



Pengembangan Instrumen Penilaian Berbasis Scaffolding untuk Mengukur Scientific Reasoning Pada Pembelajaran IPA di SMP

Yussy Miranda¹, Zainur Rasyid Ridlo², Ulin Nuha³

^{1,2,3}Universitas Jember, Indonesia

E-mail: zainur.fkip@unej.ac.id

Article Info	Abstract
Article History Received: 2023-02-12 Revised: 2023-03-23 Published: 2024-04-03 Keywords: <i>Scientific Reasoning;</i> <i>Scaffolding;</i> <i>Science Learning.</i>	The aim of this research is to determine the provision of scaffolding-based assessment instruments to measure scientific reasoning in science learning in junior high schools on temperature, heat and expansion which was carried out in November at one of the MTsN 1 Jember Schools, Jember Regency. This research uses Borg and Gall. Samples were taken using small group tests and large group tests in class VII E, F, G, H. This research instrument is in the form of scientific reasoning test questions developed by Lawson, namely the LCTSR (Lawson Classroom Test of Scientific Reasoning) type. The research results show that the validity of the assessment instrument has an average of 80.59%. The reliability of the assessment instrument obtained an index of 0.837 and can be categorized as very high reliability. Discriminating power of question items There are 3 categories, namely, 5 items in the good category, 10 items in the sufficient category, and 5 items in the bad category which is determined by the discriminating power of the 20 question items tested. The difficulty level of the 20 question items tested resulted in a percentage of 20% easy, 70% medium and 10% difficult. However, these results still need to be improved so that students can be trained to have Scientific Reasoning skills.

Artikel Info	Abstrak
Sejarah Artikel Diterima: 2023-02-12 Direvisi: 2023-03-23 Dipublikasi: 2024-04-03 Kata kunci: <i>Scientific Reasoning;</i> <i>Scaffolding;</i> <i>Pembelajaran IPA.</i>	Tujuan dari penelitian ini yaitu untuk mengetahui pemberian instrumen penilaian berbasis scaffolding untuk mengukur scientific reasoning pada pembelajaran IPA di SMP pada materi suhu, kalor dan pemuai yang dilaksanakan pada bulan november di salah satu Sekolah MTsN 1 Jember Kabupaten Jember. Penelitian ini menggunakan Borg and Gall. Sampel diambil menggunakan uji kelompok kecil dan uji kelompok besar di kelas VII E,F,G,H. Instrumen penelitian ini berupa soal tes penalaran ilmiah yang dikembangkan oleh Lawson yaitu tipe LCTSR (Lawson Classroom Test of Scientific Reasoning). Dari hasil penelitian menunjukkan bahwa validitas instrumen penilaian mendapatkan rata-rata sebesar 80,59%. Reliabilitas instrumen penilaian mendapatkan indeks sebesar 0,837 dan dapat dikategorikan reliabel sangat tinggi. Daya pembeda butir Soal Terdapat 3 kategori yaitu, 5 item dalam kategori baik, 10 item dalam kategori cukup, dan 5 item dalam kategori buruk yang ditentukan oleh daya pembeda dari 20 item pertanyaan yang diujikan. Tingkat kesukaran 20 item soal yang diujikan menghasilkan presentase mudah 20%, sedang 70% dan sukar 10%. Namun hasil ini masih harus ditingkatkan kembali agar peserta didik dapat terlatih untuk memiliki kemampuan Scientific Reasoning.

I. PENDAHULUAN

Pendidikan abad 21 sebaiknya membekali siswa berbagai kemampuan dan kompetensi untuk menghadapi tantangan global yang terus berubah, membutuhkan keragaman keterampilan, informasi, dan metode pembelajaran baru, oleh karena itu kemampuan *scientific reasoning* diperlukan (Utami et al., 2019). *Scientific reasoning* ialah kemampuan seseorang untuk mentuntaskan masalah dengan berpikir sistematis dan logis menggunakan metode secara ilmiah dengan cara mengevaluasi fakta, kemudian memprediksi dan membuat hipotesis, memberi penentuan serta pengontrolan variabel, melaksanakan perancangan yang selanjutnya

melakukan percobaan, menghimpun data dan dianalisis serta terakhir menarik simpulan (Aulia Handayani et al., 2020).

