**VALIDITAS DAN RELIABILITAS INSTRUMEN ASESMEN KEMAMPUAN BERPIKIR INTUITIF MATEMATIS SISWA**

**Tayudi1 \***

1,Universitas Pasundan

\*cs1.jagocell@gmail.com

**Abstrak**

Tujuan dari penelitian ini adalah untuk menguji validitas dan reliabilitas instrumen asesmen kemampuan berpikir intuitif matematis siswa. Penelitian ini merupakan salah satu bagian dari tahapan Model penelitian dan pengembangan Four-D yaitu tahapan design. Instrumen yang sudah dirancang dikonsultasikan kepada ahli untuk nantinya dibuktikan validitas dan reliabilitas. Uji ahli atau Validasi, dilakukan dengan responden para ahli perancangan instrumen atau produk. Selanjutnya analisis validitas konten dilakukan dengan menganalisis hasil validasi konten oleh para ahli dengan menggunakan pendekatan rasio validitas isi (Content Validity Ratio/CVR). Studi reliabilitas yang melibatkan rater biasanya dinamakan dengan kesepakatan antar rater (inter rater agreement) atau reliabilitas antar rater (inter-rater reliability). Hasil penelitian menunjukan bahwa validasi umum dari ke-empat penilai terhadap instrumen asesmen kemampuan berpikir intuitif matematis yaitu dapat digunakan dilihat dari nilai expert judgment, tidakada yang mendapatkan skor maksimal 20, skor terbesar ada pada aspek kesesuaian bahasa dengan persentase 90%. Hasil dalam validasi isi untuk perangkat asesmen unjuk kerja dianalisis dengan menggunakan Validitas isi Lawshe dimana standar kevalidan CVR tergantung pada jumlah SME. Nilai CVR harus memenuhi 0,99 agar butir dapat dinyatakan valid. Hasil dari perhitungan ICC menggunakan SPSS v.20 di dapat. Hasil analisis menunjukkan rata-rata kesepakatan antar rater sebesar 0,200 sedangkan untuk orang rater konsistensinya adalah 0.500 yang itu artinya mempunyai stabilitas yang moderate.

**Kata kunci:** Asesmen kemampuan berpikir intuitif, Instrumen, validitas, reliabilitas

**Abstract**

The purpose of this study was to test the validity and reliability of the instrument for assessing students' mathematical intuition. This research is part of one of the stages of the four-dimensional research and development model, namely the design stage. The instruments that have been designed are consulted with experts to later prove their validity and reliability. expert test or validation, carried out with respondents who are experts in instrument or product design. Furthermore, content validity analysis is carried out by analyzing the results of content validation by experts using the content validity ratio (CVR) approach. Reliability studies involving raters are usually called "inter-rater agreements" or "inter-rater reliability." The results showed that the general validation of the four assessors of the instrument for assessing mathematical intuitive thinking skills was that it could be used from the point of view of expert judgment. No one got a maximum score of 20, and the biggest score was in the aspect of language suitability with a percentage of 90%. The results of content validation for the performance assessment tool were analyzed using Lawshe's content validity, where the validity standard of the CVR depends on the number of SMEs. The CVR value must meet 0.99 so that the items can be declared valid. The results of the ICC calculations using SPSS version 20 are obtained. The results of the analysis show that the average agreement among raters is 0.200, while their consistency is 0.500, which means that they have moderate stability..

**Keywords:**Assessment of intuitive thinking skills, instruments, validity, reliability

**Pendahuluan**

Pembelajaran matematika memerlukan kreativitas, karena kreativitas ini menjadi salah satu yang tidak dapat dipisahkan dari perkembangan ilmu pengetahuan dan teknologi. Kreativitas menawarkan ide-ide baru yang imajinatif, termasuk perbaikan baru untuk masalah, solusi yang berbeda, dan formulasi masalah yang berbeda. Singh dalam (Sabbagh, 2016) menyatakan bahwa kreativitas sebagai proses pembangkit hipotesis berfokus pada kausalitas dalam situasi matematika, menguji dan menguji ulang hipotesis, dan memodifikasi penilaian berdasarkan ide yang berbeda. Matematika, memprediksi apa yang hilang, menjelaskan masalah umum dengan submasalah tertentu, dan akhirnya menyampaikan hasil. Kreativitas adalah kemampuan untuk melihat hubungan antara ide-ide yang tidak berhubungan sehingga menghasilkan ide baru. Kreativitas berperan dalam menciptakan sesuatu yang unik dan berharga (Metcalfe, 2011). Berpikir kreatif mencakup aspek kemampuan kognitif dan metakognitif (Basadur et al., 2000). Berpikir kreatif adalah pembelajaran yang mencakup emosi, intuisi, imajinasi, dan berpikir divergen (aktivitas otak kanan), serta pembelajaran analitis yang mencakup logika (aktivitas otak kiri). Berpikir kreatif adalah kemampuan untuk berpikir secara logis dan intuitif, meliputi otak kiri (logical center) dan otak kanan (*intuition, imajinasi, divergent thinking, center of creative*) (Burnett, 2007).

