan *unfavorable*. Tahap ketiga adalah penilaian validitas isi oleh panel ahli yang terdiri dari 4 orang penilai. Para ahli diminta menilai relevansi setiap butir terhadap konstruk kecemasan menggunakan skala Likert 1-5. Tahap keempat adalah analisis data menggunakan koefisien Aiken's V untuk menghitung tingkat kesepakatan antar penilai. Penelitian ini tidak melibatkan uji coba lapangan pada responden karena fokus utamanya adalah validitas isi sebagai langkah awal dan fundamental dalam pengembangan instrumen, sebagaimana direkomendasikan dalam literatur pengembangan skala (Boateng et al., 2018; Gupta et al., 2025).

2. Populasi, Sampel, atau Subjek Penelitian

Populasi dalam penelitian ini adalah panel ahli yang memiliki kompetensi untuk menilai relevansi butir-butir skala kecemasan. Sampel diambil menggunakan teknik *purposive sampling* dengan kriteria: (1) memiliki latar belakang pendidikan minimal magister di bidang psikologi, psikometri, kedokteran jiwa, atau profesi terkait; (2) memiliki pengalaman minimal 5 tahun di bidang kesehatan mental, psikologi klinis, atau pengembangan instrumen; (3) memahami konstruk kecemasan beserta dimensi-dimensinya (kognitif, afektif, fisiologis, dan perilaku); serta (4) bersedia berpartisipasi secara sukarela.

Jumlah sampel yang berhasil dihimpun dalam penelitian ini adalah 4 orang panel ahli. Penentuan jumlah panel ahli mengacu pada berbagai penelitian validitas isi terkini. (Helmiyatun et al., 2025) dalam penelitian validitas isi instrumen persepsi siswa terhadap manajemen pendidikan menggunakan 4 orang ahli (rater) dan berhasil memperoleh nilai Aiken's V yang bervariasi dari 0,50 hingga 1,00, yang membuktikan bahwa jumlah 4 ahli masih dapat menghasilkan penilaian yang andal. Sementara itu, (Saryono et al., 2025) dalam pengembangan instrumen kesiapsiagaan bencana gempa bumi menggunakan 5 orang ahli dan menghasilkan nilai Aiken's V di atas 0,87 untuk seluruh item, yang menunjukkan validitas isi yang sangat baik. (Sabaruddin et al., 2022) menggunakan 9 orang ahli untuk memvalidasi instrumen beasiswa KIP-K dan memperoleh nilai rata-rata Aiken's V sebesar 0,95 dengan kategori sangat valid. Dengan demikian, meskipun jumlah panel ahli bervariasi antara 4 hingga 9 orang, keempat penelitian tersebut membuktikan bahwa validitas isi yang baik dapat dicapai selama para ahli dipilih berdasarkan kriteria yang ketat dan relevan dengan konstruk yang diukur.

Keempat panel ahli dalam penelitian ini berasal dari latar belakang keilmuan yang relevan dengan pengukuran kecemasan. Dalam file Excel yang digunakan, keempat panel ahli diberi kode sebagai Penilai I, II, III, dan IV. Penelitian ini tidak melibatkan responden dari populasi umum karena fokus utama adalah validitas isi sebagai langkah awal pengembangan instrumen, yang hanya memerlukan penilaian ahli, bukan uji coba lapangan.

3. Teknik Pengumpulan Data dan Pengembangan Instrumen

Teknik pengumpulan data dilakukan melalui lembar validasi (*validation sheet*) yang diberikan kepada keempat panel ahli. Lembar validasi berisi 30 butir pernyataan skala kecemasan akademik yang telah disusun berdasarkan indikator kecemasan akademik (gejala

kognitif, gejala afektif, gejala fisiologis dan gejala perilaku/tindakan). Setiap ahli diminta memberikan penilaian menggunakan skala Likert 5 poin (1 = sangat tidak relevan, 2 = tidak relevan, 3 = cukup relevan, 4 = relevan, 5 = sangat relevan) terhadap tiga aspek penilaian: (a) kesesuaian butir dengan definisi operasional kecemasan akademik, (b) kejelasan redaksi butir, serta (c) keterbacaan butir untuk populasi umum .