Scientific reasoning ini penting dilakukan dalam proses pembelajaran karena merupakan dasar dari perkembangan pemecahan masalah peserta didik dalam belajar. Pembelajaran IPA membutuhkan *scientific reasoning* karena mewajibkan peserta didik untuk berpikir secara luas. Walaupun *scientific reasoning* termasuk penting dalam aspek pembelajaran siswa, namun masih banyak siswa yang *scientific reasoning* nya rendah karena siswa menganggap bahwa sesuatu yang tidak biasa dilakukan dan juga dianggap

sulit dalam pembelajarannya sehari-hari (Novianti, et al., 2019).

Scientific reasoning mempunyai peran yang amat penting dalam penyelesaian suatu masalah dan tidak terpisahkan dalam pembelajaran sehingga dapat mentuntaskan permasalahan soal yang diberikan (Konita, et al., 2019). Akan tetapi fakta dilapangan memperlihatkan bahwa tak sedikit peserta didik yang tak mendayagunakan penalarannya ketika mentuntaskan permasalahan dalam belajar sehingga kemampuan bernalarnya masih rendah. Pernyataan rendahnya *scientific reasoning* dapat dibuktikan dengan penelitian yang telah dilaksanakan Pohan (2022) yang menyatakan bahwa kemampuan penalaran di MTsN 4 Padang dikelas VIII A masih rendah. Selain itu juga sependapat dengan penelitian yang telah dilakukan oleh Yusa (2022) yaitu kemampuan penalaran VIII SMP Negeri 3 Jombang juga masih dalam kategori kurang. Pernyataan tersebut juga dibenarkan oleh *Programme for International Tests Student Assessment (PISA)* adalah kemampuan bernalar secara ilmiah dengan hasil tes PISA 2009 Indonesia menduduki peringkat ke-60 dari 65 negara dengan skor rata-rata 383 dibandingkan dengan skor rata-rata 501 OECD (Purwana & Rusdiana, 2021). Rendahnya *scientific reasoning* juga didukung oleh wawancara yang saya lakukan di MTsN 1 Jember bahwa kemampuan bernalar siswa masih rendah jika dihadapkan dengan soal fisika.

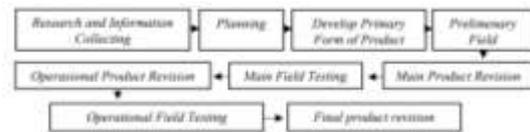
Scaffolding yang pada hakekatnya merupakan pendekatan pembelajaran untuk membantu siswa dalam proses pembelajaran kognitif. Sesuai dengan tujuannya, *scaffolding* akan menggugah siswa untuk berpikir lebih dengan memberikan bimbingan berupa instruksi dan tugas yang jelas. (Pratama & Saregar, 2019). Dalam pengembangan ini, *scaffolding* diberikan pada saat proses soal latihan siswa dengan tujuan bahwa *scaffolding* tertulis lebih membantu dalam membantu siswa memahami daripada menghafal (Affriyenni et al., 2021).

Berdasarkan penjelasan di atas, maka peneliti menawarkan inovasi baru yaitu instrumen berbasis *scaffolding* (Destiana et al., 2020). *Scientific reasoning* dihubungkan dengan berbasis *scaffolding* agar para siswa dapat bernalar dalam proses belajarnya serta dapat meningkatkan pemahaman mereka tentang soal yang diberikan (Hidayah et al., 2019). Penggunaan instrumen penilaian berbasis *scaffolding* dalam pembelajaran berpotensi untuk membuat peserta didik pemahaman tentang *scientific reasoning* dapat meningkat pada

pembelajaran IPA SMP, sehingga pernyataan tersebut perlu diujikan dalam susunan penelitian.

II. METODE PENELITIAN

Penelitian yang dilaksanakan ialah penelitian pengembangan. Produk dari penelitian pengembangan ini berupa instrumen penilaian berbasis *scaffolding* untuk mengukur *scientific reasoning*. Desain penelitian yang digunakan adalah model pengembangan *Borg and Gall*. Penggunaan model *Borg and Gall* dalam mengembangkan instrumen penilaian dinilai cocok karena langkahnya yang sistematis dan mudah untuk dipahami. Model pengembangan *Borg and Gall* yang digunakan dalam proses penelitian ini dan metodologinya sesuai dengan alur atau tahapannya. Model pengembangan *Borg and Gall* meliputi 10 langkah umum. Langkah-langkah tersebut tidak harus diikuti karena dapat menyesuaikan dengan kebutuhan peneliti, sehingga penelitian ini hanya menggunakan 8 tahapan. Tahapan penelitian dapat dilihat pada gambar dibawah ini.