Munir (2012) mengatakan bahwa menjadi intuitif dapat digunakan sebagai jembatan untuk memahami, sehingga membantu dan memfasilitasi asosiasi objek yang disajikan dengan solusi alternatif yang diinginkan. Solusi yang disukai (munir, 2012). Intuisi diperlukan dalam proses pemecahan masalah matematika untuk menggabungkan keterampilan/pengalaman yang telah dimiliki sebelumnya dalam memecahkan masalah matematika. Tentu saja dalam tahap penyelesaian, munculnya ide-ide tersebut muncul dengan segera, otomatis (segera) atau tiba-tiba (*suddenly*). Ini adalah karakteristik berpikir dengan intuisi (Dreyfus & Eisenberg, 2020).

Penggunaan intuisi dalam matematika bisa saja dapat mengakibatkan kesalahan dalam pemahaman konsep, namun demikian terdapat hasil penelitian yang mendukung pentingnya intuisi dalam pembelajaran matematika dan potensinya dalam meningkatkan pemahaman terhadap matematika serta dalam pemecahan masalah matematika. Pentingnya intuisi dalam pembelajaran matematika seperti yang diungkapkan oleh Raman dalam Erdyana Dwi Etika (2016). Raman menemukan bahwa representasi dan interpretasi intuitif dapat memandu seseorang membuat klaim matematika yang benar. Selain itu ditemukan bahwa bila seseorang menemukan ide kunci (key idea) dalam suatu pembuktian, maka ia dapat mengaitkan pemahaman terhadap klaim dalam pembuktian dengan bukti formal pernyataan matematika (Etika et al., 2016).

Penyelesaian matematika merupakan kemampuan dasar yang harus dikuasai oleh siswa. Berpikir intuitif merupakan peran penting dalam menyelesaikan masalah matematika. Dalam menyelesaikan masalah Matematika diperlukan proses berpikir analitik dan logika. Brunner dan Hart mengungkapkan bahwa dalam menyelesaikan masalah matematika ada dua pendekatan yaitu secara analitik dan intuitif. Berpikir intuitif berperan dalam proses membangun pengetahuan matematika yang tanpa disadari menghasilakan satu keyakinan yang tepat (Muniri, 2018).

Instrumen adalah alat yang dapat digunakan untuk mengumpulkan data. Dengan menggunakan data yang terkumpul, peneliti mampu menjelaskan permasalahan yang dihadapi. Data yang dikumpulkan juga memungkinkan peneliti untuk menguji hipotesis yang diajukan dalam penelitian. Oleh karena itu, keabsahan atau keabsahan data yang diperoleh sangat ditentukan oleh kualitas peralatan yang digunakan, selain prosedur pengumpulan data yang digunakan, oleh karena itu peralatan sangat menentukan kualitas suatu penelitian. memainkan peran. Hal ini karena alat membantu mengubah fakta menjadi data, jadi jika alat yang Anda gunakan berkualitas wajar dari segi validitas dan reliabilitas, data yang Anda dapatkan akan sesuai dengan fakta atau situasi nyata di lapangan (Yudha, 2020).

Asesmen adalah komponen penting dalam proses pembelajaran dan memiliki peran dalam mengetahui hasil (Clements & Cord, 2013; Ulumudin & Fujianita, 2019). Asesmen adalah serangkaian kegiatan untuk memperoleh informasi kualitatif dan kuantitatif awal, berkelanjutan, atau akhir pembelajaran, yang bertujuan untuk mengevaluasi dan mendiagnosis hal-hal yang perlu diperbaiki sehingga pendidik dan siswa dapat meninjau, merencanakan, dan menerapkan tindakan yang harus dilakukan selanjutnya untuk mencapai tujuan pembelajaran (Purnomo, 2014). Asesmen tidak hanya bertujuan untuk pemberian skor, tetapi juga *feedback* baik kepada peserta didik maupun pendidik dalam melakukan perbaikan proses belajar dan mengajar. Dengan demikian, asesmen selalu menjadi bagian integral dan tidak dapat dipisahkan dalam pembelajaran serta menjadi bagian krusial untuk membantu peserta didik dan pendidik dalam proses belajar-mengajar.

Asesmen hasil belajar pada kelompok mata pelajaran matematika dilakukan melalui pengamatan terhadap kemampuan pengetahuan dan keterampilan untuk menilai perkembangan psikomotor dan kognitif siswa (BSNP, 2010). Asesmen yang dilakukan guru mencakup semua hasil belajar siswa yaitu kemampuan kognitif atau berpikir, kemampuan psikomotor atau kemampuan praktek, dan kemampuan afektif, namun penekanan pada masing ranah tidak sama sehingga harus diperhatikan karakteristik mata pelajaran yang akan diukur (Mardapi, 2012).

