

### **BAB III METODE PENELITIAN**

#### **A. Desain Penelitian**

Penelitian ini menggunakan multimetode, yaitu: survey, kausal-komparatif, dan *delphi* yang dilaksanakan dalam dua tahap. Penelitian tahap pertama menggunakan metode survey (Heppner, Wampold, & Kivlighan, 2008). Metode survey digunakan untuk mendeskripsikan, menjelaskan, dan mengeksplorasi dimensi kearifan pribadi, kearifan multibudaya, anteseden kearifan, dan aliansi konseling menurut pandangan calon konselor, serta profil kearifan pribadi, kearifan multibudaya, dan aliansi konseling calon konselor. Metode kausal-komparatif digunakan untuk menguji pengaruh dan perbedaan faktor etno-sosio-demografi, kearifan pribadi, dan kearifan multibudaya terhadap aliansi konseling calon konselor.

Penelitian tahap kedua menggunakan metode *delphi*. *Delphi* adalah penelitian multi-tahapan untuk mendapatkan informasi dan mencapai konsensus para pakar tentang suatu isu penting (Keeney, Hasson, & McKenna, 2011). Metode *delphi* digunakan untuk mengembangkan teori (Brady, 2015; Dalkey & Helmer, 1963; Grisham, 2009), mengonstruksi skenario (Glenn & Gordon, 2009), serta mengeksplorasi, mengembangkan, menguji, dan mengevaluasi program (Day & Bobeva, 2005; Delbecq, Van de Ven, & Gustafson, 1975). *Delphi* juga dapat digunakan untuk mengembangkan model pendidikan dan kurikulum (Linstone & Turoff, 1975; 2002; 2011). Dalam penelitian ini, metode *delphi* digunakan untuk mengeksplorasi, mengembangkan, menguji, mengevaluasi, dan mencapai konsensus pakar tentang PEKA.

#### **B. Partisipan Penelitian**

##### **1. Populasi dan Sampel Penelitian Kuantitatif**

Populasi penelitian ini yaitu seluruh mahasiswa (calon konselor) semester keenam dari Program Studi Bimbingan dan Konseling (Prodi BK) jenjang S1 di Indonesia. Teknik pemilihan sampel penelitian menggunakan *two stage random*

*sampling* dengan cara diundi. Tahap pertama, pemilihan sampel perguruan tinggi yang memiliki Prodi BK jenjang S1 di Indonesia. Atas dasar pertimbangan sumber daya penelitian, ditetapkan 13% atau 11 dari 85 Perguruan Tinggi penyelenggara Prodi BK jenjang S1 di Indonesia (<https://www.abkin.org>, 2011; <http://evaluasi.dikti.go.id>, 2011). Tahap kedua, menentukan jumlah anggota sampel penelitian menggunakan rumus Slovin dengan tingkat kepercayaan 99%. Jumlah anggota sampel secara rinci disajikan pada Tabel 3.1 dan Tabel 3.2.

$$n = \frac{N}{1 + Ne^2}$$

Keterangan:

n = Ukuran sampel

N = Ukuran populasi

e = Persentase kelonggaran ketidakteelitian karena kesalahan pengambilan sampel yang masih dapat diterima atau diinginkan

**Tabel 3.1. Populasi dan Anggota Sampel Penelitian**

No.	Perguruan Tinggi	Populasi	Sampel			
			Rencana	Realisasi	RV*	%
1.	UNNES	127	116	42	40	7.7
2.	UNY	85	78	65	64	12.4
3.	UM	82	75	51	51	9.9
4.	UNJ	77	72	72	72	13.9
5.	UNM	80	73	70	68	13.2
6.	UNDIKSHA	33	30	28	28	5.4
7.	UHAMKA	126	115	58	58	11.2
8.	UAD	203	186	81	80	15.5
9.	ATMAJAYA	28	26	21	21	4.1
10.	UKI	11	11	11	11	2.1
11.	USD	65	60	24	24	4.6
		<b>917</b>	<b>842</b>	<b>523</b>	<b>517</b>	<b>100</b>

Keterangan: RV = Sampel yang mengisi lengkap seluruh instrumen penelitian

Variabel etno-sosio-demografi sampel, meliputi: konteks pendidikan, jenis kelamin, usia, agama, etnis, pengalaman hidup kritis, dan frekuensi konseling dikaji dalam penelitian ini. Hal ini dilakukan berdasarkan pertimbangan berikut. Pertama, bentuk respon nyata terhadap rekomendasi pakar tentang pentingnya

mengeksplorasi dan menguji beragam anteseden kearifan, meliputi: (a) konteks-fasilitatif, seperti: usia, jenis kelamin, pendidikan, agama, etnis, dan budaya; (b) kepakaran, di antaranya: pengalaman praktik berkelanjutan, pendidikan dan pelatihan profesional, dan pengalaman hidup kritis; dan (c) pribadi, seperti: ciri-ciri kepribadian (Kunzmann & Baltes, 2005; Stange & Kunzmann, 2008; Staudinger, 1999; Staudinger & Gluck, 2011). Kedua, respon nyata terhadap rekomendasi pakar tentang pentingnya mengases, memahami, dan mempertimbangkan faktor yang dapat meningkatkan atau menghambat efektivitas proses dan luaran konseling multibudaya, seperti: identitas religius, status ekonomi, identitas seksual/kesadaran gender, kematangan psikologis, identitas etnis-budaya-ras, tantangan perkembangan kronologis (usia), pengalaman traumatik, sejarah dan dinamika keluarga, keunikan karakteristik fisik, serta perbedaan lokasi tempat tinggal dan bahasa (Baruth & Manning, 2016; D'Andrea & Daniels, 2001; 2005; Neukrug, 2012).

**Tabel 3.2 Rincian Sampel Penelitian Kuantitatif**

<b>Kelompok</b>	<b>F</b>	<b>%</b>	<b>Kelompok</b>	<b>F</b>	<b>%</b>
Institusi Pendidikan			Jenis Kelamin		
UNNES	40	7.7	Laki-laki	135	26.1
UNY	64	12.4	Perempuan	382	73.9
UM	51	9.9	Usia		
UNJ	72	13.9	≤ 20 Tahun (Remaja Akhir)	159	30.8
UNM	68	13.2	> 20 Tahun (Dewasa Muda)	358	69.2
UNDIKSHA	28	5.4	Agama		
UHAMKA	58	11.2	Islam	427	82.6
UAD	80	15.5	Non-Islam	90	17.4
ATMAJAYA	21	4.1	Etnis		
UKI	11	2.1	Jawa	287	55.5
USD	24	4.6	Non-Jawa	230	44.5
Status Institusi Pendidikan			Pengalaman Hidup Kritis		
PTN	323	62.5	Tidak Pernah	287	55.5
PTS	194	37.5	Pernah	230	44.5
Status Akreditasi Prodi BK			Frekuensi Konseling		
A	314	60.7	≤ 3x	283	54.7
B	203	39.3	> 3x	234	45.3
			<b>Jumlah</b>	<b>517</b>	<b>100</b>

## 2. Subjek Penelitian *Delphi*

Subjek penelitian *delphi* berjumlah sembilan orang pakar. Jumlah pakar ini ditetapkan berdasarkan tujuan penelitian dan kajian literatur bahwa *delphi* akan berdaya guna jika melibatkan 7 – 15 pakar (Dalkey & Helmer, 1963) yang

homogen (Adler & Ziglio, 1996; Hussler, Muller, & Ronde, 2011). Pakar dipilih secara *purposive* dengan kriteria berikut. Pertama, memiliki kepakaran yang relevan (Brady, 2015; Goluchowicz & Blind, 2011; Donohoe & Needham, 2009). Sembilan pakar memiliki kepakaran dalam BK dengan spesialisasi di bidang: asesmen dan tes, pedagogik, orientasi profesional konseling dan praktik etik, perkembangan peserta didik, konseling individu, bimbingan dan konseling kelompok, penelitian, serta evaluasi dan pengembangan program BK. Kepakaran sembilan pakar BK ini dibuktikan dengan publikasi karya ilmiah yang relevan, seperti buku dan artikel di jurnal.

Kedua, memiliki pengalaman profesional dan keterlibatan dalam penelitian dan pembuatan kebijakan yang relevan minimal 10 – 15 tahun (Imang & Ngah, 2012; Roy, Chan, & Ahmed, 2014). Pakar BK yang terpilih memiliki pengalaman profesional sebagai pendidik konselor, tester asesmen psikologis, dan konselor selama > 15 tahun. Pakar BK memiliki pengalaman keterlibatan dalam penelitian yang relevan, baik di tingkat lokal, nasional, regional, maupun internasional dengan dana mandiri, hibah internal perguruan tinggi, hibah DRPM Kemristekdikti, maupun hibah dari lembaga lainnya. Pakar BK memiliki pengalaman dalam pembuatan kebijakan yang relevan, baik dalam bidang pendidikan maupun BK di tingkat internal perguruan tinggi dan nasional (Kementerian Pendidikan dan Kebudayaan). Pakar juga terlibat aktif dalam organisasi profesi yang relevan (Asosiasi Bimbingan dan Konseling Indonesia [ABKIN]).

Ketiga, dapat memberikan kontribusi yang bermanfaat, dan bersedia merevisi penilaiannya pada setiap putaran untuk mencapai konsensus. Para pakar memiliki kepakaran di bidang BK sehingga diprediksi dapat memberikan kontribusi yang bermanfaat bagi penyempurnaan PEKA. Selain itu, pakar bersedia berpartisipasi dalam penelitian *delphi* ini yang dibuktikan dengan menandatangani lembar kesediaan.

Pakar yang dipilih bersifat homogen dari segi latar belakang kepakaran dan institusi kerja berdasarkan dua pertimbangan utama. Pertama, pertimbangan teoretis bahwa *pilot project* dalam *delphi* akan berdaya guna jika melibatkan

pakar yang homogen (Adler & Ziglio, 1996; Hussler, Muller, & Ronde, 2011). Dalam hal ini, prinsip representatif dalam *delphi* dapat diabaikan karena yang terpenting adalah relevansi kepakarannya (Hasson, Keeney, & McKenna, 2000; Keeney, Hasson, & McKenna, 2011). Kedua, pertimbangan teknis efisiensi ekonomis seperti keterbatasan dana, waktu, dan tenaga penelitian (Dattalo, 2008). Identitas para pakar yang terpilih berdasarkan kriteria yang ditetapkan disajikan pada Tabel 3.3. Identitas setiap pakar dirahasiakan dari pakar lainnya untuk menjamin kerahasiaan dan kebebasan dalam mengemukakan gagasannya (Hsu & Stanford, 2007).

**Tabel 3.3. Subjek Penelitian *Delphi***

No.	Identitas Pakar	Bidang Kepekararan	Tingkat Pendidikan	Jabatan Fungsional
1.	P1SY	BK, Perkembangan Manusia, Kesehatan Mental	S3-BK	Guru Besar
2.	P2JN	BK, Konseling dan Hubungan Bantuan	S3-BK	Guru Besar
3.	P3MS	BK, Pedagogik	S3-BK	Lektor Kepala
4.	P4S	BK, Pedagogik, Bimbingan Belajar		Lektor Kepala
5.	P4EFZ	BK	S3-BK	Lektor Kepala
6.	P6NMS	BK, Perkembangan Manusia, BK Keluarga	S3-BK	Lektor Kepala
7.	P7YRY	Konseling dan Hubungan Bantuan	S3-BK	Lektor Kepala
8.	P8MA	BK, Perkembangan Manusia	S3-BK	Lektor Kepala
9.	P9ESY	BK, Asesmen dan Evaluasi	S3-BK	Lektor

### C. Instrumen Penelitian

#### 1. Instrumen Penelitian Kuantitatif

##### a. *The Three-Dimensional Wisdom Scale*

##### 1) *Instrumen yang Diadaptasikan*

Data kearifan pribadi calon konselor dikumpulkan menggunakan *the New Three-Dimensional Wisdom Scale-Short Form (3D-WS-SF)* versi calon konselor Indonesia (Herdi, Kartadinata, & Taufiq, 2019d) hasil adaptasi dari *3D-WS* (Ardelt, 2003). *The 3D-WS* dipilih dan digunakan untuk mengukur kearifan calon konselor didasarkan pada pertimbangan berikut: (a) memiliki konstruk teoretis yang kuat; (b) memiliki properti psikometrik yang baik; (c) banyak dirujuk dan digunakan oleh peneliti lain; dan (d) ketersediaan dan kemudahan dalam perizinan untuk penggunaannya. Thomas, Bangen, Ardel, and Jeste (2017) berpendapat bahwa *3D-WS-SF* merupakan alat asesmen kearifan

yang dikembangkan secara teliti, berdaya guna, dan memiliki properti psikometrik yang baik.

*The 3D-WS* dikembangkan oleh Ardel (2003) berlandaskan pada teori dari (Clayton & Birren, 1980) bahwa kearifan adalah integrasi dari karakteristik kepribadian, meliputi dimensi kognitif, reflektif, dan afektif. *The 3D-WS* digunakan untuk mengukur tiga dimensi kearifan, yaitu dimensi kognitif (14 item), reflektif (12 item), dan afektif (13 item). *The 3D-WS* berbentuk skala Likert lima jenjang dari 1 = sangat tidak sesuai/sangat tidak setuju sampai dengan 5 = sangat sesuai/sangat setuju. Lima belas item dari *3D-WS* menggunakan skala 5 jenjang, yaitu: 1 = sangat tidak setuju, 2 = tidak setuju, 3 = netral, 4 = setuju, atau 5 = sangat setuju. Dua puluh empat item dari *3D-WS* menggunakan skala lima jenjang, yaitu: 1 = sangat tidak sesuai dengan diri sendiri, 2 = tidak sesuai dengan diri sendiri, 3 = cukup sesuai dengan diri sendiri, 4 = sesuai dengan diri sendiri, atau 5 = sangat sesuai dengan diri sendiri.

Kajian empiris menunjukkan bahwa *3D-WS* memiliki properti psikometrik yang baik, meliputi: validitas (validitas isi, validitas konstruk, validitas prediktif, validitas diskriminan, dan validitas konvergen), dan konsistensi internal yang baik (.72 untuk dimensi kognitif, .74 untuk dimensi reflektif, dan .66 untuk dimensi afektif), dan reliabilitas *test-retest* yang tinggi dari pengukuran selama 10 bulan (Ardelt, 2003). *The 3D-WS* digunakan secara luas untuk sampel remaja (Bailey & Russell, 2008), dewasa muda/mahasiswa, dan dewasa akhir dari beragam latar belakang budaya (Ardelt & Oh, 2010b; Benedikovicova & Ardel, 2008; Ferrari, Kahn, Benayon, & Nero, 2011; Redzanowski & Gluck, 2013). *The 3D-WS* berkorelasi positif signifikan ( $r = .33$ ) dengan *SAWS* pada sampel orang Australia dari beragam usia (Ardelt, 2011a; Taylor, Bates, & Webster, 2011). Kajian terkini menunjukkan *3D-WS-12* menjadi alat asesmen kearifan yang efisien dan valid dalam konteks survey epidemiologis (Thomas, Bangen, Ardel, & Jeste, 2017).

Tabel 3.4. Kisi-kisi 3D-WS dan 3D-WS-SF

Variabel	Dimensi	Indikator	3D-WS			3D-WS-SF		
			+	-	$\Sigma$	+	-	$\Sigma$
Kearifan Pribadi	Kognitif	▪ Kemampuan atau keinginan untuk memahami peristiwa atau fenomena		C <sub>3</sub> , C <sub>2</sub> , C <sub>1</sub> , C <sub>10</sub> , C <sub>14</sub> , C <sub>11</sub> , C <sub>12</sub>	14		C <sub>11</sub> , C <sub>12</sub>	4
		▪ Pengetahuan tentang ambiguitas, kompleksitas, dan ketidakpastian dalam hidup		C <sub>8</sub> , C <sub>6</sub> , C <sub>4</sub> , C <sub>5</sub> , C <sub>7</sub> , C <sub>9</sub>			C <sub>9</sub>	
		▪ Kemampuan untuk membuat keputusan penting dalam kehidupan yang tidak dapat diprediksi		C <sub>13</sub>			C <sub>13</sub>	
Reflektif	▪	Kemampuan dan keinginan untuk melihat fenomena dan peristiwa dari perspektif yang berbeda	R <sub>9</sub> , R <sub>7</sub> , R <sub>3</sub> , R <sub>4</sub> , R <sub>5</sub> , R <sub>8</sub> , R <sub>6</sub> ,		12		R <sub>9</sub>	4
		Ketiadaan subjektivitas dan proyeksi	R <sub>2</sub> , R <sub>1</sub> , R <sub>10</sub> , R <sub>12</sub> , R <sub>11</sub>				R <sub>10</sub> , R <sub>11</sub> , R <sub>12</sub>	
Afektif	▪	Emosi positif dan kepedulian terhadap orang lain		A <sub>4</sub> , A <sub>7</sub>	13		A <sub>4</sub>	4
		Motivasi untuk mendorong kesejahteraan psikologis orang lain	A <sub>5</sub> , A <sub>8</sub> , A <sub>9</sub>	A <sub>13</sub>			A <sub>8</sub> , A <sub>9</sub>	
		Ketiadaan emosi negatif terhadap orang lain	A <sub>1</sub> , A <sub>2</sub> , A <sub>3</sub> , A <sub>10</sub> , A <sub>6</sub> , A <sub>11</sub> , A <sub>12</sub>				A <sub>11</sub>	
<b>Jumlah</b>			<b>22</b>	<b>17</b>	<b>39</b>	<b>7</b>	<b>5</b>	<b>12</b>

\*) Pada lampiran instrumen, nomor item 3D-WS-SF diurutkan kembali menjadi C<sub>1</sub>-C<sub>4</sub>, R<sub>1</sub>-R<sub>4</sub>, dan A<sub>1</sub>-A<sub>4</sub> mengikuti urutan pada nomor aslinya.

## 2) Prosedur Pengadaptasian

Pengadaptasian 3D-WS diawali dengan perizinan untuk menggunakannya kepada Prof. Monika Ardel, PhD. dari University of Florida sebagai pengembang dan pemilik hak ciptanya. Prosedur pengadaptasian 3D-WS mengikuti pedoman umum pengadaptasian instrumen (Brislin, 1986; *International Test Commission*, 2016; Bartram, Hambleton, Gregoire, & Muniz, 2018) yaitu: “penerjemahan, penyuntingan, dan penerjemahan kembali oleh penerjemah profesional yang

berbeda.” Prosedur pengadaptasian *3D-WS* secara rinci dalam penelitian ini adalah sebagai berikut.

Tahap pertama, menerjemahkan item. Penerjemahan item *3D-WS* dari bahasa Inggris ke bahasa Indonesia dilakukan oleh pakar Bahasa Inggris dan Linguistik dari Balai Bahasa Universitas Pendidikan Indonesia (UPI).

Tahap kedua, menyunting dan mengintegrasikan hasil terjemahan. Hasil terjemahan dari kedua penerjemah disunting dan diintegrasikan ke dalam rumusan saduran sementara. Kegiatan tahap ini dilakukan sendiri oleh peneliti. Tujuan tahap ini adalah mendapatkan rumusan terjemahan yang tepat dari penerjemah.

Tahap ketiga, menerjemahkan hasil saduran sementara. Hasil saduran pada tahap ketiga, diterjemahkan kembali ke dalam bahasa Inggris. Penerjemahan kembali dimaksudkan untuk menguji kecocokan isi antara hasil terjemahan dengan bahasa aslinya. Proses penerjemahan hasil saduran dilakukan oleh pakar bahasa Inggris dan Linguistik yang berbeda dari penerjemah sebelumnya dari Balai Bahasa UPI.

Tahap keempat, menyunting dan mengintegrasikan hasil terjemahan pertama dan kedua. Tahap ini dilakukan sendiri oleh peneliti. *The 3D-WS* yang digunakan dalam penelitian ini sedikit mengalami beberapa perubahan istilah dari versi aslinya. Perubahan yang dimaksud yaitu menghapus kata “layaknya binatang” karena dinilai memiliki rasa bahasa yang tidak sesuai dengan budaya Indonesia.

Tahap kelima, mengujicobakan *3D-WS*. Kegiatan uji coba *3D-WS* dilakukan terhadap 517 calon konselor di Indonesia. Analisis data hasil uji coba *3D-WS* berlandaskan pendekatan *Item Response Theory (IRT)* menggunakan *Rasch Model v. 3.73* dan *Classical Test Theory (CTT)* sebagai pembandingan.

### **3) Pengujian Properti Psikometrik**

#### **a) Analisis Kelayakan Item**

Gambaran bukti level kelayakan item *3D-WS* dianalisis menggunakan kriteria *item measure*, *item fit order*, dan *DIF*.

*Item Measure* digunakan untuk mengukur tingkat kesulitan item. Hasil pengujian menunjukkan item C<sub>7</sub> dengan + 1.05 logit menunjukkan item yang paling sukar disetujui oleh partisipan, sedangkan item A<sub>9</sub> dengan nilai -.69 logit merupakan item yang paling mudah disetujui.

*Item Fit Order* digunakan untuk menguji item yang *fit* dan *misfit*. Hasil penelitian menemukan bahwa terdapat 24 dari 39 item yang *fit*. Keputusan ini didasarkan pada kriteria bahwa item dikatakan *fit* jika nilai *Infit* dan *Outfit Mnsq* berada pada rentang nilai yang dapat diterima ( $.5 Mnsq < 1.5$ ), *Outfit Zstd* ( $-2 < Zstd < +2$ ), *Pt-MCorr* ( $.4 < pt-MCorr .85$ ) (Bond & Fox, 2015; Dimitrov, 2012; Linacre, 2019; Sumintono & Widhiarso, 2014; Sumintono & Widhiarso, 2015; Boone, Staver, & Yale, 2014). Dalam penelitian ini, kriteria *Outfit Zstd* tidak digunakan karena ukuran sampel besar  $> 500$  orang (Sumintono & Widhiarso, 2014; 2015). Item-item yang *misfit*, yaitu: C<sub>5</sub>, C<sub>6</sub>, C<sub>7</sub>, C<sub>10</sub>, R<sub>1</sub>, R<sub>2</sub>, R<sub>3</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub>, R<sub>7</sub>, R<sub>8</sub>, A<sub>7</sub>, A<sub>10</sub>, A<sub>13</sub>.

*DIF* digunakan untuk mendeteksi bias item dalam kategori partisipan tertentu. Item yang terdeteksi bias diketahui berdasarkan nilai probabilitas yang kurang dari 5% atau .05 (Bond & Fox, 2015; Dimitrov, 2012; Linacre, 2019; Sumintono & Widhiarso, 2014; 2015). Hasil pengujian menampilkan 11 item mengalami bias terhadap jenis kelamin tertentu, yaitu: C<sub>1</sub> (.000), C<sub>4</sub> (.01), C<sub>6</sub> (.01), C<sub>7</sub> (.001), C<sub>10</sub> (.04), C<sub>14</sub> (.003), R<sub>1</sub> (.002), R<sub>3</sub> (.02), R<sub>8</sub> (.009), R<sub>9</sub> (.006), A<sub>2</sub> (.01). Sebanyak dua item mengalami bias terhadap usia tertentu, yaitu: R<sub>6</sub> (.03) dan A<sub>1</sub> (.01). Satu item mengalami bias terhadap afiliasi agama tertentu, yaitu: A<sub>6</sub> (.01).

Sebagai pembanding, kelayakan item juga diuji dengan *item-total correlation* ( $r_{it}$ ) dari pendekatan *Classical Test Theory* (CTT) menggunakan *IBM SPSS 25.0 for Windows*. Hasil pengujian menemukan 12 dari 39 item yang memiliki indeks koefisien  $r_{it} \leq .300$ . C<sub>5</sub>, C<sub>7</sub>, R<sub>1</sub>, R<sub>2</sub>, R<sub>3</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub>, R<sub>8</sub>, A<sub>7</sub>, A<sub>10</sub>, A<sub>13</sub> yang tidak memenuhi kriteria yang disarankan oleh pengembang instrumen (Ardelt, 2016; melalui komunikasi pribadi via email) dan pendapat pakar psikometrik (Azwar, 2011; Drummond & Jones, 2010). Hasil analisis kelayakan

item *3D-WS* menggunakan *Rasch Model 3.73* dan *IBM SPSS v. 25.0* terlampir pada Lampiran D.

b) Analisis Abilitas Partisipan

Gambaran bukti level abilitas responden dianalisis dari kriteria *person measure* menggunakan *Rasch Model*. Hasil pengujian *Person Measure* menampilkan responden nomor urut 141 (1PDIE) dengan + 2.03 logit cenderung memiliki kearifan pribadi yang tinggi karena lebih banyak menjawab sesuai dan sangat sesuai, sedangkan responden nomor urut 201 (1LDIC) dengan nilai -1.08 logit cenderung memiliki kearifan pribadi yang rendah karena lebih banyak menjawab ke arah tidak sesuai dari item-item *3D-WS*.

c) Analisis Kualitas Instrumen

Gambaran bukti level kualitas instrumen dianalisis dari kriteria *unidimensionality*, *rating scale*, dan *reliability (test, item, dan person)* menggunakan *Rasch Model*, *Exploratory Factor Analysis (EFA)* menggunakan *IBM SPSS v. 25.0*.

*Unidimensionality* diukur menggunakan analisis komponen utama (*Principle Component Analysis [PCA]*) dari residual yakni mengukur tingkat uniformitas dari instrumen dalam mengukur apa yang seharusnya diukur (Linacre, 2019; Sumintono & Widhiarso, 2014; 2015). Hasil pengukuran menampilkan *raw variance* sebesar 23.4% serta varians tertinggi yang tidak dapat dijelaskan oleh instrumen sebesar 8.3%. Hal ini berarti bahwa persyaratan unidimensionalitas minimal 20% dan varians yang tidak dapat dijelaskan oleh instrumen  $\leq 15\%$  terpenuhi (Bond & Fox, 2015; Dimitrov, 2012; Linacre, 2019; Sumintono & Widhiarso, 2014; 2015).

*Rating Scale*. Ini merupakan pengujian untuk memverifikasi pilihan peringkat yang digunakan dalam instrumen dapat membingungkan partisipan atau tidak (Bond & Fox, 2015; Dimitrov, 2012; Linacre, 2019; Sumintono & Widhiarso, 2014; 2015). Hasil analisis *Rasch Model* menemukan bahwa rata-rata nilai pengamatan dimulai dari logit -2.70 untuk pilihan skor 1 (sangat tidak

sesuai), logit -1.15 untuk pilihan skor 2 (tidak sesuai), logit -.03 untuk pilihan skor 3 (cukup sesuai), logit +1.14 untuk pilihan skor 4 (sesuai), dan logit +2.77 untuk pilihan skor 5 (sangat sesuai). Hasil senada juga ditampilkan oleh *Andrich Threshold* yang menguji ketepatan nilai politomi yang digunakan. Nilai tersebut menunjukkan pergerakan dari *NONE* ke negatif (-1.33, -.61) dan terus mengarah ke positif (.49, 1.45) secara berurutan. Artinya, pilihan skala peringkat yang digunakan sudah valid dan tepat bagi partisipan.

*Reliability*. Tabel 3.5 Ringkasan Statistik menampilkan *person measure* = +.34 logit. Nilai rata-rata yang lebih tinggi dari logit .0 menunjukkan kecenderungan partisipan yang lebih banyak menjawab penting pada setiap item. Nilai *Alpha Cronbach* untuk mengukur reliabilitas tes yakni interaksi antara *person* dan item secara keseluruhan menunjukkan .85 yang termasuk bagus. Sementara itu, nilai *person reliability* sebesar .82 dan *item reliability* .98 sehingga dapat disimpulkan bahwa konsistensi jawaban partisipan termasuk bagus dan seperangkat tes termasuk istimewa (Sumintono & Widhiarso, 2014; 2015). Hasil senada ditampilkan oleh *Infit* dan *Outfit Mnsq* untuk *person* dan item, nilai rata-ratanya secara berurutan 1.02 dan 1.01 dan 1.00 dan 1.01 yang mendekati nilai ideal 1.00. Demikian juga *Infit* dan *Outfit Zstd* menunjukkan nilai rata-rata *person* dan item sebesar -.4 dan -.4 serta -.1 dan .2 yang berarti baik karena mendekati .0.

**Tabel 3.5 Ringkasan Statistik 3D-WS dan 3D-WS-SF**

Kelayakan	Luaran	Indeks Hasil	
		3D-WS	3D-WS-SF
Item	Reliabilitas item	.98	.98
	Indeks pengelompokkan	7.81	6.57
	Strata pengelompokkan (H)	3.15	3.15
	Nilai logit tertinggi	+ 1.05 logit (C <sub>7</sub> )	+.54 logit (R <sub>1</sub> )
	Nilai logit terendah	-.69 logit (A <sub>9</sub> )	-.83 logit (A <sub>3</sub> )
Person	Reliabilitas partisipan	.82	.81
	Indeks pengelompokkan	2.11	2.04
	Strata pengelompokkan (H)	3.15	3.15
	Nilai logit tertinggi	+ 2.03 (141; 1PDIE)	+5.79(246; 6PDIC)
	Nilai logit terendah	- 1.08 (201; 1LDIC)	-2.62 (426; 6LDIH)
Instrumen	<i>Alpha Cronbach</i>	.85	.85
	<i>Raw variance explained by measures</i>	23.4	40.0
	<i>Unexplained variance in 1<sup>st</sup> contrast</i>	8.3	11.5
	<i>Unexecpected variance in 2<sup>nd</sup> contrast</i>	6.2	9.3
	<i>Unexecpected variance in 3<sup>rd</sup> contrast</i>	5.0	6.1
	<i>Unexecpected variance in 5<sup>nd</sup> contrast</i>	3.1	4.8

Nilai *separation* digunakan untuk mengelompokkan *person* dan item. Makin besar nilai *separation* menunjukkan makin bagus kualitas instrumen karena dapat mengidentifikasi kelompok responden dan kelompok item. Persamaan yang digunakan yaitu:  $H = \frac{[(4 \times 2,11) + 1]}{3} = 3.15$  dibulatkan menjadi 3. Hasil ini bermakna terdapat tiga kelompok partisipan.

#### 4) Bentuk Akhir 3D-WS

Dengan mempertimbangkan berbagai hasil pengujian kelayakan, maka 3D-WS-SF versi calon konselor Indonesia dipilih dan digunakan dalam penelitian ini. Instrumen ini diharapkan dapat menjadi alat asesmen kearifan calon konselor yang lebih efisien dan valid. *The 3D-WS-SF* mengukur tiga dimensi kearifan, yaitu: kognitif, reflektif, dan afektif yang setiap dimensinya berjumlah 4 item. Item final 3D-WS-SF, yaitu: C<sub>9</sub>, C<sub>11</sub>, C<sub>12</sub>, C<sub>13</sub>, R<sub>9</sub>, R<sub>10</sub>, R<sub>11</sub>, R<sub>12</sub>, A<sub>4</sub>, A<sub>8</sub>, A<sub>9</sub>, dan A<sub>11</sub>. Selanjutnya nomor-nomor item diurutkan menjadi C<sub>1</sub>, C<sub>2</sub>, C<sub>3</sub>, C<sub>4</sub>, R<sub>1</sub>, R<sub>2</sub>, R<sub>3</sub>, R<sub>4</sub>, A<sub>1</sub>, A<sub>2</sub>, A<sub>3</sub>, dan A<sub>4</sub> mengikuti nomor urut aslinya. *The 3D-WS-SF* merupakan alat ukur laporan diri berbentuk *summated rating* dengan skala Likert 5 jenjang yang ketentuannya disajikan pada Tabel 3.6. Kisi-kisi 3D-WS-SF versi calon konselor Indonesia disajikan pada Tabel 3.4 dan instrumen lengkapnya dapat diperiksa pada Lampiran C.

**Tabel 3.6. Pedoman Penskoran 3D-WS-SF**

Aitem	Skor Alternatif Jawaban				
	Sangat Setuju/ Sangat Sesuai	Setuju/Sesuai	Netral/ Cukup Sesuai	Tidak Setuju/ Tidak Sesuai	Sangat Tidak Setuju/Sangat Tidak Sesuai
Positif	5	4	3	2	1
Negatif	1	2	3	4	5

#### b. *The New Wisdom Scale for Multicultural Counselor*

##### 1) *Prosedur Pengembangan*

Data kearifan calon konselor multibudaya dikumpulkan menggunakan Skala Kearifan Calon Konselor Multibudaya (SKKM). Penyebutan selanjutnya menggunakan istilah *the New Wisdom Scale for Multicultural Counselor/MC-WiSe* (Herdi, Kartadinata, & Taufiq, 2019c). *MC-WiSe* dikonstruksi sendiri

oleh peneliti melalui prosedur berikut: (a) mengkaji literatur; (b) mengonseptualisasikan dan mengoperasionisasikan konstruk ukur; (c) mengidentifikasi indikator, menyusun item, dan menetapkan format respon; (d) menganalisis isi dan *pilot testing*, merevisi, dan mengadministrasikan aitem; (e) melakukan uji coba; (f) menganalisis faktor, memfinalisasi item, dan menguji properti psikometrik (Heppner, Wampold, & Kivlighan, 2008).

Tahap pertama dan kedua, mengonseptualisasikan dan mengoperasionisasikan konstruk ukur serta mengkaji literatur. *MC-WiSe* dikonstruksi berlandaskan pada konsep kearifan konselor multibudaya dari (Hanna, Bemak, & Chung, 1999; Hanna & Ottens, 1995). Dimensi dan indikator kearifan tersebut teridentifikasi juga dari *Balance Theory of Wisdom* (Sternberg, 1985; 1990), seni penemuan dan pengentasan masalah (Arlin, 1990), penalaran relativistik dan dialektik untuk mengambil perspektif jamak (Kramer, 1990), dan teori-teori kearifan lainnya yang telah dibahas di BAB II.

Tahap ketiga, mengidentifikasi indikator, menyusun item, dan menetapkan format respon. *MC-WiSe* mengukur dua dimensi, yaitu: dimensi kognitif dengan enam indikator (metakognitif, ketajaman pandangan, penalaran dialektik, toleransi terhadap ambiguitas, pengentasan masalah, dan keterampilan  *coping*) dan dimensi afektif dan kesadaran (empati, kepedulian, mengenali emosi, deotomatisasi, dan kecerdikan). Berdasarkan indikator, selanjutnya dikonstruksi 68 item. Pola respon yang dipilih berbentuk Skala Likert lima jenjang yaitu: 1 = Sangat Tidak Sesuai; 2 = Tidak Sesuai; 3 = Cukup Sesuai; 4 = Sesuai; dan 5 = Sangat Sesuai menggambarkan diri saya. Hasil identifikasi indikator kearifan calon konselor multibudaya disajikan pada Tabel 3.7.

Tahap keempat, menganalisis isi dan *pilot testing*, merevisi, dan mengadministrasikan item. Analisis/validasi isi *MC-WiSe* dilakukan oleh tiga orang penimbang pakar dengan kepakaran di bidang BK (Prof. Dr. Syamsu Yusuf L.N., M.Pd.), BK dan psikologi pendidikan (Prof. Dr. Juntika Nurihsan, M.Pd.), dan psikologi, psikometrik, penelitian dan evaluasi pendidikan (Prof. Dr. Awaluddin Tjalla, M.Psi., M.Pd.). Pada tahap ini tidak dilakukan analisis *inter-rater reliability* karena hanya dua dari tiga penimbang pakar yang

menilai setiap item. Namun, ketiga penimbang pakar memberikan penilaian umum *MC-WiSe* seperti yang disajikan pada Tabel 3.8.

**Tabel 3.7 Kisi-kisi *MC-WiSe* dan *MC-WiSe-SF***

Var	Dimensi	Indikator	<i>MC-WiSe</i>			<i>MC-WiSe-SF</i>		
			+	-	Σ	+	-	Σ
Kearifan Konselor	Kognitif	▪ Metakognisi	1, 4, 6	2, 3, 5	6	1, 6	0	2
		▪ Ketajaman pandangan	7, 11, 12	8, 9, 10	6	7, 11	0	2
		▪ Penalaran dialektik	14	13, 15, 16, 17	5	0	13, 17	2
		▪ Toleransi terhadap ambiguitas	18	19, 20, 21	4	18	19	2
		▪ Penemuan dan pengentasan masalah	22, 25	23, 24	4	22, 25	0	2
		▪ Keterampilan <i>coping</i>	26, 29	27, 28, 30, 31	6	27	30	2
	Afektif dan kesadaran	▪ Empati	32, 34, 36	33, 35, 37	6	32	37	2
		▪ Kepedulian	38, 41	39, 40	4	38, 41	0	2
		▪ Mengenali (Mengelola emosi)	42, 44, 46, 49, 50, 51	43, 45, 47, 48	10	49, 50	0	2
		▪ Deotomatisasi	52, 53	54, 55	4	52	55	2
		▪ Kecerdikan	56, 58, 61, 62, 63, 64, 65, 66, 67	57, 59, 60, 68	13	64	57	2
				34	34	68	15	7

\*) Pada instrumen final, nomor item *MC-WiSe-SF* diurutkan kembali menjadi 1-22 mengikuti urutan nomor aslinya.

Penimbang pakar menilai item-item *MC-WiSe* dengan skor penilaian yang disediakan yaitu: 1 (kurang terbaca/sesuai), 2 (cukup terbaca/baik), atau 3 (terbaca/baik). Hasil penilaian penimbang pakar menunjukkan bahwa item-item *MC-WiSe* memiliki keterbacaan pada kategori cukup baik, item sesuai dengan indikator, dan item sesuai dengan pola respon. Penimbang pakar juga memberikan masukan: (a) perlu menghapus kata-kata yang mengarah pada kecenderungan tertentu seperti “paling” dan “lebih”; (b) perlu menghapus kata negatif “tidak”; dan (c) beberapa item perlu disesuaikan dengan konteks BK untuk meningkatkan *face validity*-nya. *Pilot testing* dilakukan terhadap tiga orang mahasiswa BK (calon konselor) di Prodi BK UNJ. Hasil penilaian calon konselor menunjukkan

bahwa item-item *MC-WiSe* memiliki keterbacaan yang baik, item sesuai dengan indikator, dan item sesuai dengan pola respon.

**Tabel 3.8 Hasil Penilaian Penimbang Pakar terhadap *MC-WiSe***

Aspek Penilaian	Skor							
	P <sub>1</sub>	P <sub>2</sub>	P <sub>3</sub>	$\bar{X}$	CK <sub>1</sub>	CK <sub>2</sub>	CK <sub>3</sub>	$\bar{X}$
Keterbacaan	2	2	3	2.3	3	2	3	2.3
Kesesuaian item dengan indikator	3	3	2	2.7	3	3	3	3.0
Kesesuaian item dengan pola respon	3	3	3	3.0	3	3	3	3.0
Rerata Total	2.7	2.7	2.7	2.7	3.0	2.7	3.0	2.8

Keterangan: P adalah penimbang pakar dan CK adalah calon konselor.

Tahap kelima, melakukan uji coba. Uji coba dilakukan setelah *MC-WiSe* hasil uji validasi direvisi. Uji coba *MC-WiSe* dilakukan terhadap 517 calon konselor dari 11 Prodi BK di Indonesia.

Tahap keenam, menganalisis faktor, memfinalisasi item, dan menguji properti psikometrik. Analisis data hasil uji coba *MC-WiSe* berlandaskan pendekatan *Item Respon Theory (IRT)* menggunakan *Rasch Model v. 3.73* dan *Classical Test Theory (CTT)* sebagai pembandingan.

### 3) Pengujian Properti Psikometrik

#### a) Analisis Kelayakan Item

Gambaran level kelayakan item *MC-WiSe* dianalisis dari kriteria *item measure*, *item fit order*, dan *DIF* menggunakan Rasch Model. *Item Measure* digunakan untuk mengukur tingkat kesulitan item. Hasil pengujian menunjukkan item X<sub>35</sub> dengan + 1.10 logit menunjukkan item yang paling sukar disetujui oleh partisipan, sedangkan item X<sub>38</sub> dengan nilai -1.09 logit merupakan item yang paling mudah disetujui.

*Item Fit Order* digunakan untuk menguji item yang *fit* dan *misfit*. Hasil penelitian menemukan bahwa terdapat 31 dari 68 item yang *fit*. Keputusan ini didasarkan pada kriteria bahwa item dikatakan *fit* jika nilai *Infit* dan *Outfit Mnsq* berada berada pada rentang nilai yang dapat diterima ( $.5 Mnsq < 1.5$ ), *Outfit Zstd* ( $-2 < Zstd < +2$ ), *Pt-MCorr* ( $.4 < pt-MCorr .85$ ) (Bond & Fox, 2015; Dimitrov, 2012; Linacre, 2019; Sumintono & Widhiarso, 2014; 2015; Boone, Staver, & Yale, 2014). Dalam penelitian ini, kriteria *Outfit Zstd* tidak digunakan karena

ukuran sampel besar > 500 orang (Sumintono & Widhiarso, 2014; 2015). Item-item yang *fit*, yaitu: X<sub>1</sub>, X<sub>3</sub>, X<sub>5</sub>, X<sub>6</sub>, X<sub>7</sub>, X<sub>8</sub>, X<sub>11</sub>, X<sub>13</sub>, X<sub>14</sub>, X<sub>17</sub>, X<sub>25</sub>, X<sub>26</sub>, X<sub>27</sub>, X<sub>30</sub>, X<sub>37</sub>, X<sub>38</sub>, X<sub>41</sub>, X<sub>47</sub>, X<sub>48</sub>, X<sub>49</sub>, X<sub>50</sub>, X<sub>51</sub>, X<sub>52</sub>, X<sub>55</sub>, X<sub>57</sub>, X<sub>59</sub>, X<sub>61</sub>, X<sub>62</sub>, X<sub>64</sub>, X<sub>67</sub>, dan X<sub>68</sub>.

*DIF* digunakan untuk mendeteksi bias item dalam kategori partisipan tertentu. Item yang terdeteksi bias diketahui berdasarkan nilai probabilitas yang kurang dari 5% (.05) (Bond & Fox, 2015; Dimitrov, 2012; Linacre, 2019; Sumintono & Widhiarso, 2014; 2015). Hasil pengujian menampilkan 10 item mengalami bias terhadap jenis kelamin tertentu, yaitu: X<sub>19</sub> (.008), X<sub>22</sub> (.01), X<sub>28</sub> (.003), X<sub>32</sub> (.02), X<sub>33</sub> (.02), X<sub>35</sub> (.000), X<sub>37</sub> (.01), X<sub>53</sub> (.01), X<sub>54</sub> (.008), dan X<sub>61</sub> (.006). Sebanyak tujuh item mengalami bias terhadap usia tertentu, yaitu: X<sub>2</sub> (.007), X<sub>9</sub> (.009), X<sub>28</sub> (.001), X<sub>39</sub> (.02), X<sub>47</sub> (.04), X<sub>54</sub> (.02), X<sub>56</sub> (.001).

Berdasarkan ketiga kriteria ketepatan item diperoleh 28 dari 31 item yang *fit*. Untuk kepentingan penyeimbangan item pada setiap indikator *MC-WiSe* dari 28 item dipilih 22 item. Item final *MC-WiSe*, yaitu: X<sub>1</sub>, X<sub>6</sub>, X<sub>7</sub>, X<sub>11</sub>, X<sub>13</sub>, X<sub>17</sub>, X<sub>18</sub>, X<sub>19</sub>, X<sub>22</sub>, X<sub>25</sub>, X<sub>27</sub>, X<sub>30</sub>, X<sub>32</sub>, X<sub>37</sub>, X<sub>38</sub>, X<sub>41</sub>, X<sub>49</sub>, X<sub>50</sub>, X<sub>52</sub>, X<sub>55</sub>, X<sub>57</sub>, dan X<sub>64</sub>. Sebaran item pada dimensi dan indikatornya disajikan pada Tabel 3.7. Item X<sub>3</sub>, X<sub>5</sub>, X<sub>8</sub>, X<sub>14</sub>, X<sub>26</sub>, X<sub>48</sub>, X<sub>51</sub>, X<sub>59</sub>, X<sub>62</sub>, dan X<sub>68</sub> tidak dipilih walaupun *fit* karena indikatornya sudah terwakili oleh item lain. Item X<sub>18</sub>, X<sub>19</sub>, X<sub>22</sub>, X<sub>32</sub>, dan X<sub>41</sub> tetap dipilih untuk keseimbangan item pada dan keterwakilan indikator. Pertimbangan yang sama juga berlaku untuk item X<sub>38</sub> tetap dipilih walaupun dinilai paling mudah disetujui oleh responden karena memenuhi ketiga kriteria *fit*. Hal ini sesuai pendapat (Sumintono & Widhiarso, 2014; 2015) bahwa item yang hanya memenuhi salah satu kriteria *fit* tetap dapat digunakan jika diperlukan. Pertimbangan lainnya didasarkan hasil pengujian *item-total correlation* ( $r_{it}$ ) dari pendekatan *Classical Test Theory* (*CTT*) menggunakan *IBM SPSS 25.0 for Windows* bahwa item X<sub>18</sub>, X<sub>19</sub>, X<sub>22</sub>, X<sub>32</sub>, X<sub>38</sub>, dan X<sub>41</sub> memiliki indeks koefisien  $r_{it} \geq .300$ . Pakar psikometrik (Azwar, 2011; Drummond & Jones, 2010) mengatakan bahwa indeks koefisien  $r_{it}$  tersebut dapat diterima.

#### b) Analisis Abilitas Partisipan

Gambaran bukti level abilitas responden dianalisis dari kriteria *person measure*, dan *person fit order* menggunakan Rasch Model. Hasil pengujian *Person Measure* menampilkan responden nomor urut 498 (8LDPI) dengan + 1.82 logit cenderung memiliki kearifan yang tinggi karena lebih banyak menjawab sesuai dan sangat sesuai, sedangkan responden nomor urut 40 (0PDID) dengan nilai -.40 logit cenderung memiliki kearifan multibudaya yang rendah karena lebih banyak menjawab ke arah tidak sesuai dari item-item instrumen *MC-WiSe*.

#### c) Analisis Kualitas Instrumen

Gambaran bukti level kualitas instrumen dianalisis dari kriteria *unidimensionality*, *rating scale*, dan *reliability (test, item, dan person)* menggunakan *Rasch Model v 3.73*.

*Unidimensionality*. Aspek ini diukur menggunakan Rasch Model dengan analisis komponen utama (*Principle Component Analysis [PCA]*) dari residual yakni mengukur tingkat uniformitas dari instrumen dalam mengukur apa yang seharusnya diukur. Hasil pengukuran menampilkan *raw variance* sebesar 18.8% serta varians yang tidak dapat dijelaskan oleh instrumen sebesar 9.8%. Hal ini berarti bahwa persyaratan unidimensionalitas minimal 20% tidak terpenuhi dan varians yang tidak dapat dijelaskan oleh instrumen  $\leq 15\%$  terpenuhi (Bond & Fox, 2015; Dimitrov, 2012; Linacre, 2019; Sumintono & Widhiarso, 2014; 2015).

*Rating Scale*. Ini merupakan pengujian untuk memverifikasi pilihan peringkat yang digunakan dalam instrumen *MC-WiSe* dapat membingungkan partisipan atau tidak (Bond & Fox, 2015; Dimitrov, 2012; Linacre, 2019; Sumintono & Widhiarso, 2014; 2015). Hasil analisis Rasch Model menemukan bahwa rata-rata nilai pengamatan dimulai dari logit -2.67 untuk pilihan skor 1 (sangat tidak sesuai), logit -1.17 untuk pilihan skor 2 (tidak sesuai), logit -.08 untuk pilihan skor 3 (cukup sesuai), logit +1.13 untuk pilihan skor 4 (sesuai), dan logit +2.87 untuk pilihan skor 5 (sangat sesuai). Hasil senada juga ditampilkan oleh *Andrich Threshold* yang menguji ketepatan nilai politomi yang digunakan. Nilai tersebut menunjukkan pergerakan dari *NONE* ke negatif (-1.27, -.68) dan

terus mengarah ke positif (.37, 1.58) secara berurutan. Artinya, pilihan skala peringkat yang digunakan sudah valid dan tepat bagi partisipan.

*Reliability.* Tabel 3.9 Ringkasan Statistik menampilkan *person measure* = +.36 logit. Nilai rata-rata yang lebih tinggi dari logit .0 menunjukkan kecenderungan partisipan yang lebih banyak menjawab penting pada setiap item. Nilai *Alpha Cronbach* untuk mengukur reliabilitas tes yakni interaksi antara *person* dan item secara keseluruhan menunjukkan .82 yang termasuk bagus. Sementara itu, nilai *person reliability* sebesar .79 dan *item reliability* .99 sehingga dapat disimpulkan bahwa konsistensi jawaban partisipan termasuk cukup baik dan seperangkat tes termasuk istimewa (Sumintono & Widhiarso, 2014; 2015). Hasil senada ditampilkan oleh *Infit* dan *Outfit Mnsq Mnsq* untuk *person* dan item, nilai rata-ratanya secara berurutan 1.02 dan 1.01 dan 1.00 dan 1.01 yang mendekati nilai ideal 1.00. Demikian juga *Infit* dan *Outfit Zstd* menunjukkan nilai rata-rata *person* dan item sebesar -.5 dan -.5 serta -.2 dan .0 yang berarti baik karena mendekati .0.

**Tabel 3.9 Ringkasan Statistik MC-WiSe dan MC-WiSe-SF**

Kelayakan	Luaran	Indeks Hasil	
		MC-WiSe	MC-WiSe-SF
Item	Reliabilitas item	.99	.98
	Indeks pengelompokkan	8.41	8.2
	Strata pengelompokkan (H)	2.92	2.6
	Nilai logit tertinggi	+ 1.10 logit (X <sub>35</sub> )	+ .86 logit (Q <sub>7</sub> )
	Nilai logit terendah	-1.09 logit (X <sub>38</sub> )	-.98 logit (Q <sub>15</sub> )
Person	Reliabilitas partisipan	.79	.74
	Indeks pengelompokkan	1.94	1.7
	Strata pengelompokkan (H)	2.92	2.6
	Nilai logit tertinggi	+ 1.82 logit (498; 8LDPI)	+5.79 logit (177; 6PDIC)
	Nilai logit terendah	-.40 logit (40; OPDID)	-.70 logit (473; 3PRIH)
Instrumen	<i>Alpha Cronbach</i>	.82	.78
	<i>Raw variance explained by measures</i>	18.8	26.0
	<i>Unexplained variance in 1<sup>st</sup> contrast</i>	9.8	8.7
	<i>Unexecpected variance in 2<sup>nd</sup> contrast</i>	6.6	7.3
	<i>Unexecpected variance in 3<sup>nd</sup> contrast</i>	3.4	5.8
	<i>Unexecpected variance in 4<sup>nd</sup> contrast</i>	2.9	4.6
	<i>Unexecpected variance in 5<sup>nd</sup> contrast</i>	2.5	4.3

Nilai *separation* digunakan untuk mengelompokkan *person* dan item. Makin besar nilai *separation* menunjukkan makin bagus kualitas instrumen karena dapat mengidentifikasi kelompok responden dan kelompok item.

Persamaan yang digunakan yaitu:  $H = \frac{[(4 \times 1,94) + 1]}{3} = 2,92$  dibulatkan menjadi 3. Hasil ini bermakna terdapat tiga kelompok responden.

#### 4) Bentuk Akhir MC-WiSe

Dengan mempertimbangkan hasil pengujian berbagai kriteria kelayakan, maka *MC-WiSe* final yang digunakan dalam penelitian berjumlah 22 item untuk mengukur dua dimensi (kognitif serta afektif dan kesadaran) dan 11 indikator (metakognitif, ketajaman pandangan, penalaran dialektik, toleransi terhadap ambiguitas, pengentasan masalah, keterampilan *coping*, empati, kepedulian, mengenali emosi, deotomatisasi, dan kecerdikan). *The 3D-WS-12* diharapkan dapat menjadi alat asesmen kearifan multibudaya calon konselor yang lebih efisien dan valid. *MC-WiSe* merupakan alat ukur laporan diri berbentuk *summated rating* dengan skala Likert 5 jenjang yang ketentuannya disajikan pada Tabel 3.10. Kisi-kisi *MC-WiSe* disajikan pada Tabel 3.9 dan instrumen lengkapnya dapat diperiksa pada Lampiran C.

**Tabel 3.10. Pedoman Penskoran MC-WiSe-SF**

Aitem	Skor Alternatif Jawaban				
	Sangat Sesuai	Sesuai	Cukup Sesuai	Tidak Sesuai	Sangat Tidak Sesuai
Positif	5	4	3	2	1
Negatif	1	2	3	4	5

### c. Skala Ajektif Kearifan Konselor

#### 1) Prosedur Pengembangan

Data konstruk baru dimensi kearifan konselor dikumpulkan menggunakan Skala Ajektif Kearifan Konselor (SAjKK). SAjKK dikonstruksi sendiri oleh peneliti berdasarkan karakteristik-karakteristik kearifan hasil kajian meta-analisis dari beragam teori kearifan. SAjKK dikembangkan melalui prosedur berikut: (a) mengkaji literatur; (b) mengonseptualisasikan dan mengoperasionisasikan konstruk ukur; (c) mengidentifikasi indikator, menyusun item, dan menetapkan format respon; (d) menganalisis isi dan *pilot testing*, merevisi, dan mengadministrasikan aitem; (e) melakukan uji coba; (f) menganalisis faktor,

memfinalisasi item, dan menguji properti psikometrik (Heppner, Wampold, & Kivlighan, 2008).

Tahap pertama dan kedua, mengonseptualisasikan dan mengoperasionalkan konstruk ukur serta mengkaji literatur. SAjKK dikonstruksi berlandaskan pada dan teori-teori kearifan (di antaranya: Ardel, 2003; Ardel & Oh, 2010; Ardel, Pridgen, & Nutter-Pridgen, 2019; Baltes, Gluck & Kunzmann, 2005; Brown & Greene, 2006; Gluck et al., 2013; Hanna & Ottens, 1995; Hanna, Bemak, & Chung, 1999; Jeste et al., 2011; Kunzmann, 2019; Levenson, Jennings, & Aldwin, 2005; Webster, 2007; 2019; Yahya, 2003) dan teori lainnya yang telah dibahas di BAB II.

Tahap ketiga, mengidentifikasi indikator, menyusun item, dan menetapkan format respon. SAjKK mengukur 44 karakteristik kearifan, yaitu: religiusitas/spiritualitas, transendensi diri, metakognisi, kaya pengetahuan faktual, kaya pengetahuan prosedural, kaya pengetahuan situasional, pengaturan kognisi, ketajaman pandangan, penalaran dialektik, timbangan reflektif, toleransi terhadap ambiguitas, mengelola ketidakpastian, keterbukaan pikiran, relativisme nilai-nilai, penemuan dan pengentasan masalah, cinta belajar, pengetahuan diri, penghargaan diri, pengintegrasian diri, penyesuaian diri, dan ketenangan.

Pola respon yang dipilih berbentuk *descriptor rating* dengan Skala Likert lima jenjang yaitu: 1 = Sangat Tidak Penting; 2 = Tidak Penting; 3 = Cukup Penting; 4 = Penting; dan 5 = Sangat Penting menggambarkan diri saya.

Tahap keempat dan kelima, menganalisis isi dan *pilot testing*, merevisi, dan mengadministrasikan item. Uji coba dilakukan setelah SAjKK hasil uji validasi direvisi. Uji coba SAjKK dilakukan terhadap 517 calon konselor dari 11 Prodi BK di Indonesia.

Tahap keenam, menganalisis faktor, memfinalisasi item, dan menguji properti psikometrik. Analisis data hasil uji coba SAjKK berlandaskan pendekatan *Item Respon Theory (IRT)* menggunakan *Rasch Model v. 3.73* dan *Classical Test Theory (CTT)* sebagai pembanding.

### 3) Pengujian Properti Psikometrik

#### a) Analisis Kelayakan Item

Gambaran level kelayakan item SAjKK dianalisis dari kriteria *item measure*, *item fit order*, dan *DIF* menggunakan Rasch Model. *Item Measure* digunakan untuk mengukur tingkat kesulitan item. Hasil pengujian menunjukkan item X<sub>40</sub> dengan + 1.81 logit menunjukkan item yang paling sukar disetujui oleh partisipan, sedangkan item X<sub>27</sub> dengan nilai -1.38 logit merupakan item yang paling mudah disetujui.

*Item Fit Order* digunakan untuk menguji item yang *fit* dan *misfit*. Hasil penelitian menemukan bahwa terdapat 44 item yang *fit*. Keputusan ini didasarkan pada kriteria bahwa item dikatakan *fit* jika nilai *Infit* dan *Outfit Mnsq* berada pada rentang nilai yang dapat diterima ( $.5 Mnsq < 1.5$ ), *Outfit Zstd* ( $-2 < Zstd < +2$ ), *Pt-MCorr* ( $.4 < pt-MCorr .85$ ) (Bond & Fox, 2015; Dimitrov, 2012; Linacre, 2019; Sumintono & Widhiarso, 2014; 2015; Boone, Staver, & Yale, 2014). Dalam penelitian ini, kriteria *Outfit Zstd* tidak digunakan karena ukuran sampel besar > 500 orang (Sumintono & Widhiarso, 2014; 2015).

*DIF* digunakan untuk mendeteksi bias item dalam kategori partisipan tertentu. Item yang terdeteksi bias diketahui berdasarkan nilai probabilitas yang kurang dari 5% (.05) (Bond & Fox, 2015; Dimitrov, 2012; Linacre, 2019; Sumintono & Widhiarso, 2014; 2015). Hasil pengujian menampilkan seluruh item tidak mengalami bias terhadap jenis kelamin tertentu.

Berdasarkan ketiga kriteria ketepatan item diperoleh seluruh (44) item *fit*. Pertimbangan lainnya didasarkan hasil pengujian *item-total correlation* ( $r_{it}$ ) dari pendekatan *Classical Test Theory* (CTT) menggunakan *IBM SPSS 25.0 for Windows* bahwa seluruh (44) item juga memiliki indeks koefisien  $r_{it} \geq .300$ . Pakar psikometrik (Azwar, 2011; Drummond & Jones, 2010) mengatakan bahwa indeks koefisien  $r_{it}$  tersebut dapat diterima.

#### b) Analisis Abilitas Partisipan

Gambaran bukti level abilitas responden dianalisis dari kriteria *person measure*, *person fit order*, dan *item-person maps* menggunakan Rasch Model.

Hasil pengujian *Person Measure* menampilkan responden nomor urut 312 dengan + 7.76 logit cenderung memiliki kearifan yang tinggi karena lebih banyak menjawab sangat penting, sedangkan responden nomor urut 476 dengan nilai -1.01 logit cenderung memiliki kearifan multibudaya yang rendah karena lebih banyak menjawab ke arah tidak penting dari item-item instrumen SAjKK.

### c) Analisis Kualitas Instrumen

Gambaran bukti level kualitas instrumen dianalisis dari kriteria *unidimensionality*, *rating scale*, dan *reliability (test, item, dan person)* menggunakan *Rasch Model v 3.73*.

*Unidimensionality*. Aspek ini diukur menggunakan Rasch Model dengan analisis komponen utama (*Principle Component Analysis [PCA]*) dari residual yakni mengukur tingkat uniformitas dari instrumen dalam mengukur apa yang seharusnya diukur. Hasil pengukuran menampilkan *raw variance* sebesar 37.4% serta varians yang tidak dapat dijelaskan oleh instrumen sebesar 6.4%. Hal ini berarti bahwa persyaratan unidimensionalitas minimal 20% terpenuhi dan varians yang tidak dapat dijelaskan oleh instrumen  $\leq 15\%$  terpenuhi (Bond & Fox, 2015; Dimitrov, 2012; Linacre, 2019; Sumintono & Widhiarso, 2014; 2015).

*Rating Scale*. Ini merupakan pengujian untuk memverifikasi pilihan peringkat yang digunakan dalam instrumen SAjKK dapat membingungkan partisipan atau tidak (Bond & Fox, 2015; Dimitrov, 2012; Linacre, 2019; Sumintono & Widhiarso, 2014; 2015). Hasil analisis Rasch Model menemukan bahwa rata-rata nilai pengamatan dimulai dari logit -2.95 untuk pilihan skor 1 (sangat tidak penting), logit -1.60 untuk pilihan skor 2 (tidak penting), logit -.33 untuk pilihan skor 3 (cukup penting), logit +1.54 untuk pilihan skor 4 (penting), dan logit +3.63 untuk pilihan skor 5 (sangat penting). Hasil senada juga ditampilkan oleh *Andrich Threshold* yang menguji ketepatan nilai politomi yang digunakan. Nilai tersebut menunjukkan pergerakan dari *NONE* ke negatif (-1.34, -.65) dan terus mengarah ke positif (.56, 2.43) secara berurutan. Artinya, pilihan skala peringkat yang digunakan sudah valid dan tepat bagi partisipan.

*Reliability*. Tabel 3.11 *Ringkasan Statistik* menampilkan *person measure* = +.36 logit. Nilai rata-rata yang lebih tinggi dari logit .0 menunjukkan kecenderungan partisipan yang lebih banyak menjawab penting pada setiap item. Nilai *Alpha Cronbach* untuk mengukur reliabilitas tes yakni interaksi antara *person* dan item secara keseluruhan menunjukkan .82 yang termasuk bagus. Sementara itu, nilai *person reliability* sebesar .79 dan *item reliability* .99 sehingga dapat disimpulkan bahwa konsistensi jawaban partisipan termasuk cukup baik dan seperangkat tes termasuk istimewa (Sumintono & Widhiarso, 2014; 2015). Hasil senada ditampilkan oleh *Infit* dan *Outfit Mnsq Mnsq* untuk *person* dan item, nilai rata-ratanya secara berurutan 1.02 dan 1.01 dan 1.00 dan 1.01 yang mendekati nilai ideal 1.00. Demikian juga *Infit* dan *Outfit Zstd* menunjukkan nilai rata-rata *person* dan item sebesar -.5 dan -.5 serta -.2 dan .0 yang berarti baik karena mendekati .0.

**Tabel 3.11 Ringkasan Statistik SAjKK**

Kelayakan	Luaran	Indeks Hasil
Item	Reliabilitas item	.99
	Indeks pengelompokkan	9.97
	Nilai logit tertinggi	+ 1.81 logit (X <sub>40</sub> )
	Nilai logit terendah	-1.38 logit (X <sub>27</sub> )
Person	Reliabilitas partisipan	.91
	Indeks pengelompokkan	3.13
	Nilai logit tertinggi	+ 7.76 logit (312)
	Nilai logit terendah	-1.01 logit (476)
Instrumen	<i>Alpha Cronbach</i>	.94
	<i>Raw variance explained by measures</i>	37.4
	<i>Unexplained variance in 1<sup>st</sup> contrast</i>	6.4
	<i>Unexecpected variance in 2<sup>nd</sup> contrast</i>	3.7
	<i>Unexecpected variance in 3<sup>rd</sup> contrast</i>	3.1
	<i>Unexecpected variance in 4<sup>nd</sup> contrast</i>	2.7
	<i>Unexecpected variance in 5<sup>nd</sup> contrast</i>	2.3

#### 4) Bentuk Akhir SAjKK

Dengan mempertimbangkan hasil pengujian berbagai kriteria kelayakan, maka SAjKK final yang digunakan dalam penelitian berjumlah 44 deskriptor karakteristik kearifan konselor. SAjKK merupakan alat ukur laporan diri berbentuk *descriptor rating* dengan skala Likert 5 jenjang yang ketentuannya disajikan pada Tabel 3.12. instrumen lengkapnya dapat diperiksa pada Lampiran C.

**Tabel 3.12. Pedoman Penskoran SAjKK**

Aitem	Skor Alternatif Jawaban				
	Sangat Penting	Penting	Cukup Penting	Tidak Penting	Sangat Tidak Penting
Positif	5	4	3	2	1
Negatif	1	2	3	4	5

#### d. Skala Anteseden Kearifan Konselor

##### 1) *Prosedur Pengembangan*

Data anteseden kearifan konselor dikumpulkan menggunakan Skala Anteseden Kearifan Konselor (SAnKK). SAnKK dikonstruksi sendiri oleh peneliti melalui prosedur berikut: (a) mengkaji literatur; (b) mengonseptualisasikan dan mengoperasionisasikan konstruk ukur; (c) mengidentifikasi indikator, menyusun item, dan menetapkan format respon; (d) menganalisis isi dan *pilot testing*, merevisi, dan mengadministrasikan aitem; (e) melakukan uji coba; (f) menganalisis faktor, memfinalisasi item, dan menguji properti psikometrik (Heppner, Wampold, & Kivlighan, 2008).

Tahap pertama dan kedua, mengonseptualisasikan dan mengoperasionisasikan konstruk ukur serta mengkaji literatur. SAnKK dikonstruksi berlandaskan pada anteseden kearifan dari *Berlin Wisdom Paradigm* (di antaranya Baltes, Gluck, & Kunzmann, 2002; Baltes & Staudinger, 2000; Kunzmann & Baltes, 2005; Stange & Kunzmann, 2008; Staudinger, 1999; Staudinger & Gluck, 2011).

Tahap ketiga, mengidentifikasi indikator, menyusun item, dan menetapkan format respon. SAnKK mengukur tiga faktor dan 28 deskriptor, yaitu: Faktor konteks-fasilitatif, meliputi: agama, usia, etnis (budaya), interaksi sosial, dan konteks pendidikan. Faktor kepakaran, meliputi: mentor/model peran, praktik berkelanjutan, pengalaman hidup, dan pelatihan profesional, kompetensi konseling, belajar filsafat, belajar berpikir dialektik, belajar dengan metode Dialog Socratic, belajar melalui pengalaman (ELT), belajar melalui teknik kesadaran diri, latihan meditasi, transformasi pengalaman (internalisasi, integrasi, dan refleksi pengalaman), latihan meditasi. Faktor pribadi, meliputi: kecerdasan, kepribadian, motivasi, jenis kelamin, orientasi nilai, pandangan hidup, dan timbangan moral.

Herdi, 2020

PEDAGOGIK UNTUK MENTRANSFORMASI KEARIFAN DAN ALIANSI KONSELING CALON KONSELOR  
Universitas Pendidikan Indonesia repository.upi.edu perpustakaan.upi.edu

Pola respon yang dipilih berbentuk Skala Likert lima jenjang yaitu: 1 = Sangat Tidak Relevan; 2 = Tidak Relevan; 3 = Cukup Relevan; 4 = Relevan; dan 5 = Sangat Relevan. Hasil identifikasi indikator kearifan calon konselor multibudaya disajikan pada Tabel 3.13.

Tahap keempat dan kelima, menganalisis isi dan *pilot testing*, merevisi, dan mengadministrasikan item. Uji coba dilakukan setelah SAnKK hasil uji validasi direvisi. Uji coba SAnKK dilakukan terhadap 517 calon konselor dari 11 Prodi BK di Indonesia. Tahap keenam, menganalisis faktor, memfinalisasi item, dan menguji properti psikometrik. Analisis data hasil uji coba SAnKK berlandaskan pendekatan *Item Respon Theory (IRT)* menggunakan *Rasch Model* v. 3.73 dan *Classical Test Theory (CTT)* sebagai pembanding.

**Tabel 3.13 Kisi-kisi Skala Anteseden Kearifan Konselor**

Variabel	Faktor	Indikator	No. Aitem	$\Sigma$
Anteseden Kearifan	Konteks-Fasilitatif	Agama	1	6
		Usia	2	
		Etnis	4	
		Konteks pendidikan (tingkat dan lembaga pendidikan)	11, 12	
	Kepakaran	Interaksi sosial	24	15
		Pelatihan profesional	13	
		Praktik berkelanjutan	14	
		Kompetensi konseling	15	
		Belajar filsafat	16	
		Belajar berpikir dialektik	17	
		Belajar dengan metode Dialog Socratic	18	
		Belajar melalui pengalaman (ELT)	19	
		Belajar melalui teknik kesadaran diri	20	
		Latihan meditasi	21	
		Pengalaman hidup	22	
		Model peran	23	
		Transformasi pengalaman	25	
	Belajar menghadapi ketidakpastian	26		
	Bimbingan spiritual	27		
	Menghadiri upacara kematian	28		
	Pribadi	Jenis kelamin	3	7
		Kecerdasan	5	
		Kepribadian	6	
		Motivasi	7	
		Orientasi nilai	8	
		Pandangan hidup	9	
		Timbangan moral	10	
		Jumlah		28

## 2) Pengujian Properti Psikometrik

### a) Analisis Kelayakan Item

Gambaran level kelayakan item *MC-WiSe* dianalisis dari kriteria *item measure*, *item fit order*, dan *DIF* menggunakan Rasch Model. *Item Measure* digunakan untuk mengukur tingkat kesulitan item. Hasil pengujian menunjukkan item  $X_{28}$  dengan + 2.17 logit menunjukkan item yang paling sukar disetujui oleh partisipan, sedangkan item  $X_6$  dengan nilai -1.35 logit merupakan item yang paling mudah disetujui.

*Item Fit Order* digunakan untuk menguji item yang *fit* dan *misfit*. Hasil penelitian menemukan bahwa terdapat 27 item dari 28 item yang *fit*. Keputusan ini didasarkan pada kriteria bahwa item dikatakan *fit* jika nilai *Infit* dan *Outfit Mnsq* berada pada rentang nilai yang dapat diterima ( $.5 Mnsq < 1.5$ ), *Outfit Zstd* ( $-2 < Zstd < +2$ ), *Pt-MCorr* ( $.4 < pt-MCorr .85$ ) (Bond & Fox, 2015; Linacre, 2019; Sumintono & Widhiarso, 2014; 2015; Boone, Staver, & Yale, 2014). Dalam penelitian ini, kriteria *Outfit Zstd* tidak digunakan karena ukuran sampel besar > 500 orang (Sumintono & Widhiarso, 2014; 2015). Item yang *misfit*, yaitu:  $X_{28}$ .

*DIF* digunakan untuk mendeteksi bias item dalam kategori partisipan tertentu. Item yang terdeteksi bias diketahui berdasarkan nilai probabilitas yang kurang dari 5% (.05) (Bond & Fox, 2015; Dimitrov, 2012; Linacre, 2019; Sumintono & Widhiarso, 2014; 2015). Hasil pengujian menampilkan 2 item yang mengalami bias terhadap jenis kelamin tertentu, yaitu:  $X_{11}$  (.0278) dan  $X_{14}$  (.0169). Sebanyak satu item mengalami bias terhadap usia tertentu, yaitu:  $X_{26}$  (.0416). Berdasarkan ketiga kriteria ketepatan item diperoleh 27 dari 28 item yang *fit*. Sebaran item pada dimensi dan indikatornya disajikan pada Tabel 3.13. Pertimbangan lainnya didasarkan hasil pengujian *item-total correlation* ( $r_{it}$ ) dari pendekatan *Classical Test Theory (CTT)* menggunakan *IBM SPSS 25.0 for Windows* bahwa seluruh item memiliki indeks koefisien  $r_{it} \geq .300$ . Pakar psikometrik (Azwar, 2011; Drummond & Jