

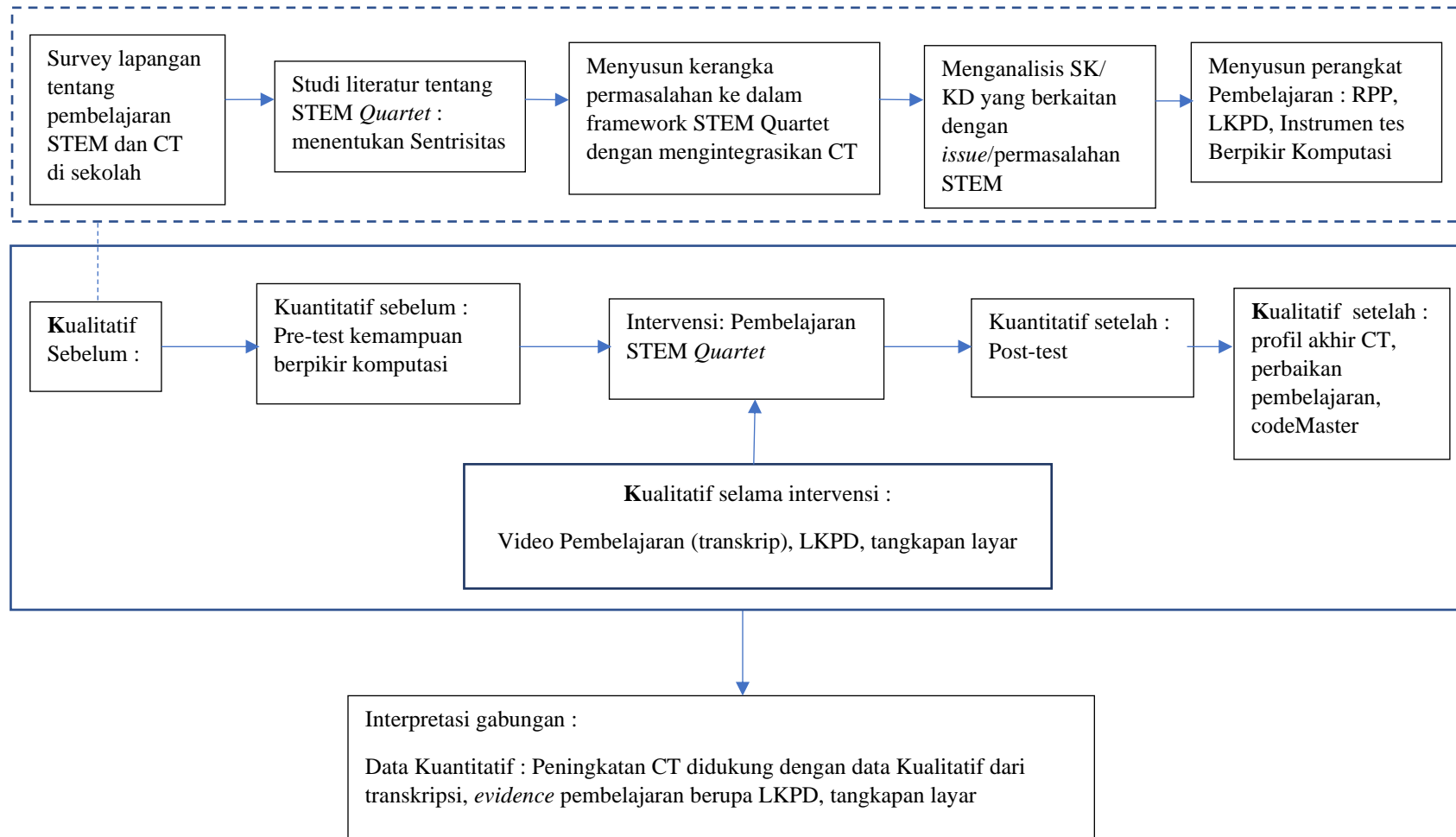
BAB III

METODE PENELITIAN

3.1 Desain Penelitian

Penelitian ini mengungkapkan bagaimana penerapan integrasi STEM *Quartet* dalam meningkatkan kemampuan berpikir komputasi siswa sehingga metode yang digunakan berupa metode campuran kuantitatif dan kualitatif. Oleh karena itu metode penelitian yang digunakan adalah *Mixed Methods*. Pada penelitian ini kedua bentuk data (data kuantitatif dan kualitatif) terintegrasi dalam analisis desain melalui penggabungan data, maka desain penelitian yang digunakan adalah *Embedded Mixed Methods Design* (Creswell & Creswell, 2018). Penelitian STEM *Quartet* dalam melatih CT relatif baru dilaksanakan, oleh karena itu metode kuantitatif yang digunakan pada penelitian ini ialah metode *pre-experimental*. Metode ini dapat digunakan untuk menilai apakah penjelasan mengenai suatu intervensi dapat diteliti lebih lanjut atau tidak. Untuk melihat bagaimana hasil dari suatu intervensi, maka dilakukan sebuah pengamatan sebelum dan sesudah intervensi. Oleh karena itu, desain yang digunakan adalah *one group pretest-posttest design*.

Metode kualitatif yang digunakan adalah metode kualitatif deskriptif. Metode ini digunakan untuk memperoleh gambaran mengenai karakteristik pembelajaran STEM *Quartet*. Karakteristik pembelajaran dianalisis melalui transkrip pembelajaran yang diperoleh dari rekaman video, serta *evidence* pembelajaran berupa LKPD, tangkapan layar serta hasil penilaian CodeMaster. Metode analisis pembelajaran berdasarkan transkrip ini dinamakan *Transcript Based Lesson Analysis*. Adapun desain penelitian yang akan dilaksanakan pada penelitian ini ditunjukkan oleh Gambar 3.1.



Gambar 3. 1 Desain Penelitian

3.2 Populasi dan Sampel Penelitian

Populasi pada penelitian ini yaitu siswa di salah satu SMA di Kota Subang. Adapun teknik sampling yang digunakan yaitu menggunakan *purposive sampling*. *Purposive sampling* merupakan teknik pengambilan sampel *non random* berdasarkan suatu pertimbangan-pertimbangan tertentu yang diyakini peneliti agar memperoleh sampel yang dapat merepresentasikan populasi (Fraenkel et al., 2012; Palinkas et al., 2015). Adapun kriteria yang menjadi pertimbangan adalah dengan memilih kelas XI IPA dimana topik kalor dan perpindahannya berada pada kurikulum kelas XI pada pelajaran Fisika. Ukuran sampel yang digunakan pada penelitian ini didasarkan pada (Fraenkel et al., 2012) yang menyatakan bahwa pada penelitian eksperimen, jumlah sampel yang diperlukan ialah lebih dari 30 orang. Penentuan jumlah sampel ini dipilih supaya sampel cukup merepresentasikan populasi.

3.3 Instrumen Penelitian

Pada penelitian ini terdapat instrumen tes dan non-tes.

1. Instrumen Tes Berpikir Komputasi

Instrumen tes dalam penelitian ini berupa soal dalam bentuk Pilihan Ganda dengan satu jawaban benar dan 4 pilihan lainnya berupa distraktor. Beberapa kelebihan soal pilihan ganda yaitu terletak pada kemudahan dalam mengevaluasi dan dapat digunakan pada kemampuan berpikir tingkat tinggi dengan *range* yang lebih luas sehingga dapat mengukur banyak indikator. Namun kelemahan soal pilihan ganda ini yaitu responden dapat mengerjakan soal dengan cara menebak. Untuk meminimalisir siswa menebak pilihan maka soal dilengkapi dengan alasan. Adapun indikator CT yang menjadi penilaian dalam tes adalah dekomposisi, abstraksi, algoritma, generalisasi, dan evaluasi. Selanjutnya kisi-kisi instrumen tes *Computational Thinking* disajikan dalam Tabel 3.1