Penilaian capaian hasil belajar pada tingkat kognitif yang lebih tinggi (*higher-order thinking*), menurut Nitko dan Brookhart (Nitko & Brookhart, 2011) dibutuhkan tes (*task*) yang menuntut peserta didik untuk menggunakan pengetahuan dan keterampilan dalam situasi yang baru (*new or novel situations*). Dengan demikian peserta didik tidak hanya dituntut untuk memahami, tetapi sampai mampu untuk menganalisis, mengevaluasi dan berkreasi.

Beberapa permasalahan ditemukan di lapangan, nampak ada ketidaksesuaian antara pembelajaran matematika di SMP dengan sistem penilaian yang digunakannya. Proses penilaian yang biasa dilakukan guru selama ini hanya mampu menggambarkan aspek penguasaan konsep peserta didik, akibatnya tujuan kurikuler mata pelajaran Matematika belum dapat dicapai dan atau tergambarkan secara autentik dan menyeluruh. Selain itu, dikemukakan masih terdapat miskonsepsi tentang pada responden yang diteliti, kurangnya pemahaman sebagian besar responden tentang tata cara membuat perangkat asesmen uraian berupa rubrik yang sesuai, dan sebagian responden yang cukup memahami asesmen dan guru juga masih kesulitan karena keterbatasan waktu dan banyaknya jumlah siswa dalam satu kelas. Juga menghitung validitas pada intrumen perlu diketahui bahwa sistem penilaian yang baik sangat berpengaruh terhadap peningkatan motivasi siswa dalam pembelajaran. Namun pada umumnya guru-guru belum memberikan perhatian yang serius dalam penyusunan perangkat penilaian, sehingga selama ini guru hanya melakukan penilaian dengan cara subjektif tanpa memikirkan kriteria apa yang seharusnya untuk dinilai.

Berdasarkan pengalaman guru-guru matematika, salah satu pokok bahasan yang sulit dipahami oleh siswa adalah pokok bahasan bangun ruang sisi datar. Bangun ruang sisi datar merupakan salah satu pokok bahasan pada standar kompetensi kelas VIII SMP semester 2 yang juga dianggap sulit oleh siswa, tidak hanya bagi siswa, tetapi sebagian guru juga menganggap geometri khususnya materi bangun ruang sisi datar merupakan materi yang sulit sehingga mereka memerlukan pelatihan tambahan. Hal ini dipertegas pada Konferensi Nasional Penelitian Matematika dan Pembelajarannya (KNPMP) tahun 2019 menyatakan dalam pembelajaran Bangun Ruang Sisi Datar terdapat beberapa kesulitan yang dialami oleh siswa yakni antara lain: menentukan bentuk jaring-jaring kubus, balok, menentukan rumus untuk mencari volume dan luas permukaan bidang bangun ruang sisi datar, menyelesaikan soal cerita yang berbasis permasalahan sehari-hari. Menurut Rohmah (2014) Salah satu penyebab siswa kurang menguasai materi bangun ruang sisi datar adalah sumber belajar yang dipakai siswa dimana hanya menekankan konsep yang mengacu pada hafalan, penggunaan rumus tanpa mengetahui asal rumus tersebut seperti pembelajaran konvensional, sehingga berakibat siswa cenderung mengesampingkan konsep dasar dan lebih mengutamakan pada hasil belajar dengan menggunakan rumus yang telah dihafalkan. Padahal konsep dasar pada materi bangun ruang sisi datar harus dikuasai oleh siswa karena dalam permendikbud No.68 tahun 2013 tentang kurikulum SMP-MTs pada bagian Matematika kelas VIII hal 45 kompetensi dasar poin 3.11 tertulis Menaksir dan menghitung luas permukaan dan volume bangun ruang yang tidak beraturan dengan menerapkan geometri dasarnya.

Asesmen kemampuan berpikir intuitif matematis menghadirkan sejumlah masalah validitas yang tidak mudah ditangani dengan pendekatan dan kriteria tradisional untuk penelitian validitas. Asesmen ini biasanya memberi siswa kebebasan yang substansial dalam menafsirkan, merespons, dan mungkin merancang tugas; mereka menghasilkan lebih sedikit respons independen, yang masing-masing kompleks, yang mencerminkan integrasi berbagai keterampilan dan pengetahuan; dan mereka membutuhkan penilaian ahli untuk evaluasi. Konsekuensinya, memenuhi kriteria yang terkait dengan masalah validitas seperti reliabilitas, generalisasi, dan komparabilitas penilaian, paling tidak seperti yang didefinisikan dan dioperasionalkan menjadi bermasalah. Permasalahan yang sering dihadapi guru matematika SMP selanjutnya adalah dalam melakukan asesmen kemampuan berpikir intuitif matematis terletak pada validitas dan reliabilitas alat ukur yang digunakan dalam hal ini adalah rubrik.

Instrumen penilaian yang dapat mendeskripsikan keterampilan siswa perlu dilakukan validasi sebagai salah satu syarat evaluasi. Instrumen penilaian harus konsisten dalam penggunaannya dan dapat mengukur sesuai sasaran ukuran, sehingga perlu dilakukan validasi. Validitas merupakan suatu proses yang dilakukan oleh pengembang untuk mengumpulkan data guna mengumpulkan kesimpulan berdasarkan skor instrumen yang diperoleh. Instrumen penilaian memiliki validitas tinggi jika alat tersebut menjalankan fungsi ukurnya atau memberikan hasil ukur yang sesuai sehingga dapat dilakukan pengukuran menggunakan instrumen penilaian tersebut. Secara garis besar, validitas dapat dibedakan menjadi dua macam, yaitu validitas logis dan validitas empiris (Farida: 2017). Validitas logis pada instrumen penilaian menunjukkan kondisi instrumen penilaian yang telah memenuhi persyaratan valid berdasarkan hasil penalaran karena telah dirancang secara baik serta mengikuti teori dan ketentuan. Validitas empiris instrumen penilaian diujikan berdasarkan pengalaman yang dapat dilakukan melalui uji coba kepada siswa (Yusup, 2018).

Sebagai jawaban atas kenyataan permasalahan diatas, maka dilakukan analisis untuk melihat instrumen asesmen kemampuan berpikir intuitif matematis siswa dalam topik pengukuran geometri bangun ruang sisi datar. Kelayakan dikembangkan instrument memenuhi tiga aspek validitas, reliabilitas, dan kepraktisan. Hal ini juga dijelaskan oleh Nieveen dan Folmer (2013) bahwa empat kriteria kualitas untuk mendapatkan solusi terbaik untuk menyelesaikan masalah yang kompleks dipendidikan: relevansi, dikenal sebagai validitas konten; konsistensi, juga disebut validitas konstruk; kepraktisan, menunjukkan bahwa dapat digunakan berdasarkan pengaturan; dan keefektifan, menyerahkan produk sesuai dengan hasil yang diinginkan (Nieveen & Folmer, 2013). Rasionalisasi pemilihan pengembangan instrumen asesmen kemampuan berpikir intuitif matematis siswa didasarkan pada hasil wawancara, analisis kepustakaan terhadap hasil penelitian terdahulu, serta analisis kebutuhan yang dilakukan peneliti terhadap guru yang telah dipaparkan di atas.

**Metode**

Penelitian ini merupakan jenis penelitian dan pengembangan (*Research and Development*). Adapun model penelitian dan pengembangan pendidikan (Research and Development) yang dilakukan mengadopsi pada model pengembangan 4-D (*Four-D*). Model penelitian dan pengembangan *Four-D* dikembangkan oleh Sivasailam Thiagarajan, Dorothy S. Semmel dan Melvyn I. Semmel (1974) dengan tahapan penelitian yaitu, *define, design, develop, and dissemination* (Thiagarajan, 1974).

Penelitian ini merupakan salah satu bagian dari tahapan Model penelitian dan pengembangan *Four-D* yaitu tahapan *design*. Instrumen yang sudah dirancang dikonsultasikan kepada ahli untuk nantinya dibuktikan validitas dan reliabilitas. Uji ahli atau Validasi, dilakukan dengan responden para ahli perancangan instrumen atau produk. Kegiatan ini dilakukan untuk *review* produk awal memberikan masukan untuk perbaikan instrumen selanjutnya.Proses validasi ini disebut dengan *Expert Judgement*. Instrumen yang sudah dihasilkan dievaluasi secara kuantitatif, apakah format yang dihasilkan sudah layak atau belum, dan bagaimana kesesuaian isi materi penilaian pembelajaran. Pengategorian persentase validitas logis mengacu pada (Yusup, 2018) dengan ketentuan hasil validitas ≤ 20% berkategori sangat rendah; 21%-40% berkategori rendah; 41%-60% berkategori sedang; 61%-80% berkategori tinggi; dan ≥81% berkategori sangat tinggi. Jika instrumen belum layak, maka instrumen direvisi kembali sehingga instrumen menjadi layak untuk diuji cobakan.

Ada empat aspek terpisah dalam validasi intrumen kemampuan berpikir intuitif matematis. Aspek pertama adalah kesesuaian aspek dengan indikator, aspek penilisan, aspek bahasa, dan aspek petunjuk pelaksanaan. Diharapkan bahwa validator akan bervariasi dalam kemampuan mereka. Tujuan dari pemeriksaan ini adalah untuk membedakan menurut pakar secara andal apakah intrumen sudah layak atau belum untuk digunakan dilapangan.

Selanjutnya analisis validitas konten dilakukan dengan menganalisis hasil validasi konten oleh para ahli dengan menggunakan pendekatan rasio validitas isi (*Content Validity Ratio*/CVR). Berdasarkan persamaan Lawshe, dapat diperoleh nilai CVR untuk setiap butir soal. Makna rumusan dari Lawshe (1975) ini adalah: 1) jika validator yang menyatakan setuju kurang dari setengah dari jumlah total validator maka nilai CVR bernilai negatif; 2) jika validator yang menyatakan setuju tepat setengah dari jumlah total validator maka nilai CVR bernilai nol; dan 3) jika validator yang menyatakan setuju lebih dari setengah jumlah total validator maka nilai CVR berada antara 0 sampai dengan 1. Nilai CVR yang diperoleh dari perhitungan dibandingkan dengan nilai kritis CVR berdasarkan jumlah validator seperti yang tercantum pada Tabel 3.4. Butir soal diterima jika butir soal memiliki nilai sama dengan atau lebih besar dari nilai kritis CVR dan butir soal ditolak apabila memiliki nilai lebih rendah dari nilai kritis CVR (Lawshe, 1975; Wilson, 2004). Setelah mengidentifikasi pertanyaan pada lembar validasi dengan menggunakan CVR, kemudian dihitunglah CVI (*Content Validity Index*). Secara sederhana CVI merupakan rata-rata dari nilai CVR untuk item tes yang dijawab.