Gambar 1. Tahapan *Borg and Gall*

Tahapan penelitian ini meliputi (1) Pencarian dan pengumpulan data (*Research and Information Collecting*) Pemilihan materi yang dipilih dalam pengembangan instrumen penilaian ini adalah suhu, kalor dan pemuai. (2) Perencanaan instrumen penilaian (*Planning*) Menentukan skala yang berisikan komponen angka dan penskoran. (3) Pengembangan draft produk (*Develop Preliminary Form of Product*) Membuat spesifikasi instrumen penilaian kinerja, kisi-kisi instrumen, format dan menyusun instrumen, skala penilaian, melakukan finishing produk. (4) Validasi logis (*Preliminary Field Testing*) Validator memberikan penilaian kinerja kesesuaian teori dan ketentuan dalam menyusun instrumen, format penilaian, serta keseluruhan instrumen. (5) Revisi produk utama (*Main Product Revision*) Instrumen penilaian yang divalidasi sebelum produk digunakan penelitian dilakukan uji dengan sampel yang lebih terbatas. (6) Uji coba produk (*Main Field Testing*) Produk yang telah melalui validasi kemudian digunakan dalam penelitian kecil untuk mengetahui apakah instrumen dapat digunakan sebanyak 20 soal. (7) Revisi produk operasional (*Operational Product*

Revision) Apabila instrumen dinilai kurang mampu menakar maka perlu dilakukan revisi sehingga hasil produk operasional dapat digunakan dalam penelitian. (8) Uji coba lapangan operasional (*Operational Field Testing*) Uji validitas melibatkan 3 validator yaitu 2 dosen dan 1 guru untuk memberikan penilaian yang terdapat pada lembar validasi.

III. HASIL DAN PEMBAHASAN

Peneliti mengembangkan instrumen penilaian berbasis scaffolding untuk mengukur *scientific reasoning* di kelas VII E, F, G, H MTsN 1 Jember. Penelitian dilakukan karena berdasarkan hasil observasi dan wawancara bersama guru IPA yang menyatakan bahwa kemampuan *scientific reasoning* masih tergolong rendah. Penelitian ini merupakan penelitian pengembangan produk yang mengukur validitas butir soal, reliabilitas, daya pembeda, dan kesukaran. Produk yang dihasilkan berupa instrumen penilaian berbasis scaffolding, terdiri atas 20 soal uraian disusun mengandung ranah untuk mengukur *scientific reasoning* pada siswa. Pengembangan instrumen penilaian berbasis scaffolding ini menggunakan model pengembangan *Borg dan Gall* dengan 8 tahapan. Pada bagian ini tidak lagi memuat rancangan penelitian, namun terfokus pada hasil penelitian yang telah dilaksanakan.

1. Pencarian dan pengumpulan data (*Research and Information Collecting*) Pengembangan ini didasarkan atas hasil dari pencarian informasi dan pengumpulan data yang dilakukan oleh peneliti di MTsN 1 Jember. Pencarian informasi dan mengumpulkan data dilakukan dengan wawancara tentang permasalahan yang terjadi di sekolah. Berdasarkan hasil wawancara yang dilakukan pada guru MTsN 1 Jember, mendapatkan informasi bahwa peserta didik belum pernah sepenuhnya diberikan soal yang berbasis scaffolding, terutama pada mata pelajaran IPA materi suhu, kalor dan pemuain. Kemampuan *scientific reasoning* peserta didik masih rendah, sehingga evaluasi pembelajaran yang digunakan masih soal tingkat rendah. Oleh karena itu, dengan adanya scaffolding untuk mengukur *scientific reasoning* sangat penting untuk dilatih.
2. Perencanaan instrumen penilaian (*Planning*) Tahap perencanaan dilaksanakan dengan menentukan jumlah butir tes yang akan diuji coba kepada peserta didik. Hasil analisis materi yang diperoleh terdiri dari beberapa capaian pembelajaran, yaitu pada akhir fase D, peserta didik dapat menganalisis skala

termometer celcius ke dalam skala termometer yang sejenis atau lainnya; peserta didik dapat menganalisis perpindahan kalor secara konduksi, konveksi dan radiasi; peserta didik dapat mengkategorikan proses pemuain zat padat, zat cair dan gas; peserta didik dapat menganalisis pengaruh kalor terhadap perubahan suhu dan wujud zat; peserta didik dapat menghitung perbandingan antar suhu zat; peserta didik dapat menghitung besarnya kalor jenis pada suatu benda; peserta didik dapat menafsirkan dalam bentuk gambar perpindahan kalor secara konduksi, konveksi dan radiasi.

