

PENGEMBANGAN *E-INSTRUMEN* TES MATERI STOIKIOMETRI MENGUNAKAN *BLOOKET GAME* UNTUK MENGUKUR KEMAMPUAN BERPIKIR KRITIS PADA SISWA SMA

Development of E- Instruments for Stoichiometric Material Tests Using Blooket Games to Measure Critical Thinking Ability in High School Students

Noni Syahidah*, Muhammad Kusasi, Rizki Nur Analita, Arif Sholahuddin

Program Studi Pendidikan Kimia, FKIP Universitas Lambung Mangkurat
Jl. Brigigen H. Hasan Basry Banjarmasin, 70123, Kalimantan Selatan, Indonesia

*email: noniisyhdh@gmail.com

Informasi Artikel	Abstrak
<p>Kata kunci: <i>E-instrumen tes, blooket game, kemampuan berpikir kritis, stoikiometri, validitas, daya pembeda butir soal</i></p> <p>Keywords: <i>e-test instrument, blooket game, critical thinking ability, stoichiometry, validity, differentiating power of question items.</i></p>	<p>Proses pembelajaran perlu dievaluasi keberhasilannya untuk mengukur sejauh mana proses belajar mengajar yang dilakukan kepada peserta didik melalui alat ukur yaitu instrumen evaluasi. Studi pendahuluan melalui observasi yang dilakukan menunjukkan bahwa sebagian besar guru tidak membuat instrumen secara eksklusif dalam melakukan evaluasi pembelajaran, dan penilaian yang dilakukan masih berbasis kertas. Alat evaluasi pembelajaran dapat dikembangkan dengan memanfaatkan teknologi, salah satunya adalah aplikasi <i>blooket game</i>. Penelitian ini bertujuan untuk mengembangkan instrumen elektronik pada materi stoikiometri menggunakan <i>blooket game</i> yang digunakan untuk mengevaluasi berpikir kritis peserta didik yang valid secara isi dan konstruk, reliabel, memiliki kesukaran dan daya pembeda yang baik, terbaca dan respon yang baik dari peserta didik. Pengembangan e-instrumen menggunakan metode <i>Research and Development (R&D)</i> dengan model ADDIE (<i>Analysis, Design, Development, Implementation, Evaluation</i>). Peserta didik SMA Negeri 4 Banjarmasin sebanyak 114 merupakan subjek dalam penelitian ini dan instrumen elektronik yang dikembangkan merupakan objek penelitian. Data dikumpulkan dengan menggunakan e-instrumen tes dan angket yang kemudian dianalisis dengan analisis deskriptif menggunakan kriteria-kriteria yang telah ditentukan. Hasil penelitian menunjukkan instrumen tes elektronik valid, reliabel, kesukaran sedang, daya pembeda sedang, memiliki tingkat keterbacaan yang baik sekali dan direspon baik oleh peserta didik. Pada uji coba kelompok kecil yang terdiri dari 34 peserta didik diperoleh 15 butir soal yang layak untuk dilanjutkan ke tahapan penelitian berikutnya berdasarkan hasil validitas dan daya pembeda butir soal yang dilakukan dari total 20 butir soal sebelumnya. Pada uji coba lapangan terdiri dari 70 peserta didik, diperoleh sebanyak 15 butir soal yang layak digunakan untuk mengukur kemampuan berpikir kritis.</p>

Abstract. *The lack of The learning process needs to be evaluated for its success to measure the extent of the teaching and learning process carried out by students through measuring tools, namely evaluation instruments. Preliminary studies through observations carried out show that the majority of teachers do not create instruments exclusively for evaluating learning, and the assessments carried out are still paper-based. Learning evaluation tools can be developed by utilizing technology, one of which is the blocket game application. This research aims to develop an electronic instrument on stoichiometric materials using a blocket game which is used to transmit students' critical thinking that is content and construct valid, reliable, has good difficulty and differentiating power, is legible and responds well from students. The development of e-instruments uses the Research and Development (R&D) method with the ADDIE (Analysis, Design, Development, Implementation, Evaluation) model. 114 students of SMA Negeri 4 Banjarmasin were the subjects of this research and the electronic instruments developed were the research objects. Data was collected using e-test instruments and questionnaires which were then analyzed descriptively using predetermined criteria. The research results show that the electronic test instrument is valid, reliable, has moderate difficulty, has moderate distinguishing power, has a very good readability level and is responded well by students. In a small group trial consisting of 34 students, 15 questions were obtained which were suitable to be continued to the next research stage based on the results of the validity and differentiating power of the questions carried out from a total of 20 previous questions. In the field trial consisting of 70 students, 15 questions were obtained that were suitable for measuring critical thinking abilities.*

PENDAHULUAN

Pendidikan saat ini menuntut peserta didik untuk memiliki keterampilan berpikir kritis. Kemampuan berpikir kritis merupakan aktivitas kognitif yang melibatkan analisis, penilaian, dan evaluasi informasi. Ini mencakup keterampilan seperti memperhatikan, mengelompokkan, memilih, dan menilai data. Belajar berpikir kritis membantu memahami masalah dari berbagai sudut pandang dan membuat kesimpulan rasional, penting untuk kehidupan dan karir (Puspita et al., 2018). Kemampuan berpikir kritis ini menjadi kunci dalam pengembangan pemikiran yang rasional dan analitis, sehingga sangat diperlukan dalam pembelajaran sains (Puspita & Parma Dewi, 2021). Pembelajaran kimia merupakan salah satu ilmu sains yang memerlukan kemampuan berpikir kritis karena materi kimia sering kali melibatkan konsep-konsep abstrak dan penerapannya dalam situasi dunia nyata.

Pembelajaran kimia mencakup analisis atom dan molekul, karakteristik bahan, reaksi kimia, serta energi yang terlibat, disertai eksperimen laboratorium. (Redhana, 2019). Kimia sering dianggap sulit karena menggabungkan perhitungan kompleks dan konsep teoritis abstrak, seperti stoikiometri yang membutuhkan pemahaman mendalam tentang atom, molekul, dan reaksi kimia. Kesulitan dalam stoikiometri ini menyebabkan banyak siswa terhalang dalam mencapai pemahaman penuh karena

topik ini melibatkan perhitungan matematis yang kompleks dan pemahaman mendalam tentang konsep kimia seperti mol dan penyetaraan persamaan reaksi. Tantangan ini menunjukkan perlunya pendekatan pengajaran yang lebih efektif untuk membantu peserta didik menguasai stoikiometri (Fitriani et al., 2024; Sutinah et al., 2016).