Tabel 3. 1 Kisi-kisi Instrumen Tes *Computational Thinking*

<i>CT Skill</i>	Bagaimana cara mengukurnya	Kode	Nomor Soal	Jumlah soal per segmen
Abstraksi (A)	Memilah detail yang tak perlu; Menemukan elemen kunci suatu masalah; Memilih representasi suatu sistem	A1 A2 A3	2 13 11	3
Pemikiran Algoritma (PA)	Berpikir dalam sekuensial dan aturan-aturan tertentu Menjalankan suatu algoritma; Membuat suatu algoritma	PA1 PA2 PA3	3,7,12	3
Dekomposisi (D)	Mengurai suatu tugas/masalah; Berpikir tentang lokus masalah dalam bagian masalah-masalah itu; Membuat keputusan tentang membagi sub-tugas/ sub masalah untuk terintegrasi dalam pikiran misalnya deduksi	D1 D2 D3	1,9	2
Evaluasi (E)	Menemukan solusi terbaik dari alternatif solusi yang ditemukan; Membuat keputusan tentang penggunaan sumber daya yang ada; (Solusi dan keputusan itu) cocok untuk tujuan	E1 E2	5 10	2
Generalisasi (G)	Mengidentifikasi pola, kesamaan dan koneksi; Memecahkan masalah baru berdasarkan masalah yang telah diselesaikan; Memanfaatkan solusi umum, misalnya induksi	G1 G2 G3	4,8 6	3
Jumlah soal				13

Terdapat tiga *issue* yang dikembangkan menjadi bacaan dalam instrumen tes. Issue tersebut mengandung sejumlah informasi terkait dengan permasalahan yang diperlukan penyelesaian dengan cara berpikir komputasi. Kerangka ini dibuat dengan menyisipkan koneksi/hubungan antara subjek STEM. Kerangka ini didasarkan pada kerangka STEM *Quartet* dimana suatu permasalahan/issue dapat diselesaikan memiliki keterkaitan dengan berbagai disiplin ilmu pengetahuan. Secara detail setiap pertanyaan pada masing-masing issue dimunculkan aspek berpikir komputasi dan koneksi antar subjek STEM ditunjukkan oleh Tabel 3.2.

Tabel 3. 2 Aspek Berpikir Komputasi dan Koneksi antar Subjek STEM dalam Instrumen Tes

<i>Issue</i>	Aspek Komputasi	Berpikir	Indikator yang diukur	Nomor Soal	Koneksi antara Subjek STEM
Wajan dan Memasak <i>Steak</i>	Dekomposisi		Mengurai suatu tugas/masalah	SO1	S-E
	Abstraksi		Memilah item yang tidak perlu/perlu	SO2	S-T
	Pemikiran Algoritma Generalisasi		Membuat algoritma	SO3	S-E
			Mengidentifikasi pola, kesamaan dan koneksi	SO4	S-E
	Evaluasi		Memecahkan masalah baru berdasarkan masalah yang telah diselesaikan; Menemukan solusi terbaik dari alternatif solusi yang ditemukan;	SO6 SO5	S-M S-E
Sistem Pendingin Cairan untuk Komputer	Pemikiran Algoritma		Membuat algoritma	SO7	S-E
	Generalisasi		Mengidentifikasi pola, kesamaan dan koneksi	SO8	S-E
	Dekomposisi		Mengurai suatu tugas/masalah	SO9	S-T
	Evaluasi		Membuat keputusan tentang penggunaan sumber daya yang ada	SO10	S-E
Membuat Simulasi	Abstraksi		Memilih representasi suatu sistem	SO11	S-T
Perpindahan Kalor	Pemikiran Algoritma		Membuat suatu algoritma	SO12	S-E
	Abstraksi		Menemukan elemen kunci suatu masalah	SO13	T-E

Hubungan antar subjek STEM :

S-E : *Science – Engineering*

S-T : *Science – Technology*

S-M: *Science – Mathematics*

T-E : *Technology – Engineering*

2. Instrumen Non-Tes

a. Video dan Transkrip Pembelajaran

Video pembelajaran diperoleh dari aktivitas pembelajaran yang direkam. Video ini kemudian dibuat transkripsi. Transkrip audio dari video pembelajaran di kelas digunakan untuk menganalisis karakteristik pembelajaran STEM *Quartet*.

b. *Evidence* Pembelajaran

Hani Sulsilah, 2023

ANALISIS PENERAPAN PEMBELAJARAN STEM QUARTET DALAM MENINGKATKAN BERPIKIR KOMPUTASI SISWA SMA PADA TOPIK KALOR DAN PERPINDAHANNYA MENGGUNAKAN TRANSCRIPT BASED LESSON ANALYSIS

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

Evidence pembelajaran berupa kumpulan data-data yang dianggap penting dalam mendukung data-data penelitian lainnya. *Evidence* ini bisa berupa tulisan dari LKPD, tangkapan layar komputer, serta penilaian proyek CT oleh CodeMaster.

3.4 Prosedur Penelitian

Prosedur penelitian yang dilakukan terdiri dari tiga tahapan utama yaitu tahap persiapan, tahap pelaksanaan, dan tahap analisis data. Ketiga tahapan tersebut diuraikan sebagai berikut.

1. Tahap Persiapan

Pada tahap ini dilakukan kegiatan berupa kajian/studi baik studi lapangan maupun literatur terkait dengan: 1) identifikasi kemampuan yang penting dimiliki siswa di Abad 21 dan di tengah pesatnya perkembangan teknologi serta relevan dengan kurikulum yang digunakan di Indonesia. 2) Melakukan studi lapangan berupa survey kepada beberapa guru fisika /IPA tentang berpikir komputasi dan pelaksanaan STEM di sekolah. 3) Mengidentifikasi permasalahan (*real world problem*) dengan solusi komputasi. 4) Menganalisis kurikulum dan materi yang sesuai dengan permasalahan yang akan disajikan. 5) Menyusun perangkat pembelajaran pembelajaran. 6) Menyusun Instrumen penelitian. 7) Memvalidasi instrumen penelitian pada ahli. 8) Mengujicoba instrumen penelitian. 9) menganalisis hasil uji coba instrumen penelitian. 10) Menentukan sampel penelitian

2. Tahap Pelaksanaan

Setelah instrumen penelitian layak digunakan, maka tahap selanjutnya yaitu melaksanakan penelitian. Pelaksanaan penelitian ini terdiri dari tiga sub tahapan , yaitu :

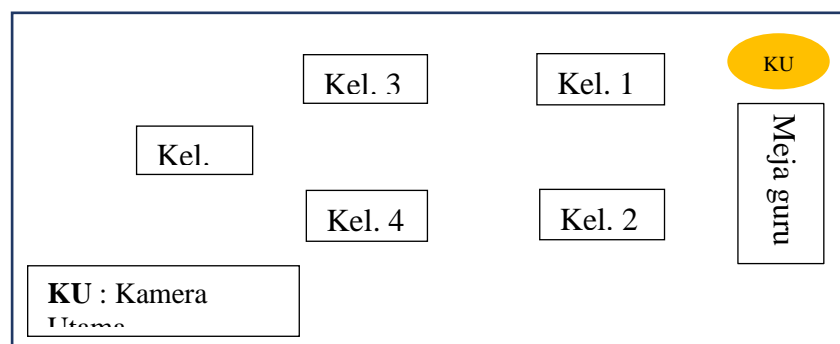
a) Pemberian *Pre-test*

Pre-test diberikan untuk memperoleh informasi awal mengenai kemampuan berpikir komputasi siswa sebelum diberi perlakuan.

a) Intervensi (pembelajaran STEM *Quartet*)

Intervensi atau perlakuan berupa pembelajaran STEM *Quartet* dengan tahapan pembelajaran *understand the problem, research the problem, scope the solution, prototype the solution, dan reflect on the solution*. Pembelajaran ini dilaksanakan sebanyak empat kali pertemuan. Pertemuan pertama dilakukan dengan tahapan *understand the problem* kemudian dilanjutkan *research the*

problem. Pertemuan kedua dan ketiga yaitu pada tahapan *scope the solution* dan pertemuan keempat pada tahapan *prototype the solution* dan *reflect on the solution*. Proyek komputasi yang dibuat yakni proyek simulasi peristiwa konduksi dan membuat aplikasi pengumpul dan penampil informasi data *coolant*. Selama intervensi berlangsung, terdapat satu kamera utama yang berfungsi untuk merekam keseluruhan aktivitas pembelajaran mulai dari awal hingga kelas berakhir dan pandangan kamera ditujukan untuk seluruh kelas. Video pembelajaran dari kamera utama ini menjadi data utama untuk ditranskripsi. Gambaran pengaturan kamera selama kegiatan pembelajaran berlangsung ditunjukkan pada Gambar 3.2.