Proses pengembangan instrumen mengacu pada prosedur baku penyusunan skala psikologi yang meliputi: (1) perumusan definisi operasional kecemasan berdasarkan kajian teori; (2) penyusunan kisi-kisi indikator; (3) penulisan butir pernyataan; (4) penilaian oleh panel ahli (*expert judgment*); dan (5) analisis validitas isi menggunakan Aiken's V .

4. Teknik Analisis Data

Data hasil penilaian dari keempat panel ahli dianalisis menggunakan koefisien validitas isi Aiken's V (Aiken, 1985) dengan rumus sebagai berikut :

$$V = \sum S / [n(c-1)]$$

Keterangan:

V = indeks validitas Aiken

S = r - lo, di mana r = skor yang diberikan ahli dan lo = skor terendah dalam kategori penilaian (dalam penelitian ini lo = 1)

n = jumlah panel ahli (n = 4)

c = angka penilaian tertinggi yang digunakan (c = 5)

Interpretasi nilai V mengacu pada kriteria yang dikemukakan oleh Aiken (1985) dan diperkuat oleh penelitian-penelitian terbaru, yaitu V > 0,80 dikategorikan sebagai validitas isi sangat baik, V = 0,60–0,80 cukup baik, dan V < 0,60 kurang baik . Seluruh perhitungan dilakukan menggunakan bantuan perangkat lunak *Microsoft Excel* untuk memastikan akurasi hasil.

Proses analisis dilakukan secara sistematis: (1) menghitung nilai S untuk setiap butir dari keempat ahli, (2) menjumlahkan seluruh nilai S ($\sum S$), (3) menghitung nilai $n(c-1) = 4 \times (5-1) = 16$, serta (4) membagi $\sum S$ dengan 16 untuk memperoleh nilai V setiap butir. Butir-butir yang memiliki nilai V < 0,60 direkomendasikan untuk direvisi atau dieliminasi, sedangkan butir dengan V \geq 0,60 dipertahankan untuk uji coba lapangan selanjutnya.

HASIL DAN PEMBAHASAN

Penelitian ini bertujuan untuk menguji validitas isi skala kecemasan menggunakan pendekatan Aiken's V. Skala kecemasan yang dikembangkan terdiri dari 30 butir pernyataan yang mencakup empat dimensi utama, yaitu dimensi kognitif (pikiran), dimensi afektif (perasaan), dimensi fisiologis (tubuh), dan dimensi perilaku (tindakan). Penilaian validitas isi dilakukan oleh 4 orang panel ahli yang memiliki kompetensi di bidangnya. Para ahli diminta menilai relevansi setiap butir terhadap konstruk kecemasan menggunakan skala Likert 1-5 (1 = sangat tidak relevan, 5 = sangat relevan). Hasil perhitungan koefisien Aiken's V untuk setiap butir disajikan pada Tabel 1 berikut.

Tabel 1. Hasil Validitas Isi Aiken's V Skala Kecemasan

| Butir | Skor Penilai (I-IV) | $\sum S$ | $n(c-1)$ | V | Keterangan |
|--------|---------------------|----------|----------|-------|--------------|
| butir1 | 5,4,4,5 | 14 | 16 | 0,875 | Sangat Valid |
| butir2 | 5,3,5,5 | 13 | 16 | 0,813 | Sangat Valid |
| butir3 | 5,4,4,5 | 14 | 16 | 0,875 | Sangat Valid |