Studi reliabilitas yang melibatkan rater biasanya dinamakan dengan kesepakatan antar rater (*inter rater agreement*) atau reliabilitas antar rater (*inter-rater reliability*). Jika pada kasus *self-report* reliabilitas ditunjukkan dengan konsistensi internal yang terlihat dari antara satu butir dan butir lainnya memiliki korelasi yang tinggi, maka dalam kasus reliabilitas antar rater yang diuji konsistensinya adalah raternya. Jadi posisi butir digantikan dengan posisi orang (rater). Rater-rater yang memiliki kesepakatan tinggi terlihat dari posisi subjek yang diobservasi. Jika urutan skor subjek dari Rater A dan B hampir sama maka kedua rater memiliki kesepakatan yang tinggi (Wess et al., 2021). Hal ini dikarenakan kesepatakan dioperasionalkan dalam bentuk korelasi.

Reliabilitas antar rater terdiri dari dua jenis, uji koefisien korelasi Kesepakatan Antar Rater dari Kappa dan uji koefisien korelasi antar kelas *(Intraclass Correlation Coefficients).* Uji reliabilitas antar rater Kappa digunakan apabila rater berjumlah 2 orang sedangkan uji reliabilitas antar rater ICC digunakan apabila rater lebih dari 2 orang. Penelitian ini menggunakan 4 *expert/rater* sehingga menggunakan koefisien korelasi antar kelas. ICC menunjukkan perbandingan antara variasi yang diakibatkan atribut yang diukur dengan variasi pengukuran secara keseluruhan. Alat ukur memiliki stabilitas memadai jika ICC antar pengukuran > 0.50, stabilitas tinggi jika ICC antar pengukuran ≥ 0.80. < .20 very low, .20-.39 low, .40-.59 moderate, .60-.79 high,.80-1.00 very high (Streiner & Norman, 2008)..

**Hasil Penelitian dan Pembahasan**

Instrumen asesmen kemampuan berpikir intuitif matematis yang berupa rubrik holistik terdiri dari 1 rubrik total sepuluh soal, yang dijabarkan dari masing-masing aspek. Untuk setiap butir instrumen tes ditentukan skala pengukurannya secara kualitatif, melalui sistem ini kualitas kinerja dapat diskor secara gradual mulai skor 1 jika hanya mampu mencapai satu kriteria, dan skor 4 jika mampu mencapai semua kriteria skoring. Langkah-langkah yang ditempuh pada pengembangan instrumen asesmen kemampuan berpikir intuitif matematis diformat dalam bentuk tabel, yang unsur utamanya terdiri dari: kolom pertama berisi aspek-aspek/indikator yang dinilai, kolom kedua berisi kriteria skoring/deskriptor, dan kolom ketiga berisi skor perolehan untuk setiap indikator. Berdasarkan hasil análisis data uji validitas isi diperoleh informasi bahwa semua butir adalah relevan. Setelah para pakar diberikan *questionnaire* I atas variabel penelitian ini, maka dapat ditabulasikan hasilnya pada tabel 1.

Hasil-hasil penilaian ke-empat penilai terhadap instrumen asesmen kemampuan berpikir intuitif matematis terangkum dalam Tabel 1. Aspek-aspek penilaian meliputi: kesesuaian aspek penilaian kemampuan berpikir intuitif matematis dengan indikator yang ada, aspek kesesuaian aspek dengan indikator, penulisan, aspek bahasa, dan aspek petunjuk penggunaan.

Pada Tabel 1, tampak bahwa hasil validasi umum dari ke-empat penilai terhadap instrumen asesmen kemampuan berpikir intuitif matematis yaitu dapat digunakan dilihat dari nilai *expert judgment*, tidak ada yang mendapatkan skor maksimal 20, skor terbesar ada pada aspek kesesuaian bahasa dengan persentase 90%