Gambar 3. 2 Pengaturan Kamera dalam Merekam Aktivitas Pembelajaran

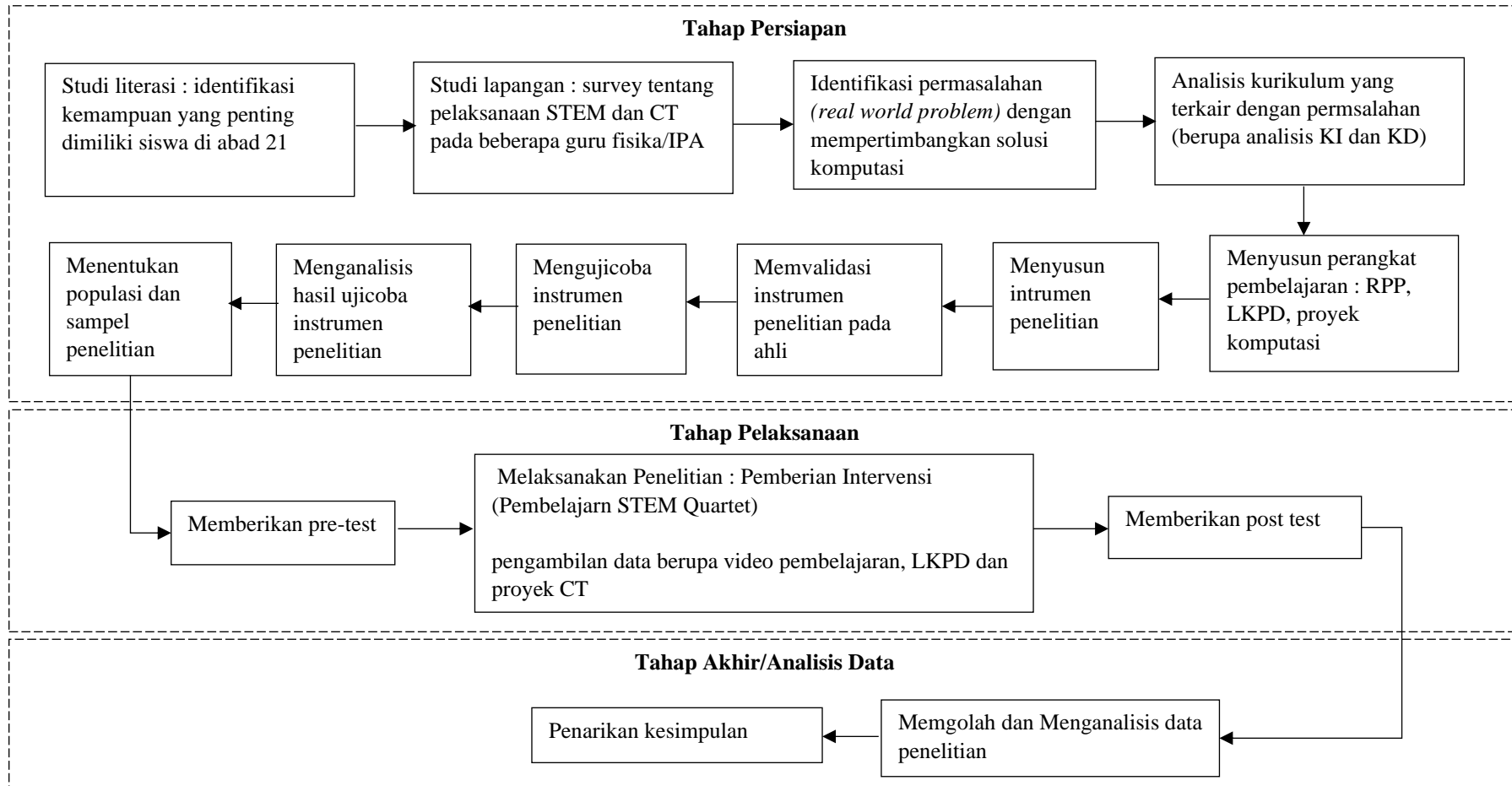
Evidence pembelajaran dikumpulkan dari LKPD serta beberapa hasil tangkapan layar komputer serta penilaian proyek simulasi dan aplikasi dari CodeMaster.

b) Pemberian *Post-test*

Tahap pelaksanaan penelitian terakhir yaitu pemberian *posttest* untuk mendapatkan informasi mengenai peningkatan kemampuan berpikir komputasi setelah diterapkannya pembelajaran dengan model STEM *Quartet*.

3. Tahap Analisis Data

Tahap ini merupakan tahapan dalam mengolah dan menganalisis data-data yang diperoleh selama penelitian untuk diambil kesimpulan. Uraian tentang analisis data dijabarkan pada poin 3.5. Selanjutnya gambaran umum mengenai prosedur penelitian ditunjukkan pada Gambar 3.3



Gambar 3. 3 Prosedur Penelitian

3.5 Analisis Data

Analisis data yang digunakan pada penelitian ini meliputi analisis instrumen dan analisis data penelitian. Analisis data instrumen meliputi uji validasi dan uji reliabilitas instrumen. Sedangkan analisis data penelitian ini meliputi analisis karakteristik pembelajaran STEM *Quartet*, analisis peningkatan kemampuan berpikir komputasi, serta analisis profil kemampuan berpikir komputasi. Selanjutnya teknik analisis kedua data tersebut, diuraikan sebagai berikut.

a. Analisis Instrumen

Suatu data penelitian yang baik ditentukan oleh kualitas sebuah instrumen (Fraenkel et al., 2012). Kualitas instrumen penelitian setidaknya harus memenuhi minimal dua syarat, yaitu instrumen bersifat valid dan reliabel. Instrumen bersifat valid berarti instrumen data mengukur apa yang hendak diukur, sedangkan reliabel berarti instrumen yang jika digunakan beberapa kali akan menghasilkan data yang konsisten (Cresswell, 2014; Fraenkel et al., 2012).

Salah satu pengukuran validitas dan reliabilitas instrumen dapat menggunakan analisis Model Rasch. Model Rasch dikenal sebagai model psikometri (cabang ilmu psikologi yang memiliki konsentrasi pada teknik dan teori pengukuran) untuk menganalisis suatu data kategori sebagai fungsi dari kemampuan/sikap dan tingkat kesulitan pernyataan (Boone & Noltemeyer, 2017). Model Rasch menyediakan berbagai teknik yang dapat digunakan untuk mengevaluasi reliabilitas dan validitas data dari instrumen seperti tes dan survei (You et al., 2018). Analisis Rasch dapat mengidentifikasi respon kesalahan, memprediksi skor data yang hilang, kemampuan tidak hanya bergantung pada jumlah jawaban yang benar, dan dapat mengidentifikasi tebakan. Pola respon yang diberikan menunjukkan ketepatan respon dari setiap responden terhadap setiap item. Keunggulan ini menyebabkan pemodelan Rasch memungkinkan kita untuk menentukan validitas konseptual setiap item. Keunggulan pemodelan Rasch yang dapat memprediksi data yang hilang menjadikan hasil analisis statistik lebih akurat dalam penelitian yang dilakukan (Sumintono & Widhiarso, 2014).

1) Uji Validitas

a) Validitas konten

Validitas konten mengacu pada penilaian atas isi dan struktur logis dari sebuah instrumen yang akan digunakan (Fraenkel et al., 2012) atau dengan kata lain menunjukkan sejauh mana isi tes sesuai dengan tujuan yang dimaksudkannya. Pengujian terhadap kelayakan isi tes dilakukan melalui analisis rasional dari seseorang yang berkompeten (ahli). Instrumen penelitian ini divalidasi oleh 5 orang ahli yang terdiri atas dosen yang merupakan peneliti ahli serta guru fisika SMA. Uji validasi dilakukan pada setiap butir soal dengan rubrik penilaian yang diadaptasi dari kriteria penilaian objektif dari (Terwee et al., 2007). Setiap kriteria diberi penilaian dengan menggunakan skor 1-5 (sangat buruk-sangat baik). Skor pada setiap item dari setiap validator kemudian dianalisis menggunakan uji rater dengan bantuan software Minifac. Hasil uji rater untuk validitas konten ditunjukkan pada Gambar 3.4.

Uji Validasi Isi Soal PG Alasan 06/07/2022 17:23:13
Table 6.0 All Facet Vertical "Rulers".

Vertical = (2A,3A,1A,S) Yardstick (columns lines low high extreme)= 160,4,-2,3,End

Measr	Item	Kriteria	Pakar	Scale
3	+	+	+ Validator 1 + (5)	
:	:	:	: Validator 2 :	
:	:	:	: Validator 5 :	
2	+	+	+ Validator 4 +	
:	:	:	: Validator 3 :	
1	A4	Butir soal tidak mengandung kalimat yang sulit dipahami		
:	A1 A2	:		
:	A9	:		
:	A3 A5	Bacaan memuat sejumlah permasalahan yang harus diselesaikan		
:	:	Penggunaan bahasa sesuai dengan kaidah Bahasa Indonesia yang baik dan benar		
:	A6 A7	Gambar/grafik/tabel yang tercantum menjadi stimulus untuk menginterpretasikan soal dengan tepat		4
0	A8	Soal dibuat sesuai dengan indikator soal		
:	A13	Hanya ada satu jawaban dan alasan yang benar		
:	:	Pilihan dan jawaban yang tersedia logis		
:	:	Soal dibuat sesuai dengan indikator CT yang diukur		3
-1	+	+		
:	A10 A11	Butir soal yang dibuat sesuai dengan konsep/materi		
:	A12	:		
-2	+	+		(2)
:	:	:		
Measr	Item	Kriteria	Pakar	Scale

Metric maintained by + or |. Spacing expanded with : to show all elements.

Gambar 3. 4 Validasi dengan uji rater Minifac

Berdasarkan hasil uji multirater di atas, terdapat 13 item soal (A1-A13) dengan masing-masing memiliki 10 kriteria penilaian yang diuji oleh 5 orang validator. Hasil pada Gambar 3.4 menunjukkan bahwa validator 1-5 berada di atas seluruh kriteria penilaian. Hal tersebut mengartikan bahwa seluruh sebaran soal memiliki interpretasi “valid” menurut ahli untuk semua kriteria validasi yang diujikan.

Hani Sulsilah, 2023

**ANALISIS PENERAPAN PEMBELAJARAN STEM QUARTET DALAM MENINGKATKAN BERPIKIR
KOMPUTASI SISWA SMA PADA TOPIK KALOR DAN PERPINDAHANNYA MENGGUNAKAN
TRANSCRIPT BASED LESSON ANALYSIS**

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

Seluruh kriteria validasi yang diujikan tidak ada yang menjadi perdebatan setiap validator di setiap item soalnya.

b) Validitas konstruk

Validitas konstruk mengacu pada hasil uji coba instrumen kepada siswa. Pada Rasch Model terdapat Unidimensionalitas yang merupakan ukuran yang penting untuk mengevaluasi apakah instrumen tes yang dikembangkan mampu mengukur apa yang seharusnya diukur. Uji validitas konstruk dengan analisis model Rasch dapat menggunakan unidimensionalitas instrumen (*item unidimensionality*) yang dilihat dari nilai *raw variance explained by measures* serta *eigenvalue*.

Adapun interpretasi unidimensionalitas berdasarkan nilai *raw variance explained by measures* ditunjukkan pada Tabel 3, sedangkan *unexplned variance in 1st contrast* harus menunjukkan eigenvalue kurang dari 3 dan nilai observednya kurang dari 15%.

Tabel 3. 3 Interpretasi Unidimensionalitas Instrumen

Interpretasi	Nilai <i>Raw Variance Explained by Measure</i>
Terpenuhi	> 20%
Bagus	> 40%
Istimewa	> 60%

(Sumintono & Widhiarso, 2014)

Ujicoba instrumen tes dilakukan pada 120 siswa SMA yang telah mempelajari materi suhu dan kalor. Hasil ujicoba instrumen kemudian dianalisis menggunakan software Winstep 5.3.2.0 yang ditunjukkan pada Gambar 3.5.

Table of STANDARDIZED RESIDUAL variance in Eigenvalue units = ITEM information units

	Eigenvalue	Observed	Expected
Total raw variance in observations =	19.3415	100.0%	100.0%
Raw variance explained by measures =	6.3415	32.8%	33.0%
Raw variance explained by persons =	4.3218	22.3%	22.5%
Raw Variance explained by items =	2.0197	10.4%	10.5%
Raw unexplained variance (total) =	13.0000	67.2%	100.0%
Unexplned variance in 1st contrast =	1.7970	9.3%	13.8%
Unexplned variance in 2nd contrast =	1.6371	8.5%	12.6%
Unexplned variance in 3rd contrast =	1.4271	7.4%	11.0%
Unexplned variance in 4th contrast =	1.2324	6.4%	9.5%
Unexplned variance in 5th contrast =	1.1173	5.8%	8.6%

Gambar 3. 5 Unidimensionalitas pada Hasil Ujicoba Instrumen Tes *Computational Thinking*

Berdasarkan hasil ujicoba instrumen tes, diperoleh nilai *Raw variance explained by measures* sebesar 32.8% dan *eigenvalue* pada *unexplned variance in*

Hani Sulsilah, 2023

ANALISIS PENERAPAN PEMBELAJARAN STEM QUARTET DALAM MENINGKATKAN BERPIKIR KOMPUTASI SISWA SMA PADA TOPIK KALOR DAN PERPINDAHANNYA MENGGUNAKAN TRANSCRIPT BASED LESSON ANALYSIS

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

1st contrast yaitu sebesar 1.7970%. Interpretasi unidimensionalitas yang ditunjukkan pada Tabel 3.2 menunjukkan bahwa instrumen tes yang dikembangkan dapat dikatakan valid dengan kriteria terpenuhi.

Rasch Model dapat digunakan untuk memeriksa butir soal/item yang tidak sesuai (*outliers/misfit*). Terdapat dua kriteria pengukuran dalam menentukan data fit atau tidak fit, yaitu nilai Infit dan Outfit. Infit (*inlier-sensitive or information-weighted fit*) merupakan kesensitifan pola respon terhadap item sasaran pada *person* atau sebaliknya (Sumintono & Widhiarso, 2014). Sedangkan Outfit (*outlier-sensitive fit*) yaitu mengukur kesensitifan pola respon *person* terhadap item dengan tingkat kesulitan tertentu atau sebaliknya (Linacre, 2002).

Infit dilaporkan sebagai *mean-square* dengan membagi chi-kuadrat dengan jumlah pembobotan. Sedangkan outfit dilaporkan dengan *mean-square* dengan cara chi-kuadrat dibagi dengan derajat kebebasan. Chi-kuadrat lebih familiar untuk ahli statistik dan chi-kuadrat sangat sensitive terhadap *outliers*. Dalam pemodelan Rasch, *outliers* seringkali merupakan tebakan yang beruntung (*lucky guesses*) dan kesalahan yang ceroboh (*careless mistakes*) sehingga karakteristik perilaku *person* yang menyimpang ini dapat membuat item yang baik terlihat buruk. Infit dirancang sebagai statistik yang dapat menurunkan bobot *outliers* dan lebih berfokus pada *string* respon dengan kesulitan item atau abilitas seseorang. Oleh karena itu nilai *outfit* lebih disarankan untuk dilaporkan pada analisis suatu pengukuran karena lebih mudah untuk dianalisis (Linacre, 2002).

Selanjutnya kriteria yang digunakan untuk memeriksa item yang tidak fit ditentukan oleh nilai *outfit* MNSQ, *outfit* ZSTD, dan *Point Measure Correlation*. MNSQ (*mean-square fit statistics*) menunjukkan ukuran keacakan suatu pengukuran yaitu jumlah distorsi suatu sistem pengukuran, nilai ekspektasinya yaitu 1. Nilai MNSQ yang mendekati 1 mengindikasikan adanya distorsi yang sedikit dalam sistem pengukuran. Nilai MNSQ yang lebih besar dari 1 mengindikasikan bahwa observasi tidak mudah diprediksi sedangkan nilai yang lebih kecil dari 1 terlalu dapat diprediksi.

ZSTD merupakan uji t dari hipotesis “apakah data sesuai/fit dengan model?”. ZSTD (Z-standardized) melaporkan signifikansi statistik (probabilitas) dari statistik chi-kuadrat yang terjadi secara kebetulan ketika data sesuai dengan model Rasch.

Nilai ini dilaporkan dalam bentuk penyimpangan nilai z. Nilai model data fit dari ZSTD yaitu 0,0. Nilai yang lebih besar dari 0,0 menunjukkan bahwa data tidak mudah ditebak sedangkan nilai yang lebih kecil dari 0,0 data mudah ditebak.

Pada analisis Rasch, korelasi item digunakan sebagai pemeriksaan langsung untuk mengetahui tingkat penilaian respon masuk akal sehingga jika korelasi bernilai negatif maka kemungkinan ada sesuatu yang salah. *Point Measure Correlation* merupakan korelasi antara pengamatan dalam data dan pengukuran Rasch. *Point Measure Correlation* juga dapat menjadi syarat untuk dapat menentukan item yang fit atau tidak fit. Selanjutnya kriteria penilaian item fit ditentukan oleh Tabel 3.4.

Tabel 3. 4 Kriteria *Item fit Order*

Kriteria	Nilai
<i>Outfit Mean Square (MNSQ)</i>	0,5 <MNSQ <1,5
<i>Outfit Z-Standard (ZSTD)</i>	-2 <ZSTD <+2
<i>Point Measure Correlation (Pt Measure Corr)</i>	0,4 < <i>Pt Measure Corr</i> <0,85

Selanjutnya data pengukuran hasil ujicoba instrumen oleh Winstep 5.3.2.0 ditunjukkan Gambar 3.6

```

INPUT: 120 PERSON 13 ITEM REPORTED: 120 PERSON 13 ITEM 4 CATS WINSTEPS 5.3.2.0
-----
PERSON: REAL SEP.: 1.61 REL.: .72 ... ITEM: REAL SEP.: 4.43 REL.: .95

ITEM STATISTICS: MISFIT ORDER
-----
|ENTRY| TOTAL| TOTAL| JMLE| MODEL| INFIT| OUTFIT| PTMEASUR-AL| EXACT MATCH|
|NUMBER| SCORE| COUNT| MEASURE| S. E. | MNSQ| ZSTD| MNSQ| ZSTD| CORR. | EXP. | OBS%| EXP%| ITEM|
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 1 | 289 | 120 | -.53 | .09 | 1.29 | 2.93 | 1.28 | 2.05 | A .40 | .45 | 16.7 | 28.9 | S01 |
| 4 | 257 | 120 | -.28 | .09 | 1.28 | 2.55 | 1.25 | 1.71 | B .38 | .48 | 24.2 | 30.9 | S04 |
| 5 | 236 | 120 | -.10 | .09 | 1.16 | 1.30 | 1.23 | 1.43 | C .32 | .50 | 25.8 | 35.8 | S05 |
| 3 | 239 | 120 | -.13 | .09 | 1.20 | 1.69 | 1.07 | .52 | D .53 | .50 | 32.5 | 35.7 | S03 |
| 6 | 302 | 120 | -.62 | .09 | 1.19 | 2.09 | 1.11 | .89 | E .49 | .44 | 20.8 | 29.1 | S06 |
| 7 | 236 | 120 | -.10 | .09 | 1.01 | .12 | 1.02 | .20 | F .50 | .50 | 36.7 | 35.8 | S07 |
| 13 | 207 | 120 | .19 | .11 | 1.02 | .15 | 1.01 | .14 | G .55 | .52 | 38.3 | 43.1 | S013 |
| 9 | 208 | 120 | .18 | .11 | .96 | -.22 | .94 | -.28 | f .49 | .52 | 36.7 | 43.0 | S09 |
| 2 | 148 | 120 | 1.31 | .19 | .74 | -.94 | .84 | -.35 | e .52 | .53 | 82.5 | 83.2 | S02 |
| 8 | 167 | 120 | .79 | .14 | .79 | -.96 | .70 | -1.15 | d .64 | .55 | 74.2 | 75.0 | S08 |
| 11 | 245 | 120 | -.18 | .09 | .76 | -2.37 | .73 | -2.01 | c .60 | .49 | 36.7 | 35.5 | S011 |
| 12 | 260 | 120 | -.30 | .09 | .68 | -3.56 | .72 | -2.21 | b .51 | .48 | 40.8 | 30.9 | S012 |
| 10 | 249 | 120 | -.21 | .09 | .60 | -4.27 | .60 | -3.19 | a .64 | .49 | 47.5 | 30.5 | S010 |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| MEAN | 234.1 | 120.0 | .00 | .10 | .98 | -.11 | .96 | -.17 | | | 39.5 | 41.3 |
| P. SD | 41.6 | .0 | .51 | .03 | .23 | 2.19 | .22 | 1.53 | | | 18.6 | 16.8 |
-----

```

Gambar 3. 6 Item Misfit Order Statistics

Gambar 3.6 menunjukkan beberapa informasi data yaitu *measure*, *SE* (*standard error*), *Infit*, *Outfit*, serta *Pt Measure*. Dalam hal ini yang akan dibahas yaitu item yang misfit sehingga penilaian berfokus pada *Outfit* dan *Pt Measure*

Hani Sulsilah, 2023

ANALISIS PENERAPAN PEMBELAJARAN STEM QUARTET DALAM MENINGKATKAN BERPIKIR
KOMPUTASI SISWA SMA PADA TOPIK KALOR DAN PERPINDAHANNYA MENGGUNAKAN
TRANSCRIPT BASED LESSON ANALYSIS

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

Corr. Linacre menyarankan dalam suatu pelaporan item *misfit*, hal yang pertama dilihat adalah nilai *Pt Measure Corr*. Nilai *Pt Measure Corr* yang negatif dinilai bahwa item tidak dapat mengukur apa yang hendak diukur. Berdasarkan data yang ditunjukkan Gambar 3.6 nilai *Pt Measure Corr* yang diperoleh yaitu antara 0,32 sampai 0,64. Hal ini menunjukkan bahwa tidak terdapat nilai *Pt Measure Corr* yang negatif sehingga dapat dikatakan bahwa seluruh item dapat mengukur apa yang hendak diukur. Namun terdapat dua item yang tidak memenuhi rentang pengukuran *Pt Measure Corr* yaitu berada pada nilai dibawah 0,4 untuk item nomor 4 (SO4) dan item nomor 5 (SO5). Oleh karena itu investigasi selanjutnya akan dilihat dari 2 kriteria lain seperti yang ditunjukkan Tabel 3.4.

Setelah meninjau nilai *Pt Measure Corr* tidak ada yang bernilai negatif, investigasi, kemudian dilakukan peninjauan atas nilai Outfit daripada Infit. Nilai Outfit Mean-Square (MNSQ) yang diperoleh dari hasil ujicoba instrumen berkisar antara 0,6 – 1,29. Berdasarkan hasil tersebut dinyatakan bahwa seluruh item memiliki nilai Outfit MNSQ pada rentang pengukuran yang produktif (0,5-1,5).

Nilai ZSTD yang diperoleh memiliki rentang nilai -3,19 sampai dengan 2,05. Nilai ZSTD yang lebih kecil dari -2 menyatakan bahwa item (item nomor 10, 11, 12) terlalu mudah diprediksi. Sedangkan nilai ZSTD yang lebih dari +2 menyatakan bahwa item sulit diprediksi.

Gambar 3.6 menunjukkan bahwa soal nomor 4,5,10,11,12 (SO4,SO5,SO10,SO11,SO12) hanya memiliki dua kriteria *item fit order* dimana soal nomor 4 dan 5 (SO4 dan SO5) memiliki *Pt Measure Corr* kurang dari 0,4 dan soal nomor 10,11, 12 memiliki nilai ZSTD tidak berada pada rentang -2 sampai dengan +2. Meski demikian, suatu item dikatakan “sesuai” jika setidaknya terdapat dua dari tiga nilai-nilai kriteria item fit order seperti pada Tabel 3.4 (Sumintono & Widhiarso, 2014). Oleh karena itu, seluruh item dapat dikatakan fit / sesuai yang berarti instrumen soal dapat mengukur apa yang hendak diukur.