| Butir | Skor Penilai (I-IV) | $\sum S$ | $n(c-1)$ | V | Keterangan |
|---------|---------------------|----------|----------|-------|--------------|
| butir4 | 5,4,5,5 | 15 | 16 | 0,938 | Sangat Valid |
| butir5 | 5,5,4,5 | 15 | 16 | 0,938 | Sangat Valid |
| butir6 | 5,5,4,5 | 15 | 16 | 0,938 | Sangat Valid |
| butir7 | 5,4,5,5 | 15 | 16 | 0,938 | Sangat Valid |
| butir8 | 4,4,5,5 | 14 | 16 | 0,875 | Sangat Valid |
| butir9 | 5,4,4,5 | 14 | 16 | 0,875 | Sangat Valid |
| butir10 | 5,5,4,5 | 15 | 16 | 0,938 | Sangat Valid |
| butir11 | 5,4,5,5 | 15 | 16 | 0,938 | Sangat Valid |
| butir12 | 5,5,5,5 | 16 | 16 | 1,000 | Sangat Valid |
| butir13 | 5,4,5,5 | 15 | 16 | 0,938 | Sangat Valid |
| butir14 | 5,3,5,5 | 13 | 16 | 0,813 | Sangat Valid |
| butir15 | 5,4,5,5 | 15 | 16 | 0,938 | Sangat Valid |
| butir16 | 5,5,3,5 | 14 | 16 | 0,875 | Sangat Valid |
| butir17 | 5,5,5,5 | 16 | 16 | 1,000 | Sangat Valid |
| butir18 | 5,4,5,5 | 15 | 16 | 0,938 | Sangat Valid |
| butir19 | 5,5,5,5 | 16 | 16 | 1,000 | Sangat Valid |
| butir20 | 5,4,5,5 | 15 | 16 | 0,938 | Sangat Valid |
| butir21 | 5,4,5,5 | 15 | 16 | 0,938 | Sangat Valid |
| butir22 | 5,4,5,5 | 15 | 16 | 0,938 | Sangat Valid |
| butir23 | 5,4,5,5 | 15 | 16 | 0,938 | Sangat Valid |
| butir24 | 5,4,5,5 | 15 | 16 | 0,938 | Sangat Valid |
| butir25 | 5,5,5,5 | 16 | 16 | 1,000 | Sangat Valid |
| butir26 | 5,5,5,5 | 16 | 16 | 1,000 | Sangat Valid |

| Butir | Skor Penilai (I-IV) | $\sum S$ | $n(c-1)$ | V | Keterangan |
|---------|---------------------|----------|----------|-------|--------------|
| butir27 | 5,4,5,5 | 15 | 16 | 0,938 | Sangat Valid |
| butir28 | 5,5,5,5 | 16 | 16 | 1,000 | Sangat Valid |
| butir29 | 5,5,4,5 | 15 | 16 | 0,938 | Sangat Valid |
| butir30 | 5,4,5,5 | 15 | 16 | 0,938 | Sangat Valid |

Rata-rata V 0,94 Sangat Valid

Berdasarkan Tabel 1, seluruh 30 butir skala kecemasan akademik memperoleh nilai V di atas 0,80. Secara rinci, terdapat 6 butir (20%) dengan nilai V = 1,00 (sangat valid sempurna), yaitu butir 12, 17, 19, 25, 26, dan 28. Sebanyak 22 butir (73,3%) memiliki nilai V antara 0,875–0,938 (sangat valid), dan 2 butir (6,7%) memiliki nilai V = 0,813 (sangat valid). Tidak ada satu pun butir yang memiliki nilai V di bawah 0,80. Nilai rata-rata V seluruh butir adalah 0,94 yang termasuk dalam kategori sangat baik.

Pembahasan

Interpretasi Hasil Validitas Isi

Hasil penelitian menunjukkan bahwa skala kecemasan akademik yang dikembangkan memiliki validitas isi yang sangat baik berdasarkan penilaian 4 orang panel ahli. Seluruh 30 butir memperoleh nilai Aiken's V > 0,80 dengan rata-rata 0,94. Temuan ini sejalan dengan kriteria yang ditetapkan (Aiken, 1985), di mana untuk jumlah panelis sebanyak 4 orang, koefisien V di atas 0,80 umumnya dianggap sebagai indikator validitas konten yang sangat baik. Dengan rata-rata V sebesar 0,94, seluruh 30 butir pernyataan dalam skala ini memenuhi standar validitas isi yang dipersyaratkan.

Keunggulan Butir dengan Nilai V Sempurna

Keenam butir yang memperoleh nilai V sempurna (1,00) merupakan butir-butir yang secara eksplisit menggambarkan manifestasi fisiologis dan kognitif dari kecemasan. Secara metodologis, nilai Aiken's V = 1,00 mengindikasikan adanya kesepakatan penuh (*perfect agreement*) di antara seluruh panel ahli bahwa butir tersebut sangat relevan dengan konstruk yang diukur (Aiken, 1985). Kesepakatan penuh ini menunjukkan bahwa indikator-indikator tersebut memiliki kejelasan konseptual yang tinggi dan dipersepsikan secara seragam oleh para ahli sebagai penciri utama kecemasan.

Butir-butir dengan nilai V sempurna mencakup indikator fisiologis seperti jantung berdebar, sesak napas, keringat dingin, dan gemetar, serta indikator kognitif seperti kekhawatiran berlebihan dan membayangkan skenario terburuk. Kesepakatan penuh para ahli terhadap butir-butir ini memperkuat bahwa dimensi fisiologis dan kognitif adalah aspek yang paling universal dan mudah dikenali dalam manifestasi kecemasan.

Variasi Penilaian Antar Ahli

Meskipun seluruh butir dalam kategori sangat valid ($V \geq 0,875$), terdapat variasi penilaian pada beberapa butir. Butir 2 dan butir 14 mendapat skor 3 dari salah satu ahli (penilai II), sehingga nilai V-nya (0,875) lebih rendah dibanding butir lain yang mendapat skor minimal 4 dari semua ahli. Hal ini wajar dalam validitas isi dan mencerminkan perbedaan perspektif ahli.

Perbandingan dengan Penelitian Sebelumnya

Dibandingkan dengan penelitian (Desstya et al., 2025) yang memvalidasi instrumen diagnostik dengan Aiken's V dan memperoleh rentang V = 0,79–0,92 (rata-rata sekitar 0,86),

hasil penelitian ini ($V = 0,94$) tergolong lebih tinggi. Hal ini kemungkinan disebabkan oleh beberapa faktor: (1) jumlah panel ahli dalam penelitian ini (4 orang) memenuhi rekomendasi umum minimal 3–5 ahli untuk validitas isi; (2) kisi-kisi indikator disusun secara rinci sebelum penulisan butir; serta (3) panel ahli yang dipilih memiliki latar belakang multidisiplin sehingga penilaian lebih komprehensif.

Penelitian (Widowati & Agustini, 2025) melaporkan bahwa validitas isi menggunakan Aiken's V dengan 3 validator menghasilkan koefisien antara 0,86–0,99 untuk instrumen yang dikembangkan dengan prosedur yang ketat. Hasil penelitian ini konsisten dengan temuan-temuan tersebut, sekaligus memperkuat bukti bahwa skala kecemasan akademik yang dikembangkan memiliki kualitas psikometri yang baik pada tahap validitas konten.

Implikasi Teoretis dan Praktis

Secara teoretis, penelitian ini mengonfirmasi bahwa kecemasan dapat diukur melalui empat dimensi utama, yaitu: dimensi kognitif (pikiran negatif dan kekhawatiran berlebihan, misalnya "Saya selalu membayangkan skenario terburuk"), dimensi afektif (perasaan tegang, gugup, dan takut, misalnya "Saya sering merasakan ketegangan yang tidak nyaman"), dimensi fisiologis (respons tubuh seperti jantung berdebar dan gemetar, misalnya "Saat saya cemas, jantung saya berdebar sangat kencang"), serta dimensi perilaku (tindakan menghindar dan menarik diri, misalnya "Saya cenderung menjauhi situasi yang bisa memicu kecemasan").