**Tabel 1.** Rangkuman Hasil Validasi Instrumen Asesmen Kemampuan Berpikir Intuitif Matematis

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **No.** | **Aspek Yang dinilai** | **Kriteria Penilaian** | **Validator Ahli** | | | | **Skor** | **Rata-rata** | **Kualitas** | **Presentase** |
| **1** | **2** | **3** | **4** |
| 1 | Kesesuaian Aspek dengan indikator | 1 | 4 | 4 | 4 | 4 | 16 | 0.775 | SB | 78% |
| 2 | 3 | 4 | 4 | 4 | 15 |
| 2 | Penulisan | 3 | 4 | 4 | 4 | 4 | 16 | 0.866 | SB | 87% |
| 4 | 5 | 4 | 4 | 5 | 18 |
| 5 | 4 | 5 | 4 | 5 | 18 |
| 3 | Bahasa | 6 | 5 | 5 | 5 | 4 | 19 | 0.9 | SB | 90% |
| 7 | 5 | 5 | 4 | 5 | 19 |
| 8 | 5 | 4 | 3 | 4 | 16 |
| 4 | Petunjuk pelaksanaan | 9 | 4 | 5 | 4 | 5 | 18 | 0.875 | SB | 87.50% |
| 10 | 4 | 4 | 4 | 5 | 17 |
| Jumlah Skor | | | 43 | 44 | 40 | 45 | 172 | 0.86 | SB | 86% |
|

Berikutnya menghitung validitas isi suatu butir dapat dibuktikan dengan menggunakan CVR dan CVI. CVR dan CVI diusulkan oleh Lawshe pada tahun 1975 dengan menggunakan 3 skala rating (LAWSHE, 1975). Pada tahun 1985, Aiken mengusulkan konsep validitas isi dengan lebih rinci. Kerincian tersebut terlihat dari standar kevalidan yang dipengaruhi oleh jumlah rater dan skala rating yang digunakan (R lewis, 1985).

Nilai CVR memiliki rentang antara -1 sampai dengan 1. Jika setengah dari SME menyatakan esensial maka nilai CVR akan bernilai 0. CVR akan bernilai 1 jika seluruh SME menyatakan esensial untuk suatu butir. Koefisien validitas Aiken didapatkan perhitungan dengan skor mentah dari ahli yang berjumlah n. Nilai koefisien V Aiken memiliki rentang antara -1 sampai dengan 1 (Fitriana & Supahar, 2019).

**Tabel 2.** Skor Rata-rata V Aiken untuk Instrumen Asesmen Kemampuan Berpikir Intuitif Matematis

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **No.** | **Aspek Yang dinilai** | **Kriteria Penilaian** | **Validator Ahli** | | | | **Rata-Rata CVR** |
| **1** | **2** | **3** | **4** |
| 1 | Kesesuaian Aspek dengan indikator | 1 | 4 | 4 | 4 | 4 | 1 |
| 2 | 3 | 4 | 4 | 4 | 1 |
| 2 | Penulisan | 3 | 4 | 4 | 4 | 4 | 1 |
| 4 | 5 | 4 | 4 | 5 | 1 |
| 5 | 4 | 5 | 4 | 5 | 1 |
| 3 | Bahasa | 6 | 5 | 5 | 5 | 4 | 1 |
| 7 | 5 | 5 | 4 | 5 | 1 |
| 8 | 5 | 4 | 3 | 4 | 1 |
| 4 | Petunjuk pelaksanaan | 9 | 4 | 5 | 4 | 5 | 1 |
| 10 | 4 | 4 | 4 | 5 | 1 |
| CVI | | |  |  |  |  | 1 |

Hasil dalam validasi isi untuk perangkat asesmen unjuk kerja dianalisis dengan menggunakan Validitas isi Lawshe dimana standar kevalidan CVR tergantung pada jumlah SME. Nilai CVR harus memenuhi 0,99 agar butir dapat dinyatakan valid. Hal tersebut berlaku untuk validasi isi dengan menggunakan 3 SME (Lawshe, 1975, p. 568). Nilai CVR yang diperoleh dari tiap butir adalah 1 dan secara lengkap disajikan dalam lampiran. Nilai CVI yang diperoleh dari rata-rata CVR sebesar 1. Berdasarkan nilai CVR yang melebihi 0,99 maka seluruh butir dinyatakan valid (R lewis, 1985) dan layak digunakan untuk penelitian lebih lanjut.

Tingkat kesepakatan (reliabilitas) antar ke-empat penilai tersebut dapat dijelaskan dengan menghitung koefisien reliabilitas Antar penilai menggunakan koefisien *Intraclass Correlation Coefficient*. Hasil penghitungannya (dengan menggunakan bantunan Programm SPSS versi 16) disajikan pada Tabel 2

Dari tabel 3 dapat dilihat bahwa hasil dari perhitungan ICC menggunakan SPSS v.20 di dapat. Hasil analisis menunjukkan rata-rata kesepakatan antar *rater* sebesar 0,200 sedangkan untuk orang *rater* konsistensinya adalah 0.500 yang itu artinya mempunyai stabilitas yang *moderate*. (Fleiss, 1975; Streiner & Norman, 2008; Williams, 2008)