3. Pengembangan draft produk (*Develop Preliminary Form of Product*) Produk di desain dengan menggunakan canva karena dapat membuat soal terlihat menarik dengan desain yang sesuai dengan peserta didik SMP, dengan desain yang menarik jadi peserta didik tertarik untuk mengerjakan soal yang diberikan. Produk diawali dengan membuat kisi kisi soal yang berbentuk tabel dengan capaian pembelajaran, level kognitif, indikator *scientific reasoning*, indikator soal, nomor soal, soal, skor, dan kunci jawaban.
4. Validasi logis (*Preliminary Field Testing*) Kevalidan dari produk yang dikembangkan dapat diketahui setelah dilakukannya validasi desain produk. Tahap validasi desain produk pengembangan instrumen penilaian berbasis scaffolding untuk mengukur *scientific reasoning* peserta didik dinilai oleh 3 validator menggunakan lembar validasi dan hasil data dari analisis validasi ahli oleh ke-3 validator yaitu dapat dilihat pada tabel berikut dibawah ini

Tabel 1. Hasil Validitas Ahli

Validator	Aspek	Presentase Validator (%)	Presentase (%)	Kriteria
Validator I	Materi	80.25	80.45	Valid
	Konstruksi	80.50		
	Bahasa	80.60		
Validator II	Materi	81.50	82.58	Sangat Valid
	Konstruksi	84.25		
	Bahasa	82.00		
Validator III	Materi	78.75	78.75	Valid
	Konstruksi	77.50		
	Bahasa	80.00		
Rata - Rata Skor			80.59	Valid

Berdasarkan tabel di atas dapat diketahui bahwa hasil validitas dari ketiga ahli untuk setiap aspek instrument penilaian memiliki nilai presentase berada pada interval nilai 78,75% - 82,58%. Dari Validator I diperoleh nilai presentase skor sebesar 80,45% dengan kriteria "Valid", Validator II diperoleh nilai

presentase skor sebesar 82,58% dengan kriteria "Sangat Valid" dan dari Validator III diperoleh nilai presentase skor sebesar 78,75% dengan kriteria "Valid". Selain itu, dari tabel di atas juga diperoleh nilai rata - rata skor sebesar 80,59%. Nilai tersebut menunjukkan bahwa secara keseluruhan instrumen penilaian yang dilihat dari ketiga aspek, yaitu Materi, Konstruksi dan Bahasa dapat dikatakan sudah "Valid" dan dapat digunakan.

5. Revisi produk utama (*Main Product Revision*) Tahap revisi dilakukan dengan melakukan perbaikan pada instrumen tes yang telah divalidasi oleh validator. Bagian yang direvisi berupa *Scaffolding* gambar, petunjuk, penguraian langkah langkah, peringatan kurang jelas. Level kognitif tidak ada kemampuan mengkreasi C6 dalam soal. Oleh karena itu, perlu adaya revisi dengan menyajikan gambar dan sumbernya dengan jelas, penguraian langkah langkah dan peringatan juga diperjelas. Dan juga pada menambahkan level kognitif C6 kemampuan mengkreasi pada soal, sehingga soal dapat diterima oleh peserta didik dan layak untuk digunakan dalam penelitian.
6. Uji coba produk awal (*Main Field Testing*) Sesudah dilakukannya kegiatan revisi validasi ahli, tahapan berikutnya yaitu melaksanakan pengujian produk pada kelompok kecil untuk mengevaluasi kelemahan produk yang telah dikembangkan agar selanjutnya dapat diujikan kembali pada kelompok yang lebih besar. Tahap uji coba produk dilaksanakan sebanyak 32 peserta didik di kelas VII F MTsN 1 Jember. Selanjutnya dilakukan analisis perhitungan validitas butir soal, reliabilitas, daya pembeda, tingkat kesukaran.