Keempat dimensi ini terwakili dalam 30 butir skala yang telah divalidasi, dengan rincian: dimensi kognitif (8 butir), afektif (8 butir), fisiologis (8 butir), dan perilaku (6 butir). Setiap dimensi terdiri atas item *favorable* dan *unfavorable* untuk meminimalkan bias respons.

KESIMPULAN

Berdasarkan hasil analisis validitas isi menggunakan koefisien Aiken's V pada 30 butir skala kecemasan yang dinilai oleh 4 orang panel ahli, diperoleh kesimpulan bahwa seluruh butir memiliki nilai V di atas 0,80 dengan rentang nilai 0,813 hingga 1,00 serta nilai rata-rata V sebesar 0,94 yang termasuk dalam kategori sangat valid. Skala kecemasan ini mengukur empat dimensi utama, yaitu dimensi kognitif (pikiran), dimensi afektif (perasaan), dimensi fisiologis (tubuh), dan dimensi perilaku (tindakan). Seluruh butir dalam skala terbagi menjadi dua jenis pernyataan, yaitu pernyataan *favorable* yang mendukung konstruk kecemasan dan pernyataan *unfavorable* yang berlawanan arah dengan konstruk kecemasan. Dengan demikian, skala kecemasan yang dikembangkan memiliki validitas isi yang sangat tinggi berdasarkan penilaian para ahli.

Pernyataan Apresiasi

Penulis menyampaikan apresiasi yang mendalam kepada para panel ahli yang telah bersedia meluangkan waktu dan memberikan penilaian profesionalnya terhadap 30 butir skala kecemasan akademik:

1. Prof. Dr. Sudarto, ME. (Dosen Universitas Jenderal Soedirman)
2. Prof. Dr. Nur'aeni, S.Psi, M.Si. (Dosen Universitas Muhammadiyah Purwokerto)
3. Ayu Kurnia S. S.Psi., M.Psi (Dosen Universitas Muhammadiyah Purwokerto)
4. Drs. Achmad Sudjadi, MSc. PhD. (Dosen Universitas Jenderal Soedirman)

Para ahli di atas telah memberikan masukan yang sangat berharga, baik melalui penilaian kuantitatif maupun saran kualitatif, yang secara signifikan meningkatkan kualitas butir-butir skala kecemasan akademik dalam penelitian ini.

Penulis menyampaikan penghargaan dan terima kasih yang sebesar-besarnya kepada keluarga tercinta yang telah memberikan dukungan moral, motivasi, semangat, serta doa yang tulus selama proses penelitian dan penyusunan artikel ini dari awal hingga akhir. Segala bentuk kontribusi yang telah diberikan merupakan bagian yang tidak terpisahkan dari keberhasilan

penyelesaian penelitian ini.

Penulis menyadari bahwa artikel ini masih memiliki berbagai keterbatasan. Namun demikian, penulis berharap penelitian ini dapat memberikan manfaat bagi pengembangan instrumen psikologi, khususnya dalam pengukuran kecemasan pada populasi umum di Indonesia.