Karakteristik materi stoikiometri di atas memerlukan kemampuan berpikir kritis yang tinggi (Agustini et al., 2021), sehingga guru dituntut untuk dapat menggali pengetahuan peserta didik untuk bisa berpikir kritis agar lebih mudah menguasai setiap pembelajaran (Gunawan et al., 2023). Kemampuan berpikir kritis pada materi stoikiometri kimia dapat dievaluasi menggunakan instrumen tes kemampuan berpikir kritis. Instrumen tes memainkan peran penting dalam proses evaluasi, karena berorientasi pada penilaian berbagai parameter penting untuk menentukan keandalan, akurasi, dan integritas (Händel et al., 2023). Oleh karena itu, diperlukan instrumen tes/evaluasi yang dapat menggali kemampuan berpikir kritis peserta didik secara komprehensif.

Berdasarkan observasi awal di SMAN 4 Banjarmasin, peserta didik mengalami kesulitan dalam memahami materi stoikiometri hal ini di tinjau dari hasil belajar peserta didik yang cenderung rendah. Hasil belajar rendah tersebut diakibatkan karena kurangnya kemampuan berpikir kritis peserta didik. Hasil observasi juga menunjukkan bahwa belum adanya instrumen evaluasi yang menggali dan mengukur kemampuan berpikir kritis di SMAN 4 Banjarmasin. Proses evaluasi hasil pembelajaran masih menggunakan format cetak konvensional, meskipun evaluasi juga dapat dilakukan dengan menggunakan teknologi baru.

Pemanfaatan teknologi dalam proses evaluasi dapat memanfaatkan penggunaan online tool *blooket game* sebagai wadah alat evaluasi akan membuat alat evaluasi lebih menarik dan efisien serta dapat memudahkan pendidik untuk menilai hasil (Putri, 2021). Media *Blooket game* merupakan salah satu platform yang digunakan dalam pembelajaran. Situs platform ini dapat diakses melalui <https://www.blooket.com>. *Blooket* digunakan sebagai alat evaluasi, seperti kuis, dalam pelaksanaan *pretest* dan *posttest*, namun alat evaluasi atau instrumen tes harus valid dan dapat diandalkan untuk memastikan pengukuran yang akurat dan konsisten dalam penilaian (Octaria et al., 2022).

Berdasarkan uraian di atas peneliti telah melakukan pengembangan instrumen tes berpikir kritis pada materi stoikiometri dengan menggunakan *Blooket game* untuk mengukur kemampuan berpikir kritis peserta didik. *Blooket game* ini memiliki perbedaan dengan platform lain seperti *Kahoot* dan *Quizizz* adalah bahwa *Blooket* menyediakan berbagai permainan yang dapat dimainkan bersamaan dengan kuis, serta memiliki keunggulan dalam memilih mode permainan secara langsung dan mode pekerjaan rumah, interaktif, dan dapat dioperasikan dengan mudah baik guru maupun peserta didik.

METODE PENELITIAN

Penelitian ini adalah jenis R&D (*Research & Development*) yang menggunakan model ADDIE, bertujuan mengembangkan dan menguji efektivitas suatu produk. Model ADDIE melibatkan lima tahap: Analisis, Desain, Pengembangan, Implementasi, dan Evaluasi untuk memastikan pengembangan yang sistematis dan efisien. (Fransisca, 2021). Adapun produk yang akan dikembangkan dalam penelitian ini adalah e-instrumen tes menggunakan *Blooket Game* untuk mengukur kemampuan berpikir kritis pada siswa SMA pada materi stoikiometri. E-Instrumen tes yang akan dirancang berupa soal pilihan ganda berjumlah 20 soal dan memiliki 5 buah opsi jawaban.

Subjek dalam penelitian ini yaitu: (1) 10 peserta didik kelas XI MIPA 1 SMA Negeri 4 Banjarmasin sebagai sampel uji keterbacaan instrumen tes, (2) 34 peserta didik kelas XI MIPA 2 SMA Negeri 4 Banjarmasin sebagai sampel uji kelompok kecil dan uji coba lapangan yang terdiri dari 35 peserta didik kelas XI MIPA 3 dan 35 peserta didik kelas XI MIPA 4 SMA Negeri 4 Banjarmasin sehingga total pada uji coba lapangan ini sebanyak 70 peserta didik, sedangkan objek dalam penelitian ini yaitu e-instrumen tes untuk mengukur kemampuan berpikir kritis dan perangkat penelitian lainnya.

Data dikumpulkan melalui teknik non-tes dan tes. Teknik non-tes menggunakan angket atau kuesioner, yang merupakan metode pengumpulan data dengan daftar pertanyaan atau isian yang diberikan kepada objek penelitian. Angket ini digunakan untuk mengukur validitas, keterbacaan, dan respons peserta didik terhadap soal-soal yang dikembangkan. Sementara itu, teknik tes menggunakan instrumen soal untuk mengukur pencapaian variabel tertentu berdasarkan cara dan aturan yang telah ditetapkan (Ndiung & Jediut, 2020). Adapun tes ini digunakan untuk mengukur reliabilitas, kesukaran dan daya pembeda soal.

Teknik analisis data yang digunakan dalam penelitian ini adalah analisis deskriptif kuantitatif dengan menggunakan kriteria seperti pada Tabel 1 berikut.

Tabel 1. Kriteria analisis data

Parameter	Metode Kalkulasi	Kategori
Validitas	$V = \frac{\sum s}{n(c - 1)}$	<ul style="list-style-type: none"> • $V < 0,4$ (Kurang) • $0,4 < V < 0,8$ (Sedang) • $V > 0,8$ (Sangat Valid) (Retnawati, 2016)
Reliabilitas	Menggunakan SPSS versi 26 dengan memperhatikan <i>output Reliability Statistics</i> pada kolom <i>Guttman Split-Half Coefficient</i> .	<ul style="list-style-type: none"> • $r_i \leq 0,20$ (Sangat rendah) • $0,20 < r_i \leq 0,40$ (Rendah) • $0,40 < r_i \leq 0,60$ (Sedang) • $0,60 < r_i \leq 0,80$ (Tinggi) • $0,80 < r_i \leq 1,00$ (Sangat Tinggi) (Ndiung & Jediut, 2020)
Kesukaran	Menggunakan SPSS versi 26 dengan memperhatikan <i>output statistics</i> pada kolom <i>Mean</i>	<ul style="list-style-type: none"> • 0,00 – 0,30 (Sukar) • 0,31 – 0,71 (Sedang) • 0,71 – 1,00 (Mudah) (Rusmayani, 2020)
Daya Pembeda	$D = P_a - P_b$	<ul style="list-style-type: none"> • $D = -$ (Jelek sekali) • $D < 0,20$ (Rendah) • 0,20 - 0,40 (Sedang) • 0,40 - 0,70 (Baik) • 0,70 - 1,00 (Baik Sekali) (Fatimah & Aifath, 2019)
Keterbacaan	$Skor = \frac{Skor\ Total}{Jumlah\ Butir}$	<ul style="list-style-type: none"> • 3,26 – 4,00 (Sangat Baik) • 2,51 – 3,25 (Baik) • 1,76 – 2,50 (Tidak Baik) • 1,00 – 1,75 (Sangat Tidak Baik)
Respon Peserta Didik	$P = \frac{s}{N} \times 100\%$	<ul style="list-style-type: none"> • $80\% \leq P \leq 100\%$ (Sangat Baik) • $67\% \leq P \leq 79\%$ (Baik) • $53\% \leq P \leq 66\%$ (Cukup Baik) • $40\% \leq P \leq 52\%$ (Kurang Baik) • $P < 39\%$ (Tidak Baik)