2) Reliabilitas

Terdapat tiga kriteria statistik dalam menentukan reliabilitas instrumen dengan menggunakan Rasch Model yaitu Cronbach's alpha, *Item and Person*

Hani Sulsilah, 2023

**ANALISIS PENERAPAN PEMBELAJARAN STEM QUARTET DALAM MENINGKATKAN BERPIKIR
KOMPUTASI SISWA SMA PADA TOPIK KALOR DAN PERPINDAHANNYA MENGGUNAKAN
TRANSCRIPT BASED LESSON ANALYSIS**

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

Reliability serta *Item and Person Separation*. Kriteria penilaian tersebut ditunjukkan pada Tabel 3.5

Tabel 3. 5 Reliabilitas dalam Analisis Rasch

Statistics	Fit Indices	Interpretation
Cronbach's alpha (KR-20)	<0.5	Low
	0.5 – 0.6	Moderate
	0.6 – 0.7	Good
	0.7 – 0.8	High
	>0.8	Very High
Item and Person Reliability	<0.67	Low
	0.67 – 0.8	Sufficient
	0.81 – 0.90	Good
	0.91 – 0.94	Very Good
	>0.94	Excellent
Item and Person Separation		A high separation value indicates that the instrument has good quality since it can identify the group and respondent

(Krishnan & Idris, 2014)

SUMMARY OF 120 MEASURED PERSON

	TOTAL		MEASURE	MODEL S. E.	INFIT		OUTFIT	
	SCORE	COUNT			MNSQ	ZSTD	MNSQ	ZSTD
MEAN	25.4	13.0	-.56	.32	.99	.10	.96	.08
SEM	.6	.0	.06	.01	.02	.06	.02	.05
P. SD	7.0	.0	.64	.09	.23	.63	.27	.51
S. SD	7.0	.0	.65	.09	.23	.64	.27	.52
MAX.	50.0	13.0	2.09	.94	1.53	1.73	1.65	1.27
MIN.	14.0	13.0	-2.58	.26	.51	-1.47	.28	-1.30
REAL RMSE	.34	TRUE SD	.55	SEPARATION	1.61	PERSON RELIABILITY	.72	
MODEL RMSE	.33	TRUE SD	.55	SEPARATION	1.69	PERSON RELIABILITY	.74	
S. E. OF PERSON MEAN	= .06							

PERSON RAW SCORE-TO-MEASURE CORRELATION = .98 (approximate due to missing data)
 CRONBACH ALPHA (KR-20) PERSON RAW SCORE "TEST" RELIABILITY = .75 SEM = 3.52 (a)

SUMMARY OF 13 MEASURED ITEM

	TOTAL		MEASURE	MODEL S. E.	INFIT		OUTFIT	
	SCORE	COUNT			MNSQ	ZSTD	MNSQ	ZSTD
MEAN	234.1	120.0	.00	.10	.98	-.11	.96	-.17
SEM	12.0	.0	.15	.01	.07	.63	.06	.44
P. SD	41.6	.0	.51	.03	.23	2.19	.22	1.53
S. SD	43.3	.0	.53	.03	.24	2.28	.23	1.60
MAX.	302.0	120.0	1.31	.19	1.29	2.93	1.28	2.05
MIN.	148.0	120.0	-.62	.09	.60	-4.27	.60	-3.19
REAL RMSE	.11	TRUE SD	.50	SEPARATION	4.43	ITEM RELIABILITY	.95	
MODEL RMSE	.11	TRUE SD	.50	SEPARATION	4.57	ITEM RELIABILITY	.95	
S. E. OF ITEM MEAN	= .15							

Gambar 3. 7 Reliabilitas *Person* dan *Item*

Berdasarkan hasil uji coba instrumen yang dilakukan pada 120 siswa SMA yang telah mempelajari materi Suhu dan Kalor diperoleh nilai reliabilitas seperti yang ditunjukkan pada Gambar 3.7. Reliabilitas person dan item diperoleh masing-masing sebesar 0,72 dan 0,95 yang berarti konsistensi jawaban siswa dalam

Hani Sulsilah, 2023

ANALISIS PENERAPAN PEMBELAJARAN STEM QUARTET DALAM MENINGKATKAN BERPIKIR KOMPUTASI SISWA SMA PADA TOPIK KALOR DAN PERPINDAHANNYA MENGGUNAKAN TRANSCRIPT BASED LESSON ANALYSIS

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

kategori cukup untuk menjawab item dengan reliabilitas yang sangat tinggi. Nilai Cronbach's Alpha sebesar 0,75 yang menunjukkan interaksi antara *person* dan item secara keseluruhan adalah tinggi. Nilai item dan *person separation* yang diperoleh dari hasil ujicoba seperti yang ditunjukkan pada Gambar 3.7 yaitu sebesar 4.43 dan 1.61. (Sumintono & Widhiarso, 2014) mengungkapkan bahwa semakin besar nilai separasi maka kualitas instrumen dalam keseluruhan dapat semakin bagus karena dapat mengidentifikasi kelompok responden dan kelompok item.

3) Tingkat Kesulitan Item

Pengelompokkan item dapat dikategori menjadi item sulit, sedang dan mudah. Tingkat kesulitan ini diperoleh dari perbandingan dari nilai *measure* dan standar deviasi pada *item fit statistic*. Selanjutnya interpretasi tingkat kesulitan item ditunjukkan oleh Tabel 3.6

Tabel 3. 6 Interpretasi Tingkat Kesulitan Item

Interpretasi	Nilai
Sulit	$M > +1 SD$
Sedang	$1SD \geq M \geq -1SD$
Mudah	$M < -1SD$

(Sumintono dan Widhiarso, 2014)

Adapun hasil dari pengolahan data untuk kualitas item dan tingkat kesulitan soal ditunjukkan pada Tabel 3.7 .

Tabel 3. 7 Tingkat Kesulitan Soal Hasil Ujicoba

Item	Measure	SD	Tingkat kesulitan soal
SO1	-0.53		Mudah
SO2	1.31		Sulit
SO3	-0.13		Sedang
SO4	-0.28		Sedang
SO5	0.10		Sedang
SO6	-0.62		Mudah
SO7	-0.10	0.51	Sedang
SO8	0.79		Sulit
SO9	0.18		Sedang
SO10	-0.21		Sedang
SO11	-0.18		Sedang
SO12	-0.30		Sedang
SO13	0.19		Sedang

4) Diskriminasi Item

Rasch model dapat menampilkan indeks diskriminasi atau daya pembeda suatu item. Diskriminasi item menunjukkan sejauh mana keberhasilan pada suatu item/butir soal sesuai dengan keberhasilan pada keseluruhan tes. Dalam pemodelan Rasch indeks diskriminasi diukur dari skor tertinggi dan terendah dari grup yang sama (Kelley et al., 2002). Nilai estimasi indeks diskriminasi dalam pemodelan Rasch yaitu bernilai 1 agar sesuai dengan model. Kisaran indeks diskriminasi bernilai antara -1 dan +1, nilai 0,4 ke atas dianggap tinggi dan kurang dari 0,2 dianggap rendah (Kelley et al., 2002).

Berdasarkan hasil ujicoba instrumen, diperoleh informasi nilai indeks diskriminasi tiap item seperti yang ditunjukkan oleh Tabel 3.8

Tabel 3. 8 Nilai Indeks Diskriminasi

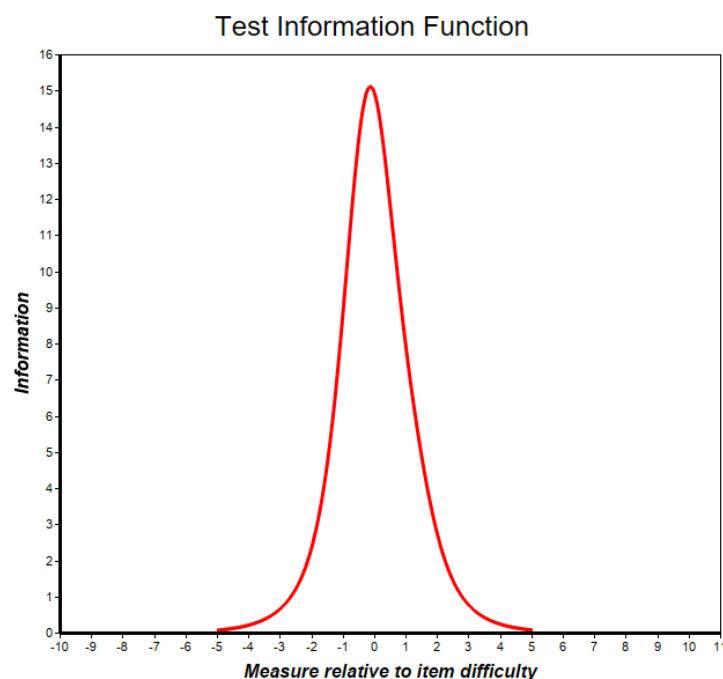
Item	Index Discrimination
SO1	1,10
SO2	0,90
SO3	1,27
SO4	0,87
SO5	0,61
SO6	1,61
SO7	1,00
SO8	1,03
SO9	0,90
SO10	1,12
SO11	1,13
SO12	0,71
SO13	1,06

Secara umum, nilai indeks diskriminasi memiliki nilai terkecil sebesar 0,61 dan nilai terbesar 1,61. Berdasarkan (Kelley et al., 2002) dapat dinyatakan bahwa seluruh item memiliki nilai indeks diskriminasi yang tinggi karena memiliki nilai yang lebih dari 0,4. Terdapat beberapa nilai indeks diskriminasi yang lebih dari 1 yaitu pada item nomor 1,3,6,8,10,11, dan 13. Nilai indeks diskriminasi yang lebih besar dari 1 mengartikan bahwa item tersebut membedakan antara yang berkinerja tinggi dan rendah lebih dari yang diharapkan untuk tingkat kesulitan pada item tersebut. Karena suatu item dalam tes dimaksudkan untuk menghasilkan skor keseluruhan tes, setiap item dengan diskriminasi negatif atau nol dapat merusak tes (Linacre,2002). Berdasarkan nilai indeks diskriminasi yang diperoleh, tidak

ditemukan adanya nilai indeks diskriminasi yang negatif atau nol sehingga seluruh item dapat dinyatakan layak digunakan sebagai instrumen tes.

5) Fungsi Informasi Tes

Informasi tes didasarkan pada kompetensi dan kemampuan siswa. Informasi tes dapat dihitung dengan menjumlahkan semua informasi kompetensi. Hasil uji coba instrumen memberikan hasil bahwa jumlah informasi memiliki nilai maksimum pada tingkat kemampuan sekitar 0 pada nilai informasi 15. Dengan kata lain, model kompetensi paling informatif ketika kemampuan siswa sama dengan tingkat kesulitan kompetensi dan kurang informatif ketika kemampuan siswa menjauh dari tingkat kesulitan kompetensi (yaitu ketika kompetensi terlalu mudah atau terlalu sulit untuk siswa)



Gambar 3. 8 Fungsi Informasi Tes

b. Analisis Data Penelitian

1) Karakteristik Pembelajaran STEM *Quartet*

Transcript Based Lesson Analysis (TBLA) merupakan suatu metode analisis berbasis transkrip yang dikembangkan dalam *Lesson Study*. TBLA dapat digunakan untuk menganalisis karakteristik dari suatu pembelajaran (Arani, 2017). Pada

Hani Sulsilah, 2023

ANALISIS PENERAPAN PEMBELAJARAN STEM QUARTET DALAM MENINGKATKAN BERPIKIR KOMPUTASI SISWA SMA PADA TOPIK KALOR DAN PERPINDAHANNYA MENGGUNAKAN TRANSCRIPT BASED LESSON ANALYSIS

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

penelitian ini, untuk memperoleh bagaimana karakteristik pembelajaran STEM *Quartet* dalam melatih CT maka transkripsi video yang difokuskan yaitu pada tahap observasi (selama pembelajaran berlangsung). Untuk memahami karakteristik pembelajaran, maka pembelajaran diklasifikasikan berdasarkan tahapan pembelajaran STEM *Quartet*. Tahapan analisis transkrip pembelajaran dilakukan dengan cara sebagai berikut:

- a) Membuat transkrip dari video pembelajaran dan merepresentasikannya ke dalam bentuk tabel

Tabel 3. 9 Transkripsi Video

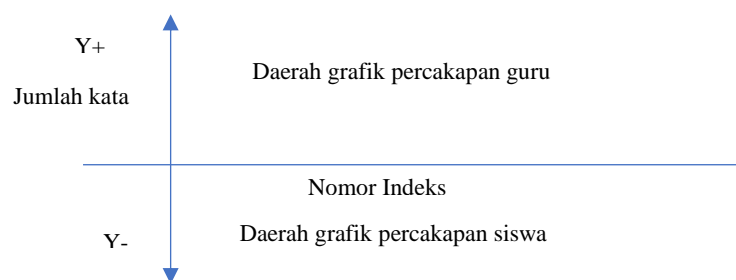
Indeks	Pembicara	Kalimat pembicaraan
1	S :Siswa	Kalimat/ucapan dalam
2	G: Guru	pembelajaran

- b) Mengklasifikasi transkripsi sesuai dengan tahapan pembelajaran STEM *Quartet*

Tabel 3. 10 Klasifikasi Transkrip

Indeks	Pembicara	Kalimat pembicaraan	Tahapan STEM <i>Quartet</i>
1	S :Siswa	Kalimat/ucapan	Tahap I
2	G: Guru	dalam pembelajaran	
dst			dst

- c) Membuat grafik jumlah kata yang diucapkan guru dan siswa



Gambar 2. 1 Representasi Grafik Jumlah kata Percakapan Guru dan Siswa

- d) Menentukan kata kunci yang berpotensi muncul
 e) Membuat grafik kata kunci
 f) Menganalisis bagaimana pembelajaran STEM *Quartet* berdasarkan klasifikasi dari tahapan-tahapan pembelajaran STEM *Quartet*

2) Peningkatan Kemampuan Berpikir Komputasi

Hani Sulsilah, 2023

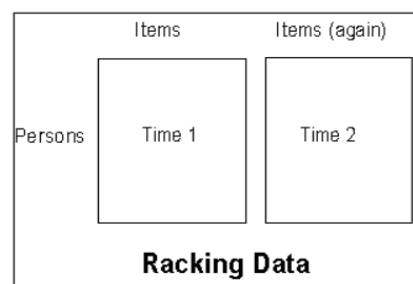
ANALISIS PENERAPAN PEMBELAJARAN STEM QUARTET DALAM MENINGKATKAN BERPIKIR KOMPUTASI SISWA SMA PADA TOPIK KALOR DAN PERPINDAHANNYA MENGGUNAKAN TRANSCRIPT BASED LESSON ANALYSIS

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

Pada penelitian ini, peningkatan kemampuan berpikir komputasi dianalisis menggunakan Rasch Model. Penggunaan Rasch Model dalam mengukur perubahan sebelum dan setelah intervensi memiliki keunggulan dapat mengungkapkan secara detail mengenai perubahan *person* dan item secara individual (Laliyo, 2021; Ling et al., 2018).

a. *Racking Analysis*

Metode *racking* digunakan untuk menganalisis perubahan tingkat kesulitan item sebelum dan sesudah intervensi. Metode ini menempatkan data *pre-test* dan *post-test* secara horizontal. Setiap item akan muncul dua kali dalam data tersebut sehingga terlihat apakah ada tingkat kesulitan item mengalami perubahan atau tidak. Penempatan data item dan *person* ditunjukkan oleh Gambar 3.9



Gambar 3. 9 *Racking Data*

Tingkat kesulitan ini kemudian akan diuraikan berdasarkan indikator-indikator CT yang terdapat dalam soal tersebut sehingga diperoleh informasi mengenai aspek-aspek CT mana saja yang mengalami perubahan sebelum dan sesudah intervensi.

Tabel 3. 11 *Pre-post item difficulties*

Item	Pre-test	Post test	Difference	Sign
SO _{xx}	X	Y	Y-X	+ / 0 /-

Tabel 3.11 menunjukkan nilai tingkat kesukaran setiap item. Item dengan kode SO_{xx} merupakan item dengan urutan nomor yang diujikan. X merupakan nilai logit untuk tingkat kesukaran item pada saat *pretest*. Y merupakan nilai logit untuk tingkat kesukaran item pada saat *posttest*. *Difference* merupakan selisih antara tingkat kesukaran item pada saat *posttest* dan *pretest*. *Sign* dapat berupa tanda + atau - atau nilai nya 0 hasil dari Y-X. tanda negative menunjukkan bahwa item menjadi berkurang tingkat kesukarannya, tanda + berarti item bertambah tingkat kesukarannya/ makin sukar dan 0 menunjukkan item tidak berubah tingkat

Hani Sulsilah, 2023

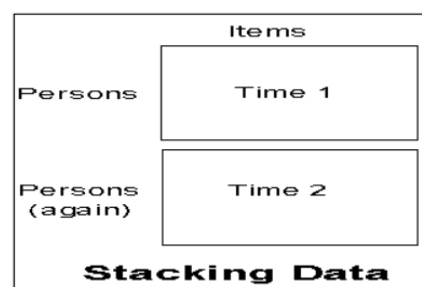
ANALISIS PENERAPAN PEMBELAJARAN STEM QUARTET DALAM MENINGKATKAN BERPIKIR KOMPUTASI SISWA SMA PADA TOPIK KALOR DAN PERPINDAHANNYA MENGGUNAKAN TRANSCRIPT BASED LESSON ANALYSIS

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

kesukarannya. Semakin besar nilai *difference* maka semakin besar pula perubahan tingkat kesukarannya (Wright, 2003).

b. *Stacking Analysis*

Metode digunakan untuk menganalisis perubahan individu setelah intervensi diberikan. *stacking* Metode *stacking* pada analisis Rasch menempatkan data *pre-posttest* secara vertikal (Laliyo, 2021; Wright, 2003). Setiap *person* (dalam hal ini siswa) akan muncul dua kali dalam kumpulan data pada item yang sama seperti yang ditunjukkan oleh Gambar 3.9. Hal ini akan memungkinkan peneliti untuk dapat menganalisis perubahan per individu setelah intervensi dilaksanakan. Analisis perubahan per individu dapat memberikan informasi pada peneliti untuk mengetahui bahwa suatu intervensi mungkin tidak memberikan hasil yang sama pada setiap individu.



Gambar 3. 10 *Stacking Data*

Stacking berfokus pada “siapa yang berubah” setelah intervensi diberikan (Wright, 2003) sehingga memuat informasi perubahan pada level individu. Perubahan ini ditunjukkan oleh nilai logit dari *person measure* pada saat *pretest* dan *posttest*. Perubahan nilai *person measure* ditunjukkan pada Tabel 3.12.

Tabel 3. 12 Perubahan *Person Measure*

<i>Person</i>	Pre-test	Post test	Difference	Sign
Pxx	X	Y	Y-X	+ / 0 /-
Lxx				

Person bertanda Pxx dan Lxx masing-masing merupakan siswa perempuan dan Laki-laki dengan xx adalah penomoran urut. Sementara itu simbol X merupakan nilai logit dari *person measure* saat *pretest* dan Y merupakan nilai logit *person measure* pada saat *posttest*. *Difference* merupakan selisih dari *person measure*. Tanda positif menunjukkan adanya peningkatan abilitas siswa dalam berpikir komputasi. Sedangkan tanda negatif dari selisih *person measure*

Hani Sulsilah, 2023

ANALISIS PENERAPAN PEMBELAJARAN STEM QUARTET DALAM MENINGKATKAN BERPIKIR
KOMPUTASI SISWA SMA PADA TOPIK KALOR DAN PERPINDAHANNYA MENGGUNAKAN
TRANSCRIPT BASED LESSON ANALYSIS

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

menandakan adanya penurunan abilitas siswa. sementara itu nilai nol pada *difference* dan *sign* menunjukkan bahwa siswa tidak mengalami perubahan abilitas baik setelah intervensi (penerapan pembelajaran STEM *Quartet*) diberikan. Semakin besar nilai perubahan *person measure* maka semakin besar pula efek suatu intervensi terhadap individu tersebut. (Wright, 2003)

Kategori Level Kemampuan Berpikir Komputasi Siswa

Level kemampuan berpikir komputasi yang dimaksud adalah level pengelompokkan berdasarkan standar pemodelan Rasch. Level kemampuan berpikir komputasi ditujukan untuk memperlihatkan pengelompokkan kemampuan siswa secara keseluruhan. Aturan level kemampuan ini berdasarkan analisis Rasch dengan cara mengukur median dari nilai 50% nilai probabilitas setiap level (Boone et al., 2014). Adapun langkah-langkah untuk menentukan nilai batas setiap level yaitu :

1. Menjalankan *Software Winstep*
2. Memilih “Item Structure file ISFILE=” dari menu file Outputs
3. Menemukan kolom kategori CAT dan 50% PROB pada Item-Structure File.
4. Menghitung nilai median dari nilai 50% PROB pada setiap level/kategori
5. Menggunakan nilai median sebagai nilai batas antar level/kategori
6. Mengulangi Langkah 3-5 untuk menentukan nilai batas untuk seluruh level/kategori. .

(Boone et al., 2014)

3) Profil Keterampilan Berpikir Komputasi

Profil keterampilan berpikir komputasi selanjutnya ditunjukkan dari penilaian proyek Komputasi. Terdapat dua proyek komputasi yang dibuat siswa yaitu proyek simulasi peristiwa konduksi serta proyek aplikasi pengumpul dan penampil informasi *database coolant*. Proyek penilaian komputasi dinilai menggunakan penilaian otomatis berbasis web yaitu CodeMaster. Beberapa penilaian Code Master didasarkan pada rubrik penilaian yang ditunjukkan pada Tabel 3.13

Tabel 3. 13 Rubrik Penilaian Proyek Siswa oleh Code Master

<i>Criteria</i>	<i>Level of Performance</i>			
	0	1	2	3
<i>Screens</i>	Tampilan layar tunggal	tampilan layar tunggal	Tampilan terdiri dari 2	Tampilan terdiri dari 2

Hani Sulsilah, 2023

ANALISIS PENERAPAN PEMBELAJARAN STEM QUARTET DALAM MENINGKATKAN BERPIKIR KOMPUTASI SISWA SMA PADA TOPIK KALOR DAN PERPINDAHANNYA MENGGUNAKAN TRANSCRIPT BASED LESSON ANALYSIS

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

Criteria	Level of Performance			
	0	1	2	3
	dengan komponen visual dan tidak terprogram untuk mengubah keadaan	dengan komponen visual yang mengubah status secara terprogram	layar dengan komponen visual yang berfungsi mengubah status secara terprogram	atau lebih layar dengan komponen visual yang berfungsi mengubah status secara terprogram
<i>User Interface</i>	Menggunakan satu komponen visual tanpa adanya perubahan	Menggunakan dua atau lebih komponen visual tanpa adanya perubahan	Menggunakan 5 atau lebih komponen visual dengan satu tipe perubahan	Menggunakan 5 atau lebih komponen visual dengan beberapa tipe perubahan
<i>Naming, components, variables, procedures</i>	Sedikit atau penamaan tidak berubah dari penamaan standar bawaan	10-25% penamaan berubah dari penamaan standar bawaan	26-75% penamaan berubah dari penamaan standar bawaan	Lebih dari 75% penamaan berubah dari penamaan standar bawaan
<i>Events</i>	Tidak menggunakan tipe <i>events</i> apapun	Menggunakan satu jenis <i>event</i>	Menggunakan 2 jenis <i>event</i>	Menggunakan lebih dari 2 jenis <i>event</i>
<i>Procedural abstraction</i>	Tidak menggunakan prosedur	Hanya ada satu penggunaan prosedur	Terdapat lebih dari satu prosedur	Terdapat beberapa prosedur dan adanya penggunaan prosedur ulang
<i>Loops</i>	Tidak ada penggunaan <i>loop</i>	Menggunakan <i>loops</i> sederhana (<i>while</i>)	Menggunakan " <i>for each</i> " <i>loops</i> dengan <i>variable</i> sederhana	Menggunakan " <i>for each</i> " <i>loops</i> dengan <i>list item</i>
<i>Data persistence</i>	Data hanya disimpan sebagai <i>variable</i> atau komponen properti UI dan tidak bertahan saat	Data disimpan dalam file	Menggunakan <i>database local</i> (TinyDB)	Menggunakan <i>web database</i>

<i>Criteria</i>	<i>Level of Performance</i>			
	0	1	2	3
	aplikasi ditutup			
<i>Sensors</i>	Tidak menggunakan sensor	Menggunakan satu jenis sensor	Menggunakan dua jenis sensor	Menggunakan lebih dari dua jenis sensor
<i>Media</i>	Tidak menggunakan komponen media	Menggunakan satu jenis komponen media	Menggunakan dua jenis komponen media	Menggunakan lebih dari dua jenis komponen media
<i>Connectivity</i>	Tidak menggunakan komponen <i>connectivity</i>	Menggunakan satu komponen <i>connectivity</i>	Menggunakan dua komponen <i>connectivity</i>	Menggunakan lebih dari dua komponen <i>connectivity</i>
<i>Drawing and Animation</i>	Tidak menggunakan komponen <i>drawing and animation</i>	Menggunakan komponen kanvas	Menggunakan komponen bola (ball)	Menggunakan komponen <i>image sprite</i>

CodeMaster memberikan kebebasan untuk menetapkan aspek-aspek mana saja yang sesuai dengan aplikasi yang dibuat, karena setiap aplikasi memiliki aspek yang berbeda untuk dikembangkan (Wangenheim et al., 2018). Selanjutnya CodeMaster membuat klasifikasi profil penilaian dalam bentuk *Numerical Grade*. *Numerical grade* merupakan nilai dalam bentuk ordinal dari skala 1-10. Konversi skor mentah dari CodeMaster ke *numerical grade* (level numerik) dilakukan melalui perhitungan sebagai berikut

$$\text{Numerical Grade} = \frac{\text{score}}{\text{maximum score of assesed criteria}} * 10$$

(Wangenheim et al., 2018)