DAFTAR PUSTAKA

- Aiken, L. R. (1985). Three Coefficients for Analyzing the Reliability and Validity of Ratings. *Educational and Psychological Measurement*, 45(1), 131–142. <https://journals.sagepub.com/doi/10.1177/0013164485451012>
- Ali, A. M., Alkhamees, A. A., Hori, H., Kim, Y., & Kunugi, H. (2021). *The Depression Anxiety Stress Scale 21 : Development and Validation of the Depression Anxiety Stress Scale 8-Item in Psychiatric Patients and the General Public for Easier Mental Health Measurement in a Post COVID-19 World*.
- Boateng, G. O., Neilands, T. B., & Frongillo, E. A. (2018). *Best Practices for Developing and Validating Scales for Health , Social , and Behavioral Research : A Primer*. 6(June), 1–18. <https://doi.org/10.3389/fpubh.2018.00149>
- Desstya, A., Sayekti, I. C., & Abduh, M. (2025). *Development of a four-tier diagnostic test for misconceptions in natural science of primary school pupils*. 22(2), 338–353.
- Gupta, S., Grover, S., Menon, V., Pv, I., Vidhukumar, K., & Chacko, D. (2025). *Item generation and establishing face and content validity of a rating scale : A primer*. <https://doi.org/10.4103/indianjpsychiatry>
- Hang, I., Suen, Y. N., Lai, C., Hui, M., Ming, S., Wong, Y., Kit, S., Chan, W., Ho, E., Lee, M., Tak, M., Wong, H., Yu, E., & Chen, H. (2022). Assessing anxiety among adolescents in Hong Kong : psychometric properties and validity of the Generalised Anxiety Disorder - 7 (GAD - 7) in an epidemiological community sample. *BMC Psychiatry*, 1–11. <https://doi.org/10.1186/s12888-022-04329-9>
- Helmiyatun, Taqjudin, M., & Ahmad. (2025). *Content Validity of Student Perception Instrument on School Educational Management*. 5(1), 93–103. <https://doi.org/10.58737/jpled.v5i1.422>
- Jeldres, M. R., & Costa, E. D. (2023). *A review of Lawshe ' s method for calculating content validity in the social sciences*. November, 1–8. <https://doi.org/10.3389/feduc.2023.1271335>
- Kurniawati, W., Chien, W. T., Lantu, N., Santi, N. F., Astuti, Y. L., & Faradina, F. (2025). *Validation of the pregnancy-related anxiety scale (PrAS) for pregnant women in Indonesia*. 7.
- Laranjeira, C., Querido, A., & Sousa, P. (2023). *Assessment and Psychometric Properties of the 21-Item Depression Anxiety Stress Scale (DASS-21) among Portuguese Higher Education Students during the COVID-19 Pandemic*. 2546–2560.
- Nguyen, B. X., Dang, A. H., Tran, H. T., & Nguyen, B. N. (2025). *Content validity of a toolkit for measuring teachers ' mental health literacy in Vietnam*. <https://doi.org/10.4103/jehp.jehp>
- Sabaruddin, Nur, M., & Fadli, M. (2022). *Content Validity of KIP-K Scholarship Instruments at Higher Education Using Aiken ' s Coefficient Validity Scale*. 06(03), 934–947.
- Saryono, Nurkamto, J., & Hidayatullah, M. F. (2025). *Aiken Index Analysis to Determine the Content Validity of an Instrument Measuring Junior High School Students ' Preparedness for Earthquake Disasters in Physical Education*. 152–162. <https://doi.org/10.18502/kss.v10i9.18479>
- Widowati, A. N., & Agustini, R. (2025). *Development and Validity Analysis of a Phenomenon-*

- Based Advance Organizer Chemistry Learning Tool for Training Science Literacy in High School Students. 11(10), 664–676. <https://doi.org/10.29303/jppipa.v11i10.12598>*
- Wu, J., Long, D., Hafez, N., Maloney, J., Lim, Y., & Samji, H. (2023). *Development and validation of a youth climate anxiety scale for the Youth Development Instrument survey. July, 1473–1483. <https://doi.org/10.1111/inm.13201>*
- Xu, C., Huang, F., Wang, S., Tan, Z., Zhang, H., Liang, Y., Lu, Z., & Shang, L. (2025). Annals of Medicine Development and validation of the reported outcome scale for patients with lumbar disc herniation. *Annals of Medicine, 57(1). <https://doi.org/10.1080/07853890.2025.2524088>*
- Zhang, C., Wang, T., Zeng, P., Zhao, M., & Zhang, G. (2021). *Reliability , Validity , and Measurement Invariance of the General Anxiety Disorder Scale Among Chinese Medical University Students. 12(May), 1–11. <https://doi.org/10.3389/fpsy.2021.648755>*