Uji validitas dilakukan untuk mengetahui kevalidan butir soal yang akan digunakan dalam mengukur kemampuan *scientiffiec reasoning*. Uji validitas yang digunakan pada penelitian ini adalah uji *product moment*. Apabila nilai korelasi r hitung $\geq r$ tabel (30, 5%) atau p - value $< 0,05$ maka butir soal dikatakan sudah valid dan begitupula sebaliknya, untuk soal yang tidak valid bisa direvisi atau juga bisa dibuang diganti soal lainnya. Hasil uji validitas butir soal disajikan oleh tabel berikut dibawah ini

Tabel 2. Hasil Validitas Uji Kelompok Kecil

No.	Butir Soal	p - value	r hitung	r tabel	Keterangan
1	Soal 1	0.027	0.3892	0.3494	Valid
2	Soal 2	<0.001	0.6616		Valid
3	Soal 3	<0.001	0.7867		Valid
4	Soal 4	0.068	0.2961		Tidak Valid
5	Soal 5	0.029	0.3864		Valid
6	Soal 6	0.090	0.2757		Tidak Valid
7	Soal 7	<0.001	0.5551		Valid
8	Soal 8	<0.001	0.7131		Valid
9	Soal 9	0.013	0.4323		Valid
10	Soal 10	<0.001	0.5768		Valid
11	Soal 11	0.002	0.5322		Valid
12	Soal 12	<0.001	0.5705		Valid
13	Soal 13	0.019	0.4100		Valid
14	Soal 14	<0.001	0.7880		Valid
15	Soal 15	0.009	0.4635		Valid
16	Soal 16	<0.001	0.7192		Valid
17	Soal 17	<0.001	0.8120		Valid
18	Soal 18	0.027	0.3946		Valid
19	Soal 19	0.918	0.3682		Valid
20	Soal 20	0.112	0.3767		Valid

Berdasarkan hasil uji validitas di atas dapat diketahui bahwa terdapat butir soal yang mempunyai nilai r hitung $< r$ tabel (0,3494) dan p - value $> 0,05$ yaitu soal no. 4 dan soal no. 6. Hal ini berarti bahwa kedua butir soal tersebut belum valid.

Uji reliabilitas dilakukan untuk mengukur sejauh mana konsistensi dan kestabilan hasil pengukuran dapat dipercaya. Uji reliabilitas yang digunakan pada penelitian ini adalah uji Cronbach's Alpha. Apabila nilai Cronbach's Alpha $> 0,60$ maka dikatakan instrumen penelitian telah reliabel dan begitupula sebaliknya, Oleh karena itu, hasil uji reliabilitas disajikan oleh tabel berikut dibawah ini:

Tabel 3. Hasil Uji Reliabilitas Uji Kelompok Kecil

Butir Soal	Cronbach's Alpha	Batasan Nilai	Keterangan
20	0.860	> 0.60	Reliabel

Berdasarkan hasil uji reliabilitas pada tabel di atas diketahui bahwa nilai Cronbach's Alpha sebesar 0,860 lebih besar dari 0,60. Dengan demikian, dapat disimpulkan bahwa semua butir soal telah reliabel dan layak digunakan untuk penelitian. Daya pembeda butir soal di analisis menggunakan nilai indeks daya pembeda. Butir soal yang dibuat pastinya harus dapat membedakan siswa yang pandai dengan siswa yang kurang pandai. Pertanyaan yang memiliki daya pembeda yang baik dapat dilihat dari hasil indeks daya pembedanya.

Tabel 4. Hasil Distribusi Frekuensi Daya Pembeda Uji Kelompok Kecil

Klasifikasi	No. Soal	Jumlah	Presentase (%)
Buruk	4,5,6,9,13,19,20	7	35.00
Cukup	1,2,7,10,11,12,15,18	8	40.00
Baik	3,8,14,16,17	5	25.00

Berdasarkan hasil daya beda yang dilakukan terhadap 20 butir soal uraian dapat diketahui bahwa sebanyak 7 (35.00%) butir soal termasuk dalam kategori Buruk yaitu soal nomor 4,5,6,9,13,19,20, yang mana soal – soal tersebut memiliki nilai indeks daya beda 0.00 – 0.20. Adapun sebanyak 8 (40.00%) butir soal termasuk dalam kategori Cukup yaitu soal nomor 1,2,7,10,11,12,15,18 dengan indeks daya beda 0.21 – 0.40, dan sisanya sebanyak 5 (25.00%) butir soal nomor 3,8,14,16,17 termasuk ke dalam kategori Baik dengan nilai indeks berada pada interval 0.41 – 0.70. Dari tabel di atas juga dapat diketahui bahwa rata – rata soal uraian siswa kelas VII MTs Negeri 1 Jember untuk data kelompok kecil sebesar 0.27 dengan kategori Cukup. Analisis tingkat kesukaran dilakukan dengan maksud untuk mendapatkan setiap butir soal tergolong mudah, sedang, atau sukar. Tingkat kesukaran yang terdapat setiap butir soal dapat dimanfaatkan untuk mengetahui kemampuan siswa dalam mengerjakan soal. Sebaiknya dalam pembuatan soal harus ada kriteria soal yang mudah, sedang, dan sukar.