HASIL PENELITIAN DAN PEMBAHASAN

Penelitian ini menggunakan model ADDIE yang mana setiap tahapannya merupakan landasan tahapan selanjutnya (Branch, 2010). Pada tahap analisis, dilakukan observasi awal di SMA N 4 Banjarmasin yang menunjukkan bahwa belum ada instrumen interaktif dalam mengevaluasi proses-proses pembelajaran yang dilakukan. Pembelajaran di sekolah ini cenderung tidak menggali kemampuan berpikir kritis secara mendalam. Identifikasi kebutuhan ini menjadi dasar penting untuk pengembangan instrumen yang lebih efektif. Analisis ini mencakup penilaian terhadap kebutuhan siswa dan guru, serta kesenjangan yang ada dalam metode pembelajaran saat ini. Berdasarkan temuan ini, ditetapkan bahwa pengembangan e-instrumen interaktif menggunakan *blooket game* dapat menjadi solusi yang inovatif.

Tahapan selanjutnya adalah tahapan desain yang melibatkan penyusunan instrumen tes berpikir kritis yang sesuai dengan aspek dan indikator yang telah ditetapkan. Aspek-aspek berpikir kritis yang diukur *Focus, Reason, Inference, Situation, Clarity* dan *overview*. Desain ini memastikan bahwa setiap item dalam instrumen mencakup berbagai tingkat kesulitan dan beragam bentuk pertanyaan yang merangsang pemikiran peserta didik untuk berpikir secara kritis. Pada tahap ini, rancangan instrumen disusun secara detail, termasuk format soal, skala penilaian, dan mekanisme penyajian dalam platform *Blooket*.

Hasil desain instrumen tes pada tahapan sebelumnya kemudian diproduksi pada tahapan pengembangan. Pada tahap ini, instrumen yang telah dirancang diinput dalam format digital dengan memanfaatkan fitur-fitur interaktif yang disediakan oleh *Blooket*. Selain itu, dilakukan validasi instrumen untuk memastikan kesesuaian dan ketepatan setiap item soal. Pengujian reliabilitas dilakukan untuk memastikan konsistensi hasil yang diperoleh dari penggunaan instrumen ini. Uji kesukaran dan daya pembeda setiap soal juga dilakukan untuk memastikan bahwa instrumen mampu mengukur berbagai tingkat kemampuan siswa dengan baik. Revisi dan perbaikan perlu dilakukan yang didasarkan pada hasil-hasil uji.

Instrumen tes berpikir kritis yang valid dan reliabel serta memenuhi kriteria kesukaran dan daya pembeda selanjutnya kemudian memasuki tahapan implementasi. Pada tahap ini, e-instrumen yang telah dikembangkan diujicobakan kepada peserta didik di SMA N 4 Banjarmasin. Uji coba ini bertujuan untuk mendapatkan data keterbacaan dan respon siswa terhadap instrumen yang dikembangkan. Implementasi ini juga menjadi kesempatan untuk mengidentifikasi kendala teknis maupun pedagogis yang mungkin muncul saat penggunaan instrumen. Data yang diperoleh dari implementasi ini dianalisis untuk mengevaluasi efektivitas instrumen dalam mengukur kemampuan berpikir kritis.

Tahap terakhir adalah evaluasi yang dilakukan untuk menilai keberhasilan setiap tahapan dalam proses pengembangan e-instrumen. Evaluasi ini mencakup analisis terhadap data yang diperoleh selama tahap implementasi, termasuk tanggapan siswa dan hasil tes berpikir kritis. Berdasarkan evaluasi ini, dilakukan revisi dan penyempurnaan instrumen untuk memastikan bahwa instrumen yang dikembangkan benar-benar efektif dan sesuai dengan tujuan pembelajaran. Evaluasi juga mempertimbangkan umpan balik dari guru dan ahli untuk memastikan kualitas dan keterandalan instrumen sebelum diterapkan secara luas. Adapun hasil-hasil evaluasi yang dilakukan adalah disajikan sebagai berikut.

Validasi Instrumen Tes

Instrumen tes perlu divalidasi untuk memastikan bahwa mereka secara akurat dan konsisten mengukur apa yang dimaksudkan untuk diukur (Farbmacher et al., 2022). Validasi adalah proses kritis dalam penelitian yang menjamin bahwa

instrumen pengukuran memiliki reliabilitas dan validitas yang memadai, sehingga data yang dihasilkan dapat diandalkan untuk mendukung pengambilan keputusan yang tepat (Tavakol & Wetzel, 2020). Proses ini penting untuk mengidentifikasi dan mengurangi kesalahan atau bias yang mungkin ada dalam instrumen, serta memastikan bahwa hasil pengukuran dapat diandalkan dan berlaku umum (Sihotang et al., 2023). Pada penelitian ini dilakukan validasi terkait dan validasi konstruk.

Validasi isi merupakan validasi yang dilakukan untuk memastikan bahwa instrumen mencakup semua aspek atau domain yang relevan dengan konsep yang diukur. Proses ini melibatkan penilaian oleh ahli di bidang terkait untuk mengevaluasi apakah setiap item dalam instrumen benar-benar mewakili konsep yang ingin diukur dan mencakup seluruh spektrum konsep tersebut (Alek, 2021). Validasi ini penting untuk memastikan bahwa instrumen tidak melewatkan elemen penting dari konsep yang diukur dan tidak memasukkan elemen yang tidak relevan (Ayudia et al., 2023). Adapun validasi isi dilakukan ahli materi dalam bidang kimia yaitu dosen pendidikan kimia dan guru kimia. Instrumen yang divalidasi sebanyak 20 butir soal, kemudian divalidasi oleh ahli melalui angket yang diberikan. Hasil validasi diolah menggunakan persamaan aikens kemudian dikategorikan pada Tabel 2 berikut ini.