**Tabel** 3 Uji Realibilitas *Intraclass Correlation Coefficient* Validasi Ahli

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Intraclass Correlation Coefficient** | | | | | | | |
|  | Intraclass Correlationb | 95% Confidence Interval | | F Test with True Value 0 | | | |
| Lower Bound | Upper Bound | Value | df1 | df2 | Sig |
| Single Measures | .200a | -.064 | .607 | 2.000 | 9 | 27 | .079 |
| Average Measures | .500c | -.315 | .861 | 2.000 | 9 | 27 | .079 |
| Two-way mixed effects model where people effects are random and measures effects are fixed. | | | | | | | |
| a. The estimator is the same, whether the interaction effect is present or not. | | | | | | | |
| b. Type C intraclass correlation coefficients using a consistency definition. The between-measure variance is excluded from the denominator variance. | | | | | | | |
| c. This estimate is computed assuming the interaction effect is absent, because it is not estimable otherwise. | | | | | | | |

Hasil validasi keseluruhan diperoleh instrument asesmen kemampuan berpikir intuitif matematis yang valid dan reliabel untuk dapat mendeskripsikan berpikir intuitif matematis melalui penalaran formal dalam memecahkan permasalahan geometri sebanyak 10 butir intrumen penilaian. Instrumen penilaian yang telah valid dapat mendeskripsikan kemampuan berpikir intuitif matematis dengan mempresentasikan aspek yang dapat dikembangkan untuk mengasah kemampuan pemecahan masalah dalam kehidupan sehari-hari melalui penalaran formal siswa. Permasalahan pada instrumen asesmen dapat diselesaikan dengan memecahkan masalah yang termuat di dalam pertanyaan di setiap butir instrumen penilaian, sebab dibutuhkan kemampuan bernalar sehingga mampu berpikir intuitif matematis dengan mengumpulkan informasi yang relevan dalam menganalisis, menemukan hubungan antar fakta/informasi yang diberikan, mengidentifikasi dan merencanakan strategi penyelesaian soal untuk mendapatkan jawaban yang tepat serta menyadari perlunya meneliti kembali hasil yang telah diperoleh (Pratiwi et al., 2015).

Peluang pengembangan intrumen asesmen kemampuan berpikir intuitif matematis terbuka luas karena belum adanya kesepakatan pasti mengenai asesmen kemampuan berpikir intuitif matematis yang bertujuan untuk melihat keberhasilan dalam pemberdayaan keterampilan berpikir intuitif matematis (Adamura & Susanti, 2018). Instrumen asesmen ini dapat mendeskripsikan kemampuan berpikir intuitif matematis dalam setiap tahapannya yang didasari dari kemampuan bernalar secara formal siswa dalam memahami permasalahan dan mengaitkan dengan konsep yang dimiliki, sintesis, dan evaluasi melalui keterampilan berpikir kritis yang dimiliki sehingga mampu memecahkan permasalahan pada materi geometri bangu ruang yang terus diajarkan dan dikembangkan secara rutin berdasarkan keterkaitannya dengan materi lainnya.

**Simpulan**

Instrumen asesmen kemampuan berpikir intuitif matematis pada pembelajaran matematika disekolah menengah pertama adalah valid menggunakan bukti validitas isi oleh 4 ahli. Hasil penilaian yang diperoleh dari validasi ahli menyatakan bahwa penilaian unjuk kerja ini layak dan termasuk dalam kategori sangat baik digunakan sebagai bentuk penilaian. Dengan demikian instrumen asesmen kemampuan berpikir intuitif matematis yang telah dikembangkan telah valid dan reliabel secara konstruk melalui validasi ahli sebanyak 10 butir instrument penilaian. Instrumen asesmen yang telah validdan reliabel tersebut dapat mendeskripsikan kemampuan berpikir intuitif matematis dengan mempresentasikan aspek yang dapat dikembangkan untuk mengasah kemampuan pemecahan masalah dalam kehidupan sehari-hari melalui kemampuan berpikir intuitif matematis siswa. Instrumen asesmen ini dapat diimplementasikan guna menguji efektivitasnya melalui penelitian selanjutnya.

**Referensi**

Adamura, F., & Susanti, V. D. (2018). Penalaran Matematis Mahasiswa dengan Kemampuan Berpikir Intuitif Sedang dalam Memecahkan Masalah Analisis Real. *Jurnal Edukasi Matematika Dan Sains*. https://doi.org/10.25273/jems.v6i2.5366

Basadur, M., Runco, M. A., & Vegaxy, L. A. (2000). Understanding how creative thinking skills, attitudes and behaviors work together: A causal process model. *Journal of Creative Behavior*. https://doi.org/10.1002/j.2162-6057.2000.tb01203.x

BSNP. (2010). Paradigma Pendidikan Nasional Abad XXI. In *Paradigma Pendidikan Nasional Abad XXI*.

Burnett, C. (2007). *Deliberate intuition : Giving intuitive insights their rightful place in the creative problem solving thinking skills model*. *3*, 236–253. http://www.bridgepointeffect.com/wp-content/uploads/2012/12/Deliberate-Intuition-CIM.pdf

Clements, M. D., & Cord, B. A. (2013). Assessment Guiding Learning: Developing Graduate Qualities in an Experiential Learning Programme. *Assessment and Evaluation in Higher Education*, *38*(1), 114–124. https://doi.org/10.1080/02602938.2011.609314

Dreyfus, T., & Eisenberg, T. (2020). Intuitive Functional Concepts: A Baseline Study on Intuitions. *Journal for Research in Mathematics Education*. https://doi.org/10.5951/jresematheduc.13.5.0360

Etika, E. D., Sujadi, I., & Subanti, S. (2016). Intuisi Kelas VII SMP Negeri 1 Nganjuk dalam Pemecahan Masalah Matematika Ditinjau Dari Adversity Quotient ( Aq ). *Jurnal Elektronik Pembelajaran Matematika*, *4*(5), 563–574.

Fitriana, D. A., & Supahar, S. (2019). Developing an Assessment Instrument of Mathematical Problem-Solving Skills in Senior High School. *International Journal of Trends in Mathematics Education Research*. https://doi.org/10.33122/ijtmer.v2i3.81

Fleiss, J. L. (1975). Measuring Agreement between Two Judges on the Presence or Absence of a Trait. *Biometrics*. https://doi.org/10.2307/2529549

LAWSHE, C. H. (1975). A QUANTITATIVE APPROACH TO CONTENT VALIDITY. *Personnel Psychology*. https://doi.org/10.1111/j.1744-6570.1975.tb01393.x

Mardapi, D. (2012). Pengukuran Penilaian dan Evaluasi Pendidikan. *Yogyakarta: Nuha Medika*, *45*.

METCALFE, A. W. S. (2011). Cognitive Psychology: A Student’s Handbook (6th Ed.) by M. W. Eysenck and M. T. Keane. *British Journal of Psychology*. https://doi.org/10.1111/j.2044-8295.2010.02010.x

munir. (2012). *Model Penalaran Intuitif Siswa*. *November*, 978–979.

Muniri, M. (2018). Peran Berpikir Intuitif dan Analitis dalam Memecahkan Masalah Matematika. *Jurnal Tadris Matematika*. https://doi.org/10.21274/jtm.2018.1.1.9-22

Nieveen, N., & Folmer, E. (2013). Formative evaluation in educational design research. *Educational Design Research. Part A: An Introduction*.

Nitko, A. J., & Brookhart, S. M. (2011). Educational assessment of students. *Human Movement Science*. https://doi.org/10.1016/j.humov.2005.02.001

Pratiwi, N. D., Setyarsih, W., Fisika, J., Matematika, F., Ilmu, D., Alam, P., Surabaya, U. N., & Kunci, K. (2015). Pengembangan Instrumen Evaluasi Berbasis Taksonomi Structure of the Observed Learning Outcome (SOLO) Untuk Menentukan Profil Kemampuan Siswa dalam Memecahkan Masalah Fluida Statis. *Jurnal Inovasi Pendidikan Fisika (JIPF) Nurul Dwi Pratiwi*.

Purnomo, Y. W. (2014). Assessment-Based Learning : Sebuah Tinjauan untuk Meningkatkan Motivasi Assessment-Based Learning : Sebuah Tinjauan untuk Meningkatkan Motivasi Belajar dan Pemahaman Matematis. *Sigma Journal*, *VI*(June).

R, A. L. (1985). Three Coefficients for Analysing Reliability and Validity of Rating. *Educational and Psychological Measurement*, *45*, 131–142. https://doi.org/10.1177/07399863870092005

Sabbagh, S. A. (2016). Childhood Students’ Creativity in Mathematics Class in Jordan. *American Journal of Educational Research*.

Streiner, D. L., & Norman, G. R. (2008). Health Measurement Scales: A practical guide to their development and use. In *Health Measurement Scales: A Practical Guide to their Development and Use*. https://doi.org/10.1093/acprof:oso/9780199231881.001.0001

Thiagarajan, S. (1974). Instructional development for training teachers of exceptional children. In *A sourcebook*.

Ulumudin, I., & Fujianita, S. (2019). The Implementation of Attitude Assessment in Curriculum 2013 at Elementary Schools. *Jurnal Penelitian Dan Evaluasi Pendidikan*, *23*(1), 46–56. https://doi.org/10.21831/pep.v23i1.23391

Wess, R., Klock, H., Siller, H. S., & Greefrath, G. (2021). Test Quality. In *International Perspectives on the Teaching and Learning of Mathematical Modelling*. https://doi.org/10.1007/978-3-030-78071-5\_4

Williams, G. (2008). Introduction to research in the health sciences (5th edition). *British Dental Journal*. https://doi.org/10.1038/sj.bdj.2008.313

Wilson, M. (2004). Constructing measures: An item response modeling approach. In *Constructing Measures: An Item Response Modeling Approach*. https://doi.org/10.4324/9781410611697

Yudha, R. P. (2020). *Asesmen Unjuk Kerja Geometri,*. CV Budi Utama.

Yusup, F. (2018). Uji Validitas dan Reliabilitas Instrumen Penelitian Kuantitatif. *Jurnal Tarbiyah : Jurnal Ilmiah Kependidikan*. https://doi.org/10.18592/tarbiyah.v7i1.2100