Persebaran tingkat kesukaran soal lebih mengarah pada proporsi perbandingan 3-5-2 atau 30% dengan kategori mudah, 50% dengan kategori sedang, dan 20% dengan kategori sukar. Oleh karena itu, hasil uji kesukaran dapat dilihat pada tabel berikut:

Tabel 5. Distribusi Frekuensi Kesukaran Soal Uji Kelompok Kecil

Klasifikasi	No. Soal	Jumlah	Presentase (%)
Mudah	1, 5, 9, 13, 18	5	25.00
Sedang	2, 3, 4, 6, 7, 8, 10, 12, 14, 17	13	65.00
Sukar	19, 20	2	10.00

Berdasarkan hasil analisis yang dilakukan terhadap 20 butir soal uraian dapat diketahui bahwa sebanyak 5 (25.00%) butir soal termasuk ke dalam kategori mudah yang mana soal-soal tersebut memiliki nilai indeks 0.71 – 1.00, sebanyak 13 (65.00%) butir soal termasuk ke dalam kategori sedang dengan nilai indeks 0.31 – 0.70. Adapun sisanya sebanyak 2 (10.00%) butir soal termasuk ke

dalam kategori sukar dengan nilai indeks berada pada interval 0.00 – 0.30.

- Revisi produk operasional (*Operational Product Revision*) Tahapan revisi produk awal merupakan tahap dimana butir soal yang belum memenuhi kriteria valid pada uji validitas butir soal, akan dilakukan pembuangan soal dengan tujuan untuk mengganti soal yang tidak valid sehingga dapat diujikan pada tahap selanjutnya. Berdasarkan hasil uji coba produk awal terdapat 2 soal yang tidak valid yaitu nomor soal 4 dan 6 sehingga perlu untuk menggantikannya dengan soal yang lain.
- Uji coba lapangan operasional (*Operational Field Testing*) Uji coba produk akhir dilaksanakan setelah melakukan perbaikan pada instrumen penilaian yang dikembangkan. Uji coba produk akhir dilaksanakan menggunakan sampel yang lebih besar, yaitu melibatkan 128 peserta didik dari kelas VII (E, F, G, dan H). Hasil uji coba produk akhir dianalisis menggunakan Microsoft excel dan Spss untuk mengetahui validitas, reliabilitas, daya pembeda dan kesukaran instrumen penilaian berbasis *scaffolding*.

Tabel 6. Hasil Uji Validitas Butir Soal Kelompok Besar

No.	Butir Soal	p-value	r hitung	r tabel	Keterangan
1	Soal 1	0.002	0.270	0.146	Valid
2	Soal 2	<0.001	0.698		Valid
3	Soal 3	<0.001	0.647		Valid
4	Soal 4	<0.001	0.548		Valid
5	Soal 5	<0.001	0.348		Valid
6	Soal 6	<0.001	0.489		Valid
7	Soal 7	<0.001	0.477		Valid
8	Soal 8	<0.001	0.542		Valid
9	Soal 9	<0.001	0.177		Valid
10	Soal 10	<0.001	0.429		Valid
11	Soal 11	0.046	0.604		Valid
12	Soal 12	<0.001	0.505		Valid
13	Soal 13	<0.001	0.546		Valid
14	Soal 14	<0.001	0.532		Valid
15	Soal 15	<0.001	0.607		Valid
16	Soal 16	<0.001	0.628		Valid
17	Soal 17	<0.001	0.560		Valid
18	Soal 18	<0.001	0.298		Valid
19	Soal 19	<0.001	0.453		Valid
20	Soal 20	<0.001	0.413		Valid

Berdasarkan hasil uji validitas butir soal untuk data kelompok besar di atas menunjukkan bahwa setiap butir soal mempunyai nilai r hitung \geq r tabel (0,146) dan p – value < 0,05. Hal ini berarti bahwa semua butir soal yaitu sebanyak 20 butir soal dalam penelitian telah valid dan dapat digunakan

sebagai instrument penelitian untuk mengukur kemampuan *scientific reasoning*.