Tabel 2. Hasil validasi isi instrumen oleh ahli

No. Butir	Skor Aiken	Kriteria	No. Butir	Skor Aiken	Kriteria
1	0,90	Sangat Valid	11	0,95	Sangat Valid
2	0,90	Sangat Valid	12	0,95	Sangat Valid
3	0,90	Sangat Valid	13	0,95	Sangat Valid
4	0,95	Sangat Valid	14	0,95	Sangat Valid
5	0,95	Sangat Valid	15	0,90	Sangat Valid
6	0,90	Sangat Valid	16	0,95	Sangat Valid
7	0,95	Sangat Valid	17	0,95	Sangat Valid
8	0,95	Sangat Valid	18	0,95	Sangat Valid
9	0,90	Sangat Valid	19	0,95	Sangat Valid
10	0,95	Sangat Valid	20	0,85	Sangat Valid

Berdasarkan Tabel 2 di atas menunjukkan hasil penilaian instrumen tes oleh validator menunjukkan 20 butir soal tersebut memiliki kriteria sangat valid. Sehingga seluruh soal dianggap layak untuk dilanjutkan untuk uji atau validasi konstruk. Instrumen yang valid berdasarkan validasi isi oleh ahli menunjukkan bahwa instrumen tersebut secara komprehensif dan tepat mencakup semua aspek relevan dari konsep yang diukur (Pérez-Rivas et al., 2023). Ahli memastikan bahwa setiap item dalam instrumen relevan, representatif, dan jelas, sehingga tidak ada elemen penting yang terlewatkan atau item yang tidak sesuai. Validasi ini memastikan bahwa hasil pengukuran benar-benar mencerminkan konsep yang dimaksud tanpa bias atau kekurangan signifikan.

Validasi konstruk menilai sejauh mana alat evaluasi tersebut benar-benar mengukur konstruk teoretis yang dimaksudkan. Ini melibatkan pengujian hipotesis yang berkaitan dengan bagaimana konstruk tersebut harus berperilaku relatif terhadap konstruk lain (Stenner et al., 2023). Validasi konstruk dilakukan pada dua kelompok yaitu kelompok kecil (*small group*) dan lapangan (*field group*). Butir soal dapat dikatakan valid jika nilai $r_{pbis} > r_{tabel}$, dimana pada penelitian ini nilai r_{tabel} untuk 34 sampel (*Small group*) pada taraf signifikansi 5% adalah 0,349. Hasil analisis uji validitas pada uji coba *small group* dengan rumus korelasi point biserial menggunakan software IBM SPSS Statistic 26. Hasil analisis validitas dapat dilihat pada Tabel 3 berikut ini.

Tabel 3. Validasi Konstruk Kelompok Kecil

No. Butir	R _{pbis}	Kriteria	No. Butir	R _{pbis}	Kriteria
1	0,711	Valid	11	0,073	Tidak Valid
2	0,728	Valid	12	0,583	Valid
3	0,668	Valid	13	-0,061	Tidak Valid
4	0,474	Valid	14	0,304	Tidak Valid
5	0,768	Valid	15	0,547	Valid
6	0,366	Valid	16	0,241	Tidak Valid
7	0,505	Valid	17	0,030	Tidak Valid
8	0,479	Valid	18	0,562	Valid
9	0,684	Valid	19	0,902	Valid
10	0,441	Valid	20	0,841	Valid

Berdasarkan hasil analisis validitas terhadap 20 butir soal pada Tabel 3 di atas, didapat butir soal nomor 11, 13, 14, 16, dan 17 memiliki nilai koefisien korelasi kurang dari 0,349 ($r_{pbis} < r_{tabel}$) sehingga butir soal tersebut tidak dapat diikutsertakan dalam tahapan penelitian analisis reliabilitas. Pada tahap uji coba kelompok kecil (small group), diminta 34 peserta didik kelas XI MIPA 2 SMA Negeri 4 Banjarmasin untuk mengerjakan instrumen tes yang dikembangkan. Dari 20 butir soal pilihan ganda yang diuji cobakan pada tahap ini, terdapat 5 butir soal dinyatakan tidak valid, yaitu pada butir soal nomor 11, 13, 14, 16, dan 17. Hal tersebut dikarenakan, nilai koefisien korelasi (r_{pbis}) butir tersebut kurang dari nilai r_{tabel} dengan *degree freedom* (df) 32 yaitu $0,349$ ($r_{pbis} < r_{tabel}$, $df = N - 2 = 34 - 2 = 0,349$). Hamimi et al. (2020) menyatakan bahwa butir soal yang tidak valid dapat dibuang atau direvisi kembali, karena butir soal tersebut tidak dapat mengukur apa yang seharusnya diukur sesuai dengan variabel yang dimaksud, sehingga dalam penelitian ini butir soal yang tidak valid tersebut akan dibuang dan tidak diikutkan pada tahap uji lapangan (*field group*). Adapun hasil uji coba terbatas disajikan pada Tabel 4 berikut.

Tabel 4. Validasi konstruk kelompok terbatas

No. Butir	R _{pbis}	Kriteria
1	0,576	Valid
2	0,524	Valid
3	0,516	Valid
4	0,565	Valid
5	0,630	Valid
6	0,539	Valid
7	0,591	Valid
8	0,591	Valid
9	0,563	Valid
10	0,554	Valid
12	0,583	Valid
15	0,433	Valid
18	0,365	Valid
19	0,509	Valid
20	0,564	Valid

Berdasarkan hasil analisis validitas pada uji coba lapangan di atas, diperoleh sebanyak 15 buah butir soal dinyatakan valid. Sehingga untuk tahapan analisis reliabilitas berikutnya, butir soal tersebut diikutsertakan semua. Berdasarkan hasil analisis uji coba instrumen tes pada kelompok kecil, didapat 15 butir soal yang telah direvisi dianggap layak untuk lanjut diimplementasikan pada uji coba lapangan. Uji coba lapangan dilakukan setelah melalui pengujian pada uji coba kelompok kecil dan melakukan perbaikan. Uji coba ini melibatkan kelompok yang lebih besar, yakni total

sebanyak 70 peserta didik yang terdiri dari peserta didik kelas XI MIPA 3 dan XI MIPA 4 SMA Negeri 4 Banjarmasin.

Uji coba lapangan membantu dalam mengidentifikasi kualitas instrumen tes dengan melibatkan berbagai sampel yang bervariasi untuk membantu mengungkap permasalahan atau kekurangan yang belum terungkap pada tahap uji coba kelompok kecil (Lourenço, 2019). Pada uji coba lapangan (*field test*), peserta didik diminta untuk mengerjakan instrumen tes yang dikembangkan sebanyak 15 butir soal. Berdasarkan hasil analisis validitas secara empiris, dari 15 butir soal pilihan ganda yang diuji cobakan semua butir soal dinyatakan valid. Hal tersebut dikarenakan, nilai koefisien korelasi (r_{pbis}) butir tersebut lebih dari nilai r_{tabel} dengan *degree freedom* (df) 68 yaitu 0,239 ($r_{pbis} > r_{tabel}$) ($df = N - 2 = 70 - 2 = 68$).

Reliabilitas Instrumen Tes

Instrumen tes dikatakan reliabel jika nilai koefisien reliabilitas (r_{hitung}) lebih besar atau sama dengan r_{tabel} ($r_{hitung} > r_{tabel}$) dimana pada penelitian ini nilai r_{tabel} untuk 70 sampel pada taraf signifikansi 5% adalah 0,239. Tingkat reliabilitas instrumen tes yang dikembangkan dianalisis menggunakan software IBM SPSS Statistic 26 dengan metode *split – half* menggunakan rumus *Spearman - Brown*, yaitu sebanyak 15 butir soal pada nomor 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 15, 18, 19, 20. Berdasarkan hasil analisis reliabilitas didapat nilai koefisien reliabilitas sebesar 0,885. Hasil tersebut menunjukkan berdasarkan pengujian pada uji coba kelompok kecil, instrumen tes yang dikembangkan ialah reliabel dengan klasifikasi tinggi. Widoyoko (2020) menjelaskan apabila r_{hitung} lebih besar daripada r_{tabel} diartikan ada korelasi yang signifikan, instrumen dianggap reliabel. Hal ini berarti instrumen tes yang dikembangkan jika dilakukan uji coba berulang kali pada objek yang sama, maka akan memberikan hasil yang tetap atau tidak berbeda secara signifikan.

Tingkat Kesukaran dan Daya Pembeda Instrumen Tes

Tingkat kesukaran soal mengukur seberapa sulit suatu soal bagi peserta tes, biasanya dinyatakan dalam proporsi peserta yang menjawab soal tersebut dengan benar; soal dengan tingkat kesukaran rendah dijawab benar oleh banyak peserta, sementara soal dengan tingkat kesukaran tinggi dijawab benar oleh sedikit peserta (Ningsih et al., 2018). Sedangkan daya pembeda soal mengukur kemampuan soal untuk membedakan antara peserta yang memiliki kemampuan tinggi dan rendah dalam keseluruhan tes. Soal dengan daya pembeda tinggi dapat menunjukkan perbedaan yang jelas antara peserta yang benar-benar memahami materi dan yang tidak, sementara soal dengan daya pembeda rendah tidak mampu membedakan dengan baik antara kedua kelompok tersebut. Kedua aspek ini penting untuk memastikan bahwa soal dalam tes tidak hanya mencakup berbagai tingkat kesulitan tetapi juga efektif dalam menilai kemampuan peserta secara akurat.

Analisis tingkat kesukaran butir soal dianalisis menggunakan software IBM SPSS Statistic 26 dan analisis daya pembeda butir dianalisis secara manual menggunakan software Microsoft Excel. Berdasarkan analisis kesukaran dan daya pembeda butir soal pada instrumen tes yang telah diuji cobakan, diperoleh data kesukaran daya pembeda butir soal pada Tabel 5 berikut ini.

Tabel 5. Tingkat Kesukaran dan Daya Pembeda Butir Soal

No. Butir	Daya Beda	Kriteria	No. Butir	Kesukaran	Kriteria
1	0,31	Sedang	1	0,70	Sedang
2	0,33	Sedang	2	0,65	Sedang
3	0,33	Sedang	3	0,71	Mudah
4	0,33	Sedang	4	0,71	Mudah

No. Butir	Daya Beda	Kriteria	No. Butir	Kesukaran	Kriteria
5	0,31	Sedang	5	0,67	Sedang
6	0,31	Sedang	6	0,55	Sedang
7	0,31	Sedang	7	0,55	Sedang
8	0,31	Sedang	8	0,55	Sedang
9	0,31	Sedang	9	0,68	Sedang
10	0,31	Sedang	10	0,58	Sedang
11	0,31	Sedang	11	0,64	Sedang
12	0,31	Sedang	12	0,50	Sedang
13	0,31	Sedang	13	0,78	Mudah
14	0,33	Sedang	14	0,57	Sedang
15	0,33	Sedang	15	0,40	Sedang

Berdasarkan hasil analisis tingkat kesukaran butir pada tabel 5 di atas, diperoleh sebanyak 3 butir soal diklasifikasikan sebagai butir soal mudah dikarenakan memiliki indeks kesukaran di atas 0,71 dan 12 butir lainnya diklasifikasikan sebagai butir soal yang sedang karena berada dalam skala indeks kesukaran di antara 0,31 – 0,71. Kualitas butir soal pada instrumen ini juga diukur dengan mengetahui tingkat kesukaran butir soal. Berdasarkan hasil analisis tingkat kesukaran butir soal yang dilakukan diperoleh semua butir soal, berjumlah 20 butir soal diklasifikasikan sedang. Rusmayani (2020) dalam penelitiannya menyebutkan bahwa indeks kesukaran pada rentang 0,30 – 0,70 (sedang) merupakan rentang suatu butir tes tersebut diterima dan dapat digunakan. Adanya pendapat bahwa soal – soal dianggap baik, yaitu soal – soal yang diklasifikasikan sedang atau memiliki indeks kesukaran 0,30 – 0,70. Perlu dicatat bahwa tingkat kesulitan soal yang terlalu mudah atau terlalu sulit tidak selalu berarti tidak boleh digunakan. Penggunaannya tergantung pada kegunaan atau tujuan dari tes tersebut (Widiyanto, 2018). Fungsi dari tingkat kesulitan suatu soal biasanya dikaitkan dengan tujuan tes yang ditetapkan (Depdiknas, 2008).

Rezigalla (2022) menjelaskan bahwa indeks kesukaran sebuah soal memiliki nilai antara 0,00 hingga 1,00, dimana semakin tinggi nilai indeks kesukaran menunjukkan tingkat kesulitan soal yang lebih rendah, dan sebaliknya. Hasil analisis tingkat kesukaran butir soal dalam uji coba lapangan menunjukkan bahwa 80% (12 butir soal) diklasifikasikan memiliki tingkat kesukaran sedang, sementara 20% (3 butir soal) lainnya diklasifikasikan sebagai soal yang mudah. Hasil ini berbeda dengan uji coba pada kelompok kecil dimana 100% (20 butir soal) memiliki tingkat kesukaran sedang dan tidak ada soal yang diklasifikasikan sebagai mudah.

Berdasarkan analisis daya pembeda yang tertera pada Tabel 5, sebanyak 15 butir soal memiliki daya pembeda yang sedang. Penurunan nilai indeks diskriminan dapat terjadi karena beberapa faktor seperti pengecoh yang tidak efektif, tingkat kesulitan soal yang terlalu tinggi sehingga banyak peserta yang menebak, serta adanya ketidaksesuaian informasi dalam soal menurut siswa yang menguasai materi (Depdiknas, 2008). Semakin tinggi nilai indeks diskriminasi menandakan kualitas butir soal yang lebih baik, karena mampu membedakan antara peserta yang memiliki pengetahuan mendalam dan yang tidak.

Tindak lanjut yang dilakukan terhadap butir soal dengan daya pembeda baik (sedang hingga baik sekali) adalah mempertahankan butir soal tersebut. Menurut Rusmayani (2020), sebuah butir soal yang baik harus dapat membedakan antara peserta tes yang menguasai materi dengan yang tidak. Sehubungan dengan itu, peneliti melakukan uji ulang terhadap daya pembeda butir soal, dan hasilnya menunjukkan bahwa 15 butir soal memiliki klasifikasi daya pembeda sedang. Hasil ini sesuai dengan pandangan Retnawati (2016) bahwa butir soal yang ideal memiliki indeks kesukaran antara 0,31 hingga 0,70.

Keterbacaan Instrumen Tes

Uji keterbacaan instrumen butir soal adalah proses evaluasi yang bertujuan untuk memastikan bahwa soal-soal dalam instrumen tes dapat dibaca dan dipahami dengan mudah oleh peserta tes (Chiodi et al., 2023). Uji ini melibatkan pemeriksaan aspek-aspek seperti kejelasan bahasa, struktur kalimat, panjang paragraf, dan penggunaan istilah yang sesuai dengan tingkat pemahaman peserta. Menurut Alkhurayyif & Weir (2017) uji keterbacaan penting dilakukan untuk mengidentifikasi dan memperbaiki potensi kesulitan yang tidak terkait dengan kemampuan yang diukur, seperti kebingungan atau salah pengertian akibat kalimat yang rumit atau terminologi yang tidak familiar. Adapun uji keterbacaan dilakukan dengan 7 butir penilaian yaitu sebagai berikut (hasil uji keterbacaan disajikan pada Tabel 6):

1. Petunjuk pengerjaan soal yang diberikan mudah dipahami dan jelas
2. Bahasa yang digunakan dalam instrumen tes secara umum mudah dipahami
3. Bahasa yang digunakan dalam instrumen tes jelas sesuai kaidah penulisan Ejaan Yang Disempurnakan (EYD)
4. Kata - kata yang digunakan mudah dipahami atau tidak ambigu
5. Bentuk, model, ukuran huruf dan spasi kalimat yang digunakan mudah dibaca
6. Soal – soal yang terdapat dalam instrumen tes dapat dipahami kejelasannya
7. Soal – soal yang terdapat dalam instrumen tes sesuai dengan materi yang diajarkan oleh guru

Tabel 6. Hasil uji keterbacaan instrumen tes

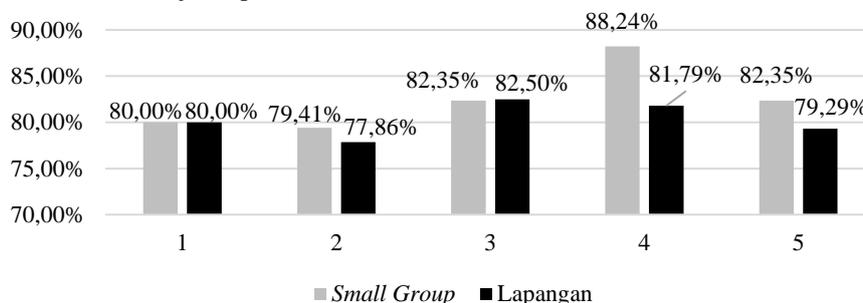
Butir Penilaian	Skor	Kriteria
1	3,33	Sangat Baik
2	3,12	Baik
3	3,11	Baik
4	3,22	Baik
5	3,88	Sangat Baik
6	3,77	Sangat Baik
7	3,66	Sangat Baik

Berdasarkan Tabel 6 di atas menunjukkan bahwa semua butir penilaian keterbacaan berada pada kategori Baik dan Baik sekali. Hal ini menunjukkan bahwa hasil evaluasi keterbacaan dari butir-butir soal menunjukkan kualitas yang tinggi. Ini mengindikasikan bahwa setiap soal dalam instrumen tes dapat dipahami dengan jelas dan mudah oleh peserta, tanpa kebingungan atau kesalahpahaman. Penilaian dalam kategori "Baik" dan "Baik sekali" menunjukkan bahwa aspek-aspek seperti kejelasan bahasa, struktur kalimat, panjang paragraf, dan penggunaan istilah telah memenuhi standar yang diperlukan untuk memastikan keterbacaan yang optimal. Dengan demikian, peserta tes dapat fokus pada konten dan kemampuan yang diuji, bukan pada upaya memahami soal.

Respon Peserta Didik

Pengumpulan data terhadap respon peserta didik dilakukan untuk membantu memastikan bahwa soal-soal tersebut sesuai dengan tingkat kemampuan dan pemahaman peserta didik, sehingga dapat memberikan penilaian yang akurat terhadap kompetensi mereka. *Feedback* dari peserta didik juga dapat mengungkapkan masalah-masalah potensial dalam soal, seperti ketidakjelasan, kesulitan yang tidak sesuai, atau bias, yang mungkin tidak terdeteksi oleh penguji. Pemahaman terhadap respons peserta didik dapat dilakukan penyesuaian dan perbaikan instrumen untuk penggunaan di masa depan, meningkatkan validitas dan reliabilitas tes. Respon peserta didik diukur menggunakan lima aspek yaitu (1) Pemahaman, (2)

Menyelesaikan soal, (3) Interaksi, (4) Persepsi, dan (5) Adaptasi. Adapun hasil respon peserta didik disajikan pada Gambar 1 berikut.



Gambar 1. Hasil respon peserta didik

Berdasarkan Gambar 1 di atas, terlihat bahwa respons peserta didik terhadap butir soal atau instrumen yang telah dikembangkan menunjukkan hasil yang positif baik saat uji coba pada kelompok kecil maupun pada uji coba lapangan. Respons positif ini mengindikasikan bahwa secara umum, soal-soal tersebut dapat dipahami dengan baik dan dianggap relevan oleh peserta didik. Ini mencerminkan kualitas instrumen yang baik, karena telah mampu memberikan penilaian yang akurat dan valid terhadap kemampuan peserta didik dalam kondisi uji coba yang berbeda (Mayuri & Aswirna, 2021).

Pada hasil respon peserta didik terjadi penurunan respons dari kelompok kecil ke kelompok besar di semua aspek yang diuji. Penurunan ini disebabkan oleh berbagai faktor, seperti peningkatan jumlah peserta yang memperkenalkan lebih banyak variabilitas dalam pemahaman dan persepsi terhadap soal-soal tersebut. Dalam kelompok besar, variasi dalam kemampuan, latar belakang, dan pengalaman belajar peserta didik lebih beragam, yang bisa mengungkapkan kekurangan atau ketidakjelasan dalam instrumen yang tidak terlihat pada kelompok kecil. Penurunan respons ini mengindikasikan perlunya penyesuaian dan perbaikan lebih lanjut pada instrumen agar dapat mempertahankan kualitas dan keterbacaan yang tinggi di berbagai skala penggunaan.

SIMPULAN

Berdasarkan hasil dan analisis pengembangan instrumen tes berpikir kritis berbasis game Blooket di SMA Negeri 4 Banjarmasin menunjukkan hasil yang positif. Validitas isi instrumen diperoleh dari penilaian ahli dengan skor Aiken rata-rata 0,90 hingga 0,95, menunjukkan bahwa instrumen ini sangat valid. Dari 20 butir soal yang diuji, 15 butir soal dinyatakan valid dengan nilai koefisien korelasi (rpbis) lebih dari 0,349. Reliabilitas instrumen diuji menggunakan koefisien alpha Cronbach, yang menunjukkan hasil tinggi yaitu 0,88, menandakan konsistensi yang baik. Uji daya beda menghasilkan butir soal dengan daya beda berkisar antara 0,30 hingga 0,68, yang menunjukkan bahwa instrumen ini mampu membedakan dengan baik antara siswa berkemampuan tinggi dan rendah. Tingkat kesukaran butir soal bervariasi antara 0,41 hingga 0,78, mencakup kategori sedang hingga sulit. Uji keterbacaan menunjukkan bahwa instrumen ini memiliki tingkat keterbacaan yang baik sekali menurut respon siswa. Respon peserta didik terhadap penggunaan instrumen berbasis Blooket juga sangat positif, dengan siswa merasa lebih termotivasi dan tertarik untuk mengikuti tes.

DAFTAR RUJUKAN

- Agustini, Rery, R. U., & Anwar, L. (2021). The Development and Validity of CPS-Based Assessment Instrument for Critical Thinking Ability on Stoichiometry Materials. *Journal of Physics: Conference Series*, 1788(1). <https://doi.org/10.1088/1742-6596/1788/1/012026>
- Alek. (2021). *Perspektif Baru Bagi Calon Peneliti*. CV. Mutiara Galuh.
- Ayudia, I., Bhoke, W., Oktari, R., Carmelita, M., Salem, V., Khairani, M., Mamontho, F., & Setiawati, M. (2023). *Pengembangan Kurikulum*. PT. Mifandi Mandiri Digital.
- Bakti, I., Kusasi, M., & Samsinar, S. PERBEDAAN KEMAMPUAN BERPIKIR KRITIS SISWA PADA PEMBELAJARAN MENGGUNAKAN MODEL PROJECT BASED LEARNING PADA MATERI KOLOID BERKONTEKS LAHAN BASAH. *Quantum: Jurnal Inovasi Pendidikan Sains*, 14(1), 130-141. <http://dx.doi.org/10.20527/quantum.v14i1.15694>
- Branch, R. M. (2010). Instructional design: The ADDIE approach. In *Instructional Design: The ADDIE Approach*. Springer US. <https://doi.org/10.1007/978-0-387-09506-6>
- Farbmacher, H., Guber, R., & Klaassen, S. (2022). Instrument Validity Tests With Causal Forests. *Journal of Business and Economic Statistics*, 40(2), 605–614. <https://doi.org/10.1080/07350015.2020.1847122>
- Fatimah, U., & Aifath, K. (2019). Analisis Kesukaran Soal, Daya Pembeda dan Fungsi Distraktor. *Jurnal Komunikasi Dan Pendidikan Islam*, 8(2), 37–64.
- Fitriani, F., Priatmoko, S., Wardani, S., & Nurhayati, S. (2024). Pengembangan Instrumen Tes Three-Tier Multiple Choice Berbantuan Google Forms untuk Mengukur Kemampuan Berpikir Kritis Siswa pada Topik Stoikiometri. *Jurnal Inovasi Pendidikan Kimia*, 18(1), 50–57. <https://doi.org/10.15294/jipk.v18i1.46778>
- Fransisca, S. M. P. L. A. (2021). Pengembangan Video Pembelajaran Berbasis Kinemaster pada Materi Larutan Elektrolit dan Non Elektrolit Terintegrasi Etnosains untuk Kelas X SMA/MA. *Jurnal Beta Kimia*, 1(2), 89–91. <https://doi.org/10.35508/jbk.v1i2.5585>
- Gunawan, A., Fitriani, N., & Setiawan, W. (2023). Kemampuan Berpikir Kritis Siswa dalam Menyelesaikan Soal Persamaan Linier Satu Variabel Berdasarkan Perbedaan Gender. *Jurnal Pembelajaran Matematika Inovatif*, 6(5). <https://doi.org/10.22460/jpmi.v6i5.18844>
- Händel, M., Artelt, C., & Weinert, S. (2023). Assessing metacognitive knowledge: development and evaluation of a test instrument. *Journal for Educational Research Online*, 5(2), 162–188. <https://doi.org/10.25656/01:8429>
- Khairunnisa, K., Saadi, P., & Leny, L. (2022). Implementasi Model Problem Based Learning Berbasis Stem Dengan Media Virtual Reality Terhadap Kemampuan Berpikir Kritis Dan Self Regulation Peserta Didik. *Quantum: Jurnal Inovasi Pendidikan Sains*, 13(1), 96-108. <http://dx.doi.org/10.20527/quantum.v13i1.12109>
- Lourenço, S. M. (2019). Field experiments in managerial accounting research. *Foundations and Trends in Accounting*, 14(1), 1–72. <https://doi.org/10.1561/14000000059>
- Mayuri, D., & Aswirna, P. (2021). Development of assessment instruments assisted by iSpring quizmaker to measure students' conceptual understanding. *Jurnal Inovasi Pendidikan IPA*, 7(2), 213. <https://doi.org/10.21831/jipi.vXiY.00001>
- Ndiung, S., & Jediut, M. (2020). Pengembangan instrumen tes hasil belajar matematika peserta didik sekolah dasar berorientasi pada berpikir tingkat tinggi.

- Premiere Educandum : Jurnal Pendidikan Dasar Dan Pembelajaran*, 10(1), 94.
<https://doi.org/10.25273/pe.v10i1.6274>
- Ningsih, N., Saadi, P., & Irhasyuarna, Y. (2018). Penerapan Model Learning Cycle 5e Dalam Meningkatkan Motivasi Dan Hasil Belajar Siswa Pada Mata Pelajaran Koloid Di Kelas Xi Mia-3 Man 1 Banjarmasin. *Jcae (Journal Of Chemistry And Education)*, 1(3), 209-217.
- Octaria, D., Ningsih, Y. L., Destiniar, D., & Fitriasari, P. (2022). Development of Higher Order Thinking Skills (HOTS) Test Instruments in Geometry. *JTAM (Jurnal Teori Dan Aplikasi Matematika)*, 6(4), 834.
<https://doi.org/10.31764/jtam.v6i4.8808>
- Pérez-Rivas, F. J., Jiménez-González, J., Bayón Cabeza, M., Belmonte Cortés, S., de Diego Díaz-Plaza, M., Domínguez-Bidagor, J., García-García, D., Gómez Puente, J., & Gómez-Gascón, T. (2023). Design and Content Validation using Expert Opinions of an Instrument Assessing the Lifestyle of Adults: The ‘PONTE A 100’ Questionnaire. *Healthcare (Switzerland)*, 11(14).
<https://doi.org/10.3390/healthcare11142038>
- Puspita, V., & Parma Dewi, I. (2021). Efektifitas E-LKPD berbasis Pendekatan Investigasi terhadap Kemampuan Berfikir Kritis Siswa Sekolah Dasar. *Jurnal Cendekia: Jurnal Pendidikan Matematika*, 5(1), 86–96.
- Puspita, V., Yuhelman, N., & Rifandi, R. (2018). Dampak Pendekatan Realistic Mathematics Education Terhadap Keterampilan Berpikir Kritis Pada Siswa Sekolah Dasar. *JUSTEK: Jurnal Sains Dan Teknologi*, 1(2), 20–25.
<https://doi.org/10.31764/justek.vXiY.3735>
- Putri, Y. D., & Dwijayanti, R. (2021). Pengembangan alat evaluasi berbantuan aplikasi android pada mata pelajaran penataan produk kelas XI BDP di SMK Negeri 10 Surabaya. *Jurnal Pendidikan Tata Niaga (JPTN)*, 9(1), 1041-1047.
<https://doi.org/10.26740/jptn.v9n1.p1041-1047>
- Redhana, W. (2019). Mengembangkan Keterampilan Abad Ke-21 dalam Pembelajaran Kimia. *Jurnal Inovasi Pendidikan Kimia*, 13(1), 2239–2253.
<https://doi.org/10.15294/jipk.v13i1.17824>
- Rosalinda, E., Hamid, A., & Kusasi, M. (2023). PENGEMBANGAN E-LKPD BERBASIS MODEL PROBLEM SOLVING UNTUK MENINGKATKAN KEMAMPUAN BERPIKIR KRITIS PESERTA DIDIK PADA MATERI LARUTAN ELEKTROLIT DAN NON ELEKTROLIT DALAM KONTEKS LAHAN BASAH. *JCAE (Journal of Chemistry And Education)*, 6(2), 64-76.
<https://doi.org/10.20527/jcae.v6i2.1685>
- Rusmayani. (2020). Analisis Butir Soal Penilaian Akhir Semester Genap Mata Pelajaran Pendidikan Agama Islam di SMP Bintang Persada Tabanan-Bali. *Jurnal Widya Balina*, 5(1), 41–49.
- Sihotang, H., Pd, M., Penerbitan, P., Buku, P., & Tinggi, P. (2023). *Metodologi Penelitian Kuantitatif*. UKI Press.
- Stenner, A. J., Smith, M., & Burdick, D. S. (2023). Toward a Theory of Construct Definition. In *Explanatory Models, Unit Standards, and Personalized Learning in Educational Measurement* (pp. 43–55). Springer Nature Singapore.
https://doi.org/10.1007/978-981-19-3747-7_4
- Sutinah, Fariati, & Sulistina, O. (2016). Efektivitas Penerapan Modul Stoikiometri Berdasarkan Konsep Sukar dan Kesalahan Konsep pada Pemahaman dan Persepsi Peserta Didik Kelas X MIA SMAN 1 Malang. *Jurnal Pembelajaran Kimia (J-PEK)*, 01(02), 1–8.
- Tavakol, M., & Wetzel, A. (2020). Factor Analysis: a means for theory and instrument development in support of construct validity. In *International journal of medical*

- education* (Vol. 11, pp. 245–247). NLM (Medline).
<https://doi.org/10.5116/ijme.5f96.0f4a>.
- Retnawati, H. (2016). *Analisis kuantitatif instrumen penelitian (panduan peneliti, mahasiswa dan psikometrian)*. Yogyakarta: Parama Publishing.
- Sholahuddin, A., Analita, R. N., Syahmani, S., Winarti, A., Hamid, A., Suharto, B., & Bakti, I. (2021). Penguatan Kompetensi Profesional Guru MGMP Kimia: Pengembangan Instrumen Evaluasi Diagnostik Multi-tier. *Bubungan Tinggi: Jurnal Pengabdian Masyarakat*, 3(2), 113-119.
<https://doi.org/10.20527/btjpm.v3i2.2809>
- Widoyoko, E. P. (2020). *Teknik Penyusunan Instrumen Penelitian*. Yogyakarta: Pustaka Belajar.
- Winarti, A., Tirto, F. R., Aprilia, A. D., Raihana, E., & Hidayati, N. (2017, May). The Development of "Chemtective" Game-Based Medium on Chemistry Learning. In *5th SEA-DR (South East Asia Development Research) International Conference 2017 (SEADRIC 2017)* (pp. 18-21). Atlantis Press.