

BAB III

METODE PENELITIAN

Metode penelitian yang digunakan adalah metode penelitian *developmental research*. Penelitian yang secara sistematis merancang, mengembangkan dan mengevaluasi program instruksional, proses, dan produk yang harus memenuhi kriteria konsistensi dan efektivitas internal. Bab ini terdiri dari desain penelitian menjelaskan bagaimana prosedur penelitian yang dilakukan. Berikutnya dijelaskan lokasi dan partisipan penelitian, kemudian bagian akhir bab ini menjelaskan bagaimana menguji kualitas instrumen yang dikembangkan. Uji kualitas instrumen meliputi uji validitas, ujicoba awal untuk mendapatkan informasi tentang kepraktisan instrumen serta ujicoba utama untuk menguji reliabilitas (konsistensi) instrumen (Richey & Klein, 2014; Tracey, 2009; Richey, Klein, & Nelson, 2004).

3.1 Desain Penelitian

Desain pengembangan yang digunakan pada penelitian ini adalah *Design and Development Research* yang memiliki fokus pengembangan dengan orientasi desain spesifik konten yang merupakan ciri penelitian *Design and Development Research* tipe 1. Penelitian *Design and Development Research* tipe 1 fokus memeriksa proses perencanaan, pengembangan, dan evaluasi yang dibahas, metode penelitian yang digunakan dan sifat kesimpulan. Kategori dari *Design and Development Research* tipe 1 biasanya digunakan mengikuti prinsip *instructional systems design* (ISD) yang mencakup analisis *front-end* melalui evaluasi. Desain pengembangan tipe 1 berfokus pada pengajian kebutuhan dan fase desain aktual dibandingkan keseluruhan proses desain-pengembangan-evaluasi. Tahap implementasi ISD pada *Design and Development Research* tipe 1 bersifat iteratif, sehingga fase desain penelitian *Design and Development Research* tipe 1 terdiri *design*, *development*, dan *evaluation* (Richey & Klein, 2014; Richey, Klein, & Nelson, 2004).

3.1.1. Perencanaan (*Design*)

Pengembangan instrumen penilaian dimulai dengan penentuan pola instrumen penelitian. Instrumen awal berisikan definisi operasional *framework* TPACK, indikator serta *task* dalam merancang rencana pembelajaran. *Framework*

TPACK berisikan 7 domain yaitu CK, TK, PK, PCK, TC, TPK, dan TPACK yang berisikan kode 1,2,, dan 3 seperti CK1, CK2, CK3... hingga TPACK1, TPACK2, dan TPACK3.

3.1.2. Pengembangan (*Development*)

Draf awal instrumen tersebut berupa definisi TPACK, *task* and rubrik. Pada kisi-kisi instrumen penilaian kinerja ditentukan definisi pada 7 domain *Content Knowledge* (CK), *Pedagogical Knowledge* (PK), *Knowledge Technology* (TK), *Pedagogical Content Knowledge* (PCK), *Technology Content Knowledge* (TCK), *Pedagogical Technology Knowledge* (TPK), dan *Pedagogical Technology and Content Knowledge* (TPACK). Indikator penilaian kinerja ditentukan pada setiap domain TPACK. Instrumen *self assessment* di lengkapi dengan item pernyataan di setiap masing-masing domain TPACK. Instrumen penilaian observasi dilengkapi dengan *task* dan rubrik penskoran. Instrumen *self assessment* dan observasi masing-masing dilengkapi dengan lembar validasi untuk dinilai oleh pakar.

3.1.3. Evaluasi (*Evaluation*)

Draft instrumen penilaian kinerja TPACK yang telah dirancang kemudian divalidasi dengan meminta lima orang pakar yaitu pakar asesmen dan pembelajaran kimia. Setelah dilakukan telaah dan *judgement* oleh para pakar, instrumen direvisi dan diperbaiki sesuai dengan saran para pakar. Tahap uji coba dilakukan dengan memberikan *task* kepada mahasiswa calon pendidik dalam membuat perencanaan pembelajaran. Selanjutnya calon pendidik diminta mencoba melengkapi *task* sesuai dengan arahan yang diberikan.

Tahap uji kualitas rubrik dilakukan setelah rubrik selesai disusun. Uji kualitas rubrik meliputi uji validitas dan uji reliabilitas (konsistensi). Uji validitas dilakukan dengan meminta pertimbangan para pakar untuk memvalidasi rubrik, revisi rubrik dilakukan sesuai dengan saran-saran yang diberikan para pakar, sampai akhirnya para pakar menilai rubrik tersebut baik. Uji reliabilitas dilakukan dengan uji konsistensi penilaian antar penilai, menggunakan uji statistik *ICC* dan *W Kendall*. Revisi rubrik kembali dilakukan pada saat hasil uji konsistensi penilaian antar penilai belum memenuhi syarat minimum, sama dengan interpretasi koefisien korelasi yaitu kategori sedang. Selanjutnya pengujian konsistensi penilaian antar penilai kembali dilakukan setelah diadakan perbaikan pada rubrik, sampai diperoleh indeks konsistensi penilaian yang memadai. Tahap

akhir pengembangan instrumen menyimpulkan kualitas instrumen berdasarkan uji kualitas yang dilakukan.

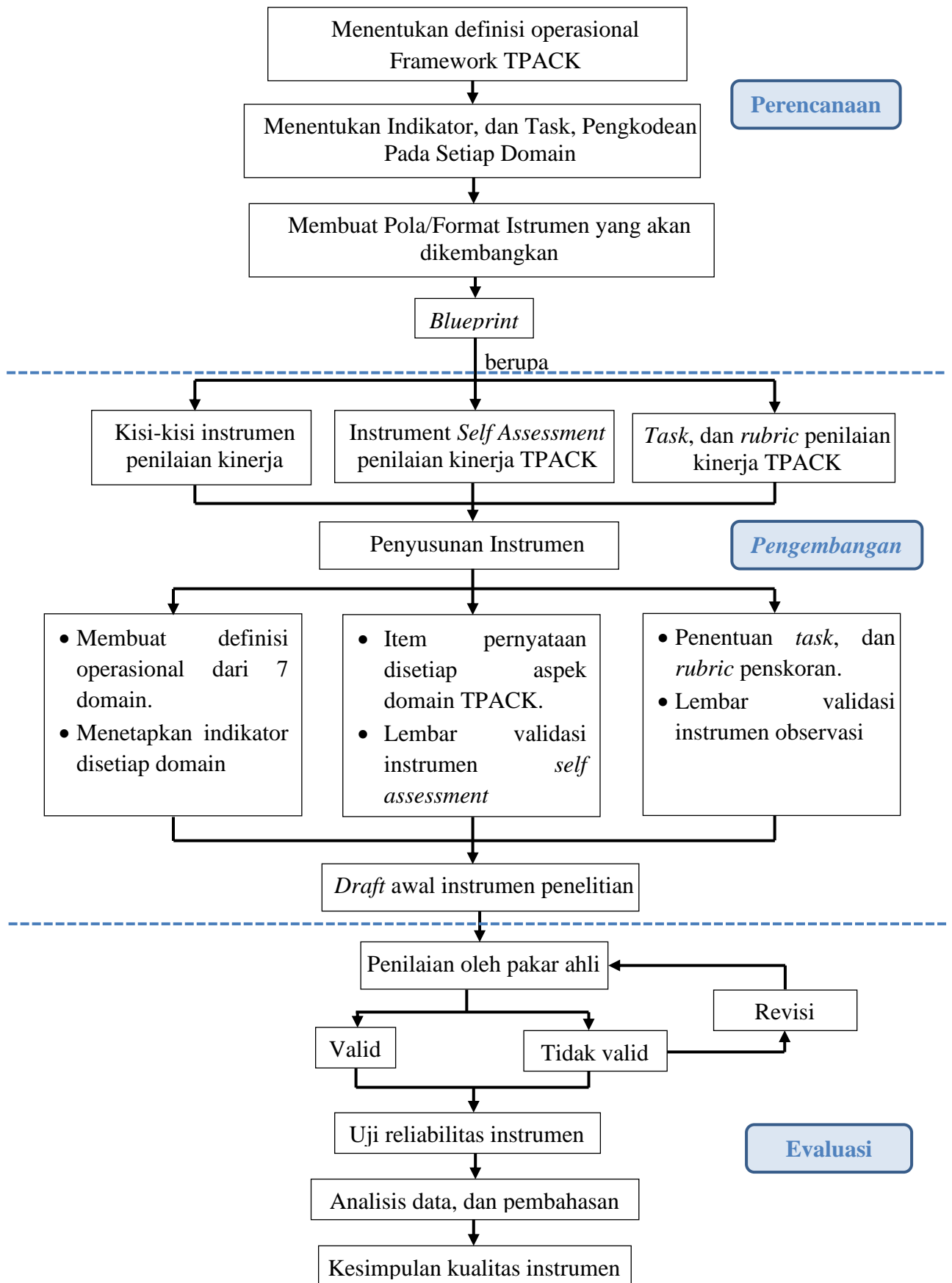
Penilaian diri dan lembar observasi dibandingkan untuk mengetahui bagaimana hubungan kedua penilaian. Kedua hasil dinormakan pada 4 pada lembar observasi dan pada *self assessment* dibuat berdasarkan skala likert 4 poin untuk skala penilaian diri. Penilaian TPACK dapat dilakukan dengan berbagai cara termasuk survei penilaian diri, dan instrumen penilaian kinerja (Koehler, Punya, & Cain, 2013). Memahami seberapa baik berbagai jenis metode pengukuran ini saling berkorelasi tetap menjadi pertanyaan penelitian yang penting. Untuk memungkinkan perbandingan seperti itu, survei penilaian diri diberikan kepada mahasiswa calon pendidik di akhir semester setelah menyelesaikan rencana pembelajaran.

3.2 Subjek, Tempat, dan Lama Penelitian

Penelitian ini dilakukan Universitas Pendidikan Indonesia pada program studi pendidikan kimia (Strata-1). Calon pendidik yang menjadi partisipan penelitian adalah mahasiswa yang telah atau sedang mengikuti perkuliahan Perencanaan Pembelajaran Kimia (PPK). Partisipan penelitian pada berjumlah 14 calon pendidik kimia, dan lima orang dosen pada program studi pendidikan kimia sebagai pakar ahli dalam menilai instrumen lembar observasi dan pernyataan angket *self assessment*. Peneliti dan tiga orang penilai dengan lulusan magister pendidikan kimia melakukan penilaian untuk menilai 14 RPP agar diperoleh indeks konsistensi instrumen. Angket *self assessment* diberikan setelah mahasiswa membuat rencana pembelajaran. Penilaian kinerja TPACK yang disusun calon pendidik merupakan bentuk rencana pembelajaran yang akan dijadikan penilaian kinerja.

3.3 Prosedur Penelitian

Pengembangan instrumen penilaian kinerja digambarkan dalam alur penelitian pada Gambar 3.1.



Gambar 3. 1 Model Pengembangan Instrumen Penilaian Kinerja dalam Mengukur TPACK

Prosedur penelitian yang diusulkan dalam rangka pengembangan instrumen penilaian kinerja TPACK calon pendidik kimia merancang pembelajaran kimia. Setiap tahapan terdiri dari langkah-langkah tertentu, setiap langkah tidak selalu merupakan urutan yang bersifat mutlak. Satu langkah dapat kembali ke langkah sebelumnya apabila diperlukan. Bagan model pengembangan instrumen penilaian kinerja TPACK dalam merancang pembelajaran kimia diusulkan dalam penelitian.

3.4 Instrumen dan Data Penelitian

Instrumen dalam penelitian ini digunakan oleh peneliti untuk menjawab pertanyaan penelitian, sehingga data yang diperoleh dari instrumen tersebut dapat dilakukan pengolahan. Instrumen/alat pengumpulan data penelitian yang akan digunakan disajikan pada tabel berikut.

Tabel 3. 1
Instrumen dan Data Penelitian

No	Rumusan Masalah	Instrumen	Data penelitian
1.	Bagaimana indikator dan <i>task</i> yang digunakan dalam instrumen penilaian kinerja TPACK?	<ul style="list-style-type: none"> • Format pemetaan setiap domain CK, TK, PK, PCK, TCK, TPK, TPACK. 	Draf awal indikator dan <i>task</i> .
2.	Bagaimana kualitas instrumen penilaian kinerja TPACK calon pendidik kimia berdasarkan validitas <i>expert judgement</i> ?	<ul style="list-style-type: none"> • Lembar validasi kinerja lembar observasi. • Lembar validasi kinerja <i>self assessment</i>. 	Hasil validasi instrumen TPACK.
3.	Bagaimana kualitas instrumen penilaian kinerja TPACK calon pendidik kimia berdasarkan reliabilitas <i>inter-rater</i> ?	<ul style="list-style-type: none"> • Lembar observasi penilaian kinerja mahasiswa. • Angket <i>self assessment</i>. 	• Nilai kinerja mahasiswa pada penilaian lembar observasi dan survei penilaian diri mahasiswa.
4.	Bagaimana hubungan antara penilaian <i>self assessment</i> dan lembar observasi pada instrumen penilaian TPACK yang telah dikembangkan?	Lembar observasi penilaian kinerja mahasiswa dan <i>self assessment</i> .	Nilai dari lembar observasi dan <i>self assessment</i> 14 orang responden.

3.5 Analisis Data Pengembangan Instrumen

Analisis data yang digunakan untuk menjawab rumusan masalah penelitian dipaparkan dalam tabel berikut:

Tabel 3. 2
Analisis Data Pengembangan Instrumen

No	Rumusan Masalah	Teknik Pengumpulan /Analisis Data	Analisis Data
1.	Bagaimana indikator dan <i>task</i> yang digunakan dalam instrumen penilaian kinerja TPACK?	<ul style="list-style-type: none"> Dokumen permendikbud Perencanaan Pembelajaran. Dokumen cpps pendidikan kimia UPI. 	<ul style="list-style-type: none"> Analisis dokumen untuk pemetaan <i>Framework</i> TPACK
2.	Bagaimana kualitas instrumen penilaian kinerja TPACK calon pendidik kimia berdasarkan validitas?	<ul style="list-style-type: none"> Analisis dokumen hasil validasi lembar observasi dan <i>self assessment</i> 	<ul style="list-style-type: none"> Nilai CVR Pola tindak lanjut dari saran berupa penambahan atau pengurangan kata atau kalimat.
3.	Bagaimana kualitas instrumen penilaian kinerja TPACK calon pendidik kimia berdasarkan reliabilitas?	<ul style="list-style-type: none"> Observasi pada Perencanaan pembelajaran oleh Penilai. Kuesioner pada instrumen <i>Assessment</i>. 	<ul style="list-style-type: none"> Indeks ICC dan W Kendall. Nilai <i>Alpha Cronbach</i>, Median, Standar Deviasi.
4.	Bagaimana hubungan antara penilaian <i>self assessment</i> dan lembar observasi pada instrumen penilaian TPACK yang telah dikembangkan?	<ul style="list-style-type: none"> Uji korelasi nonparametrik <i>Spearman</i>. Perbandingan <i>Self Assessment</i> dan lembar observasi. 	<ul style="list-style-type: none"> Nilai korelasi Spearman Mean SD

3.5.1 Analisis Dokumen

Salah satu sumber informasi berharga dalam penelitian kualitatif boleh jadi dokumen. Dokumen berisikan catatan publik dan pribadi yang didapatkan peneliti kualitatif pada tempat atau partisipan dalam suatu penelitian termasuk surat kabar, catatan harian pribadi atau surat, memo resmi, catatan di ranah publik, dan bahan

arsip perpustakaan (Creswell, 2015). Dokumen yang digunakan berupa aturan Permendikbud Standar Proses Nomor 22 Tahun 2016 dan Permendikbud Nomor 16 Tahun 2022 aturan-aturan dalam membuat perencanaan pembelajaran. Dokumen RPS perkuliahan Perencanaan Pembelajaran Kimia (PPK) digunakan dalam menentukan aspek yang akan dinilai.

3.5.2 Validitas Instrumen

Instrumen direvisi sesuai dengan saran dan masukan dari para pakar. Selanjutnya dilakukan validasi kembali sampai pakar menyatakan bahwa instrumen yang dikembangkan valid. Validasi instrumen dilanjutkan secara kuantitatif dengan cara menghitung Indeks *Content Validity Ratio* (CVR), menggunakan persamaan yang dikemukakan oleh Lawshe (1975).

Keterangan:

$$CVR = \frac{(ne - \frac{N}{2})}{\frac{N}{2}}$$

CVR = indeks *Content Validity Ratio*

ne = jumlah validator yang menyatakan setuju

N = Jumlah total validator.

Berdasarkan kriteria yang dikemukakan Lawshe (1975), item dinyatakan valid untuk jumlah pakar (panelis) lima orang, apabila indeks CVR minimum 0,99. Kriteria tersebut dilakukan pada uji satu pihak, dengan taraf signifikansi 0,05. Mencermati kriteria tersebut, maka *instrumen* akan valid apabila semua pakar setuju dengan item instrumen. Kriteria tersebut kemudian direvisi berdasarkan kajian statistik, hasil revisi merekomendasi nilai kritis CVR untuk jumlah panelis lima orang dan taraf signifikansi 0,05 pada uji satu pihak sebesar 0,736. Apabila berdasarkan hasil validasi pakar dan perhitungan indeks CVR telah memenuhi syarat sebagai *instrumen* yang valid, selanjutnya dilakukan ujicoba awal penggunaan instrumen, uji coba awal dilakukan bertujuan untuk mendapatkan masukan dari responden pengguna instrumen, guna perbaikan selanjutnya.

Menilai validitas konstruk, digunakan korelasi pearson. Korelasi pearson umumnya digunakan untuk memverifikasi intensitas hubungan linier yang ada

antara variabel dan mengukur hubungan linier antara variabel kuantitatif. Koefisien ini adalah angka antara -1 dan 1 yang menyatakan tingkat ketergantungan linear antara dua variabel kuantitatif. Jika negatif, ini menunjukkan bahwa satu variabel menurun saat variabel lainnya meningkat. Jika positif, menunjukkan bahwa satu variabel meningkat seiring dengan peningkatan aspek lainnya. Nilai r didistribusikan sebagai berikut:

Tabel 3. 3
Interpretasi Nilai r Tabel

Nilai r tabel	Keterangan
0-0,25	korelasi sangat rendah
0,26-0,49	korelasi rendah
0,5-0,69	korelasi sedang
0,7-0,89	korelasi tinggi atau kuat
0,9-1,0	korelasi sangat tinggi atau sangat kuat

korelasi pearson digunakan karena instrumen menyajikan hubungan linier antara kriteria yang disajikan (Ahrens, Lirani, & De Francisco, 2020).

3.5.3. Reliabilitas Instrumen (*Interrater Reliability*)

Uji coba instrumen dilakukan setelah dilakukan perbaikan instrumen berdasarkan masukan para dosen pengguna rubrik. Uji coba bertujuan untuk menilai tingkat reliabilitas (konsistensi) rubrik. Uji coba melibatkan 14 mahasiswa calon pendidik kimia yang sedang mengambil mata kuliah perencanaan pembelajaran kimia di Universitas Pendidikan Indonesia. Sebuah alat ukur yang baik harus mengukur dengan tepat apa yang akan diukur (valid) dan konsisten (andal, reliabel). Reliabilitas alat ukur ada dua aspek yaitu: konsistensi internal dan stabilitas. Jika sebuah instrumen terdiri dari beberapa item pertanyaan, maka skor dari setiap item pertanyaan seharusnya berkorelasi dengan skor total yang diperoleh, inilah yang dimaksud dengan konsistensi internal. Jika sebuah pengukur massa mengukur berulang kali 10 ± 0 kg beras yang terdapat dalam kemasan, sedangkan sebuah alat pengukur massa lainnya mengukur beras dalam kemasan yang sama dengan hasil ukur 10 ± 2 kg, maka alat pengukur massa pertama lebih stabil dari yang kedua (Gisev, Hons, Bell, & Chen, 2013).

Indeks konsistensi antar penilai untuk data berbentuk nominal dapat

menggunakan indeks Cohen's kappa, *Intraclass Correlation Coefficient (ICC)* dan *Weighted kappa* jika penilainya (*rater*) hanya dua orang, jika penilai lebih dari dua orang dapat menggunakan *Fleiss' kappa* dan *ICC*. Data berbentuk ordinal dengan jumlah penilai dua orang dapat menggunakan *Weighted kappa dan ICC*, jika penilai lebih dari dua orang, indeks konsistensi antar penilainya dapat menggunakan *Kendall Coefficient Concordance W dan ICC*. Sedangkan data berbentuk interval maupun rasio dapat menggunakan *Bland-Altman plots dan ICC* apabila jumlah penilai dua orang, jika jumlah penilai lebih dari dua orang dapat menggunakan *ICC* (Gisev *et al.*, 2013). Jenis indeks konsistensi antar penilai yang dapat digunakan berdasarkan jenis data dan jumlah penilai ditampilkan pada Tabel 3.4.

Tabel 3. 4
Contoh Indeks Konsistensi antar Penilai berdasarkan Jenis Data

		Jenis Data				
		Nominal		Ordinal		Interval dan Rasio
		2 penilai	> 2 penilai	2 penilai	> 2 penilai	2 penilai > 2 penilai
Indeks konsistensi antar penilai	Cohen's kappa	Fleiss kappa	Weighted kappa	Kendall's coefficient of concordance	Bland-Altman plots	ICC
	ICC	ICC	ICC	ICC	ICC	ICC
		Weighted kappa				

(diadaptasi dari Gisev *et al.*, 2013)

Pada penelitian ini, uji reliabilitas rubrik dilakukan menggunakan uji konsistensi indeks *Intraclass Correlation Coefficients (ICC)* dan indeks *Kendall's Coefficient Concordance W*. Kedua jenis uji ini digunakan untuk melihat kesesuaian skor antar individu yang diberikan penilai (*rater*). Disamping itu, data penelitian yang berbentuk ordinal dan penilaian dilakukan lebih dari 2 orang, sehingga digunakan indeks ICC dan W kendall. Data yang diuji konsistensinya berupa skor TPACK calon pendidik yang diberikan penilai berpedoman pada rubrik TPACK.

3.6.3.1 Intraclass Coefficient Correlation (ICC)

Indeks ICC diperoleh dari rasio antara varians antar kelompok dengan varians total. Varians total pada suatu pengukuran berasal dari 3 sumber, yaitu: 1) varians yang dinilai (subjek); 2) varian penilai (rater); dan *variens random error*. Jika variasi penilai diasumsikan random, maka persamaan untuk menghitung indeks ICC adalah:

$$ICC = \frac{\sigma_s^2}{\sigma_s^2 + \sigma_o^2 + \sigma_e^2}$$

Persamaan diatas digunakan untuk menghitung indeks kesepakatan antar penilai (*interrater agreement*). Kesepakatan antar penilai merupakan tingkat kecocokan antara penilai dalam memberikan skor pada objek yang dinilai. *interrater agreement* disebut juga reliabilitas absolut (*absolute reliability*). Apabila variasi penilai diasumsikan *fixed* pada persamaan di bawah, maka variasi penilai tidak diperhitungkan, sehingga persamaan ICC menjadi:

$$ICC = \frac{\sigma_s^2}{\sigma_s^2 + \sigma_e^2}$$

di mana varians (σ^2) adalah ukuran variasi, subskrip s adalah subjek, o adalah observer (penilai) dan e adalah *random error*. Persamaan tersebut digunakan untuk menghitung indeks konsistensi penilaian antar penilai (*interrater reliability*). Konsistensi penilaian antar penilai menunjukkan sejauh mana penilai mengkategorikan objek yang dinilai dengan cara yang sama. Indeks konsistensi ini juga sering disebut reliabilitas relatif (*relative reliability*). Pada penelitian ini menghitung besar indeks ICC dan indeks konkordansi W Kendall, keduanya menggunakan aplikasi SPSS versi 20. Estimasi *Intraclass Correlation Coefficients* dengan aplikasi IBM SPSS versi 26, dilakukan dengan langkah sebagai berikut.

1. Memasukkan nama penilai 1, penilai 2,dan seterusnya sampai penilai 5 pada *sheet variable view*
2. Memasukkan data pada *sheet data view*
3. Memilih menu *analyze-scale-reliability analysis*
4. Memasukkan variabel penilai1, penilai2,... dan seterusnya sampai penilai4 ke dalam kotak item

5. Memilih *statistics*, dan mencentang *F-test* dan *intraclass correlation coefficients*. Memilih jenis analisis *two way mixed* dan *consistency* dengan *confidence interval 95%*
6. Memilih *continue*, lalu mengklik ok. Setelah itu muncul *output* olahan data statistik
7. Langkah selanjutnya adalah menginterpretasi *output* olahan data dari aplikasi SPSS

Interpretasi indeks ICC dilakukan berdasarkan kriteria yang dikemukakan Koo dan Li pada tabel berikut;

Tabel 3. 5

Interpretasi Nilai Indeks ICC

Nilai Indeks ICC	Reliabilitas
<0,5	buruk
0,5-0,75	sedang
0,75-0,9	baik
0,9-1	istimewa

apabila indeks ICC lebih kecil dari 0,5 menunjukkan reliabilitas yang jelek, indeks ICC berada antara 0,5 sampai dengan 0,75 menunjukkan reliabilitas yang sedang, indeks ICC berada pada nilai antara 0,75 sampai 0,9 menunjukkan reliabilitas yang baik, dan indeks ICC dengan nilai lebih besar dari 0,9 menunjukkan reliabilitas yang istimewa (Koo & Li, 2016).

3.6.3.2 Koefisien Konkordansi W Kendall (*Kendall's Coefficient of Concordance W*)

Koefisien konkordansi W Kendall digunakan untuk mengukur kesepakatan di antara para pengamat. Kelompok yang terdiri dari empat pengamat itu sepakat ketika dinilai sebagai kelompok. Ketika masing-masing pengamat dinilai secara independen terhadap tiga lainnya, menghitung indeks W Kendall digunakan persamaan berikut:

$$W = \frac{12S}{p^2(n^3 - n) - PT}$$

Keterangan:

S = jumlah kuadrat dari jumlah baris peringkat Ri

n = jumlah objek

p = jumlah penilai (rater)

T = faktor koreksi untuk peringkat terikat

Jumlah kuadrat dari jumlah baris peringkat R_i , dihitung menggunakan persamaan di atas selanjutnya faktor koreksi untuk peringkat terikat dihitung menggunakan persamaan di bawah ini:

$$S = \sum_{i=1}^m R_i^2 = SSR$$

$$S = \sum_{k=1}^m (t_k^3 - t_k)$$

Keterangan:

m = jumlah kelompok

t_k = jumlah pangkat terikat dalam setiap (k) kelompok m

Pada penelitian ini, uji konsistensi penilaian antar penilai menggunakan indeks *Kendall's Coefficient Concordance W*, dengan memanfaatkan aplikasi IBM SPSS versi 26, dilakukan dengan langkah-langkah berikut.

1. Memasukkan nama mahasiswa1, mahasiswa2, ... dan seterusnya pada *sheet variable view*
2. Memasukkan data pada sheet data view
3. Memilih menu *analyze-nonparametric test-legacy dialogs-k related samples*
4. Memasukkan variabel mahasiswa1, mahasiswa 2, dan seterusnya ke dalam kotak *test variables* dan centang *Kendall's W*
5. Mengklik ok. Setelah itu muncul output olahan data statistik
6. Langkah selanjutnya adalah menginterpretasi output olahan data dari aplikasi SPSS

Interpretasi kategori korelasi *Spearman* rank telah diperluas untuk menginterpretasi indeks W (Levitan *et al.*, 2008). Indeks W juga diinterpretasikan dengan menganalogikannya dengan koefisien korelasi yang lain (Gisev, Bell, O'Reilly, Rosen, & Chen, 2010). Namun demikian interpretasi kategori indeks W secara khusus dikemukakan dalam tulisan yang berjudul *Managing Delphi Surveys Using Nonparametric Statistical Techniques*.

Nilai-nilai Kendall W dapat bervariasi dari 0 (tidak ada kesepakatan sama sekali) hingga 1 (kesepakatan sempurna), dengan nilai lebih tinggi dari 0,6 yang

menunjukkan kesepakatan substansial. Indeks W Kendall diinterpretasi dengan cara mencocokkan nilai indeks W hitung dengan tabel kriteria pada tabel berikut:

Tabel 3. 6

Interpretasi Nilai Indeks W Kendall

Nilai Indeks W Kendall	Keterangan
<0,11	sangat lemah
0,11-0,30	lemah
0,31-0,50	sedang
0,51-0,70	kuat
>0,71	sangat kuat

nilai indeks W lebih kecil dari 0,11, diinterpretasikan kesepakatan sangat lemah; nilai indeks W antara 0,11 sampai dengan 0,30, diinterpretasikan kesepakatan yang lemah; nilai indeks W antara 0,31 sampai dengan 0,50, kesepakatan yang sedang; nilai indeks W antara 0,51 sampai dengan 0,70, diinterpretasikan adanya kesepakatan yang kuat; dan nilai indeks W lebih besar dari 0,71, diinterpretasikan adanya kesepakatan yang sangat kuat.

3.6.3.3 Reliabilitas Alpha Cronbach

Berbagai deskriptor kualitatif yang berbeda digunakan oleh penulis untuk menginterpretasikan nilai alfa yang dihitung. Nilai alfa digambarkan dalam tabel berikut;

Tabel 3. 7

Interpretasi Nilai Alpha Cronbach

Nilai Alpha	Keterangan	Nilai Alpha	Keterangan
0,11	rendah	0,70-0,77	relatif tinggi
0,4-0,55	tidak memuaskan	0,71-0,91	baik
0,45-0,96	cukup	0,73-0,95	tinggi
0,45-0,98	dapat diterima	0,76-0,95	cukup tinggi
0,58-0,97	memuaskan	0,81	kuat
0,61-0,65	sedang	0,84-0,90	dapat diandalkan
0,64-0,85	memadai/cukup	0,91-0,93	kuat
0,67-0,87	masuk akal		sangat baik

Nilai Alpha	Keterangan
0,68	sedikit rendah

penelitian berbagai tes diterapkan pada instrumen, termasuk penggunaan alfa Cronbach, di mana skor Cronbach lebih besar dari 0,70 dianggap sebagai indikasi reliabilitas yang dapat diterima (Taber, 2018). Perhitungan instrumen penilaian kinerja TPACK *Self Assessment* menggunakan Alpha Cronbach menggunakan IBM SPSS 26.

3.5.4 Hubungan Penilaian Diri dan Lembar Observasi (*langsung ICC dan W Kendall*)

Tes korelasi Spearman pada dasarnya adalah versi nonparametrik dari koefisien korelasi Pearson dan menyediakan untuk menyelidiki hubungan linear antara dua variabel. Koefisien korelasi ini juga bisa beradaptasi dengan data ordinal. Uji korelasi Spearman adalah uji nonparametrik, maka akan tepat untuk digunakan terutama ketika data tidak terdistribusi normal, ukuran sampel yang kecil dan ada *outlier* dalam pengumpulan data (Temizhana, Mirtagioglu & Mendes, 2022). Analisis data menggunakan IBM SPSS 26 dengan langkah-langkah sebagai berikut:

1. Buka program SPSS klik *Variable View* pada kolom nama pada bagian label tambahkan label CKA,PKA hingga TPACKA *self assessment* dan label CKB, PKB hingga TPACKB untuk lembar observasi.
2. Klik data *view*, masukkan data penelitian untuk masing-masing label.
3. Klik *analyze*, pilih *correlate* kemudian *bivariate*.
4. Memasukkan variabel CKA dan CKB ke kolom *variables* centang pada pilihan Spearman kemudian pada bagian *test of significance* pilih *Two-Tailed*.
5. Klik *options*, pada *Missing Values* pilih *Exclude cases pairwise* lalu *continue*.
6. Klik Ok, akan muncul *output* korelasi *rank Spearman*.
7. Lakukan hal yang sama untuk mencari korelasi pada domain PK, TK, PCK, TCK, TPK, dan TPACK.

Interpretasi koefisien korelasi disajikan pada tabel berikut;

Tabel 3. 8
Interpretasi Nilai Koefisien Korelasi

Nilai Korelasi	Keterangan
0,00-0,10	dapat diabaikan
0,10-0,39	lemah
0,40-0,69	sedang
0,70-0,89	kuat
0,90-1,00	sangat kuat

Meskipun sebagian besar peneliti mungkin setuju bahwa koefisien $<0,1$ menunjukkan hubungan yang dapat diabaikan dan $>0,9$ sangat kuat, nilai-nilai di antara keduanya masih dapat diperdebatkan. Misalnya, koefisien korelasi sebesar 0,65 dapat diartikan sebagai korelasi “baik” atau “sedang”, bergantung pada aturan praktis yang diterapkan. Hal ini juga cukup keliru untuk mengklaim bahwa koefisien korelasi sebesar 0,39 mewakili hubungan yang “lemah”, sedangkan 0,40 adalah hubungan yang “moderat” atau ditengah-tengah (Schober, Boer, & Schwarte, 2018).