Tabel 7. Hasil Uji Reliabilitas Data Kelompok Besar

Butir Soal	Cronbach's Alpha	Batasan Nilai	Keterangan
20	0.837	> 0.60	Reliabel

Berdasarkan hasil uji reliabilitas pada tabel di atas diketahui bahwa nilai *Cronbach's Alpha* sebesar 0,837 lebih besar dari 0,60. Dengan demikian, dapat disimpulkan bahwa semua butir soal telah reliabel dan layak digunakan untuk penelitian.

Tabel 8. Distribusi Frekuensi Daya Pembeda Soal Kelompok Besar

Kategori	No. Soal	Jumlah	Presentase (%)
Buruk	1, 9, 18, 19, 20	5	25.00
Cukup	5, 6, 7, 10, 11, 12, 13, 14, 16, 17	10	50.00
Baik	2, 3, 4, 8, 15	5	25.00

Berdasarkan hasil daya beda yang dilakukan terhadap 20 butir soal uraian dapat diketahui bahwa sebanyak 5 (25.00%) butir soal termasuk dalam kategori Buruk yaitu soal 1, 9, 18, 19, 20 yang mana soal – soal tersebut memiliki nilai indeks daya beda 0.00 – 0.20. Adapun sebanyak 10 (50.00%) butir soal termasuk dalam kategori Cukup yaitu soal nomor 5, 6, 7, 10, 11, 12, 13, 14, 16, 17 dengan indeks daya beda 0.21 – 0.40, dan sisanya sebanyak 5 (25.00%) butir soal yaitu 2, 3, 4, 8, 15 termasuk ke dalam kategori Baik dengan nilai indeks berada pada interval 0.41 – 0.70. Dari tabel di atas juga dapat diketahui bahwa rata – rata soal uraian siswa kelas VII MTs Negeri 1 Jember untuk data kelompok besar sebesar 0.33 dengan kategori Cukup.

Tabel 9. Distribusi Frekuensi Kesukaran Soal Uji Kelompok Besar

Kategori	No. Soal	Jumlah	Presentase (%)
Mudah	1, 5, 9, 18	4	20.00
Sedang	2, 3, 4, 6, 7, 8, 10, 11, 12, 13, 14, 15, 16, 17	14	70.00
Sukar	19, 20	2	10.00

Berdasarkan hasil analisis yang dilakukan terhadap 20 butir soal uraian dapat diketahui bahwa sebanyak 4 (20.00%) butir soal termasuk ke dalam kategori mudah yang mana soal-soal tersebut memiliki nilai indeks 0.71 – 1.00, sebanyak 14 (70.00%) butir soal termasuk ke dalam kategori sedang dengan nilai indeks 0.31-070. Adapun sisanya

sebanyak 2 (10.00%) butir soal termasuk ke dalam kategori sukar dengan nilai indeks berada pada interval 0.00 – 0.30.

IV. SIMPULAN DAN SARAN

A. Simpulan

Validitas instrumen penilaian yang dibuat mendapatkan rata-rata sebesar 80,59% dalam kategori valid dari tiga validator, sedangkan pada validitas butir soal dinyatakan valid dan layak pada semua nomer soalnya. Berdasarkan hal tersebut, 20 butir soal pada instrumen berbasis scaffolding yang dibuat dinyatakan valid dan layak digunakan untuk mengukur *scientific reasoning* pada pembelajaran IPA di SMP. Reliabilitas instrumen penilaian yang dibuat mendapatkan indeks sebesar 0,837 dan dapat dikategorikan reliabel sangat tinggi. Berdasarkan hal tersebut, 20 butir soal pada instrumen berbasis scaffolding dinyatakan telah konsisten dan terbilang reliabel sangat tinggi untuk dapat digunakan mengukur *scientific reasoning* pada pembelajaran IPA SMP.

Daya Pembeda Butir. Soal Terdapat 3 kategori yaitu, 5 item dalam kategori baik, 10 item dalam kategori cukup, dan 5 item dalam kategori buruk yang ditentukan oleh daya pembeda dari 20 item pertanyaan yang diujikan. Berdasarkan hal tersebut, 15 soal pada instrumen berbasis scaffolding dapat digunakan kembali, sedangkan 5 soal perlu diubah agar dapat dimanfaatkan dalam penilaian pembelajaran selanjutnya. Tingkat kesukaran 20 item soal yang diujikan menghasilkan klasifikasian 4 soal mudah, 14 soal sedang, dan 2 soal sukar. Berdasarkan hal tersebut, tingkat kesukarannya di presentase mudah 20%, sedang 70% dan sukar 10%.

B. Saran

Sebaiknya untuk pengembangan instrumen penilaian berbasis scaffolding ini lebih bervariasi tentang pengembangan scaffolding, sehingga diharapkan para siswa siswi lebih terlatih dan juga dapat mengembangkan kemampuan *scientific reasoning*nya.

DAFTAR RUJUKAN

Affriyenni, Y., Indrianingrum, D., Se-tiawan, A. M., Nida, S., & Ma-lang, U. N. 2021. Pengembangan e-scaffolding berbasis pembelajaran hybrid pada materi gerak lurus kelas VIII SMP / MTs. 5.

- Aulia Handayani, G., Windyariani, S., & Yanuar Pauzi, R. 2020. Profil Tingkat Penalaran Ilmiah Siswa Sekolah Menengah Atas Pada Materi Ekosistem. *Biodik*, 6(2), 176-186.
- Azwar, S. 2012. *Sikap Manusia: Teori dan Pengukurannya*. Yogyakarta: Liberty.
- Balqis, D., Kusairi, S., & Supriana, E. 2019. Analisis Kemampuan Penalaran Ilmiah pada Pembelajaran Interactive Demonstration disertai Formative Assessment. *Jurnal Pendidikan: Teori, Penelitian, Dan Pengembangan*, 4(11), 1485.
- Destiana, D., Suchyadi, Y., & Anjaswuri, F. 2020. Pengembangan instrumen penilaian untuk meningkatkan kualitas pembelajaran produk-tif di sekolah dasar. 03(September), 119-123.
- Fatimah, L, dan K. Alfath. 2019. Analisis kesukaran soal, daya pembeda dan fungsi distraktor. *Jurnal Komunikasi dan Pendidikan Is-lam*. 8(2): 37-64
- Handayani, G.A. Windyarian, S. Pauzi, R.Y. 2020. Profil Tingkat Penalaran Ilmiah Siswa Sekolah Menengah Atas Pada Materi Ekosistem. *Jurnal Ilmiah Pendidikan Biologi*. 6 (2): 176-186.
- Hidayah, I., Adji, T. B., & Setiawan, N. A. 2019. Development and evaluation of adaptive metacognitive scaffolding for algorithm. *IET Software*, 13(4), 305-312.
- Hanifah. 2014. Perbandingan Tingkat Kesukaran, Daya Pembeda Butir Soal dan Reliabilitas Tes Bentuk Soal Pilihan Ganda Biasa dan Pilihan Ganda Asosiasi Mata Pelajaran Ekonomi. *Sosio e-Kons*. 6(1): 41-55
- Muliastrini, N.K.E 2019. Pengaruh Mod-el Pembelajaran Inkuiri dengan Teknik Scaffolding terhadap Ke-mampuan Literasi Sains dan Pres-tasi Belajar IPA. *Jurnal Ilmiah Sekolah Dasar*. 2(3): 230-240.
- Novanti, S. K. E., Yulianti, E., & Mustikasari, V. R. 2019. Pengembangan Instrumen Tes Literasi Sains Siswa Smp Materi Tekanan Zat Dan Penerapannya Dalam Kehidupan Sehari-Hari. *Jurnal Pembelajaran Sains*, 2(2009), 7.
- Pohan, I. D.R.R. & Dewi, I. 2022. Upaya Meningkatkan Kemampuan Penalaran Siswa VIII A MTsN 4 Pa-dang Lawas Menggunakan Kooperatif Tipe (NHT) Berbantu-an Goegebra. *Jurnal Ilmiah multi disiplin indonesia*, 2(1), 121-134.
- Utami, P., Supeno, & Bektiarso, S. 2019. Lembar Kerja Siswa (LKS) Ber-basis Inkuiri dengan Bantuan Scaffolding Konseptual untuk Meningkatkan Keterampilan Penalaran Ilmiah Fisika Siswa SMA. *Seminar Nasional Pendidikan Fisika 2019*, 4(1), 134-140.
- Yusa, I. W., Hadi, W. P., Suwandi. 2022. Analisis Profil Scientific Reasoning Ability dan Korelasi terhadap Hasil Penilaian Akhir Semester Peserta Didik pada Pembelajaran IPA. *Jurnal Pendidikan MIPA*. 12(3): 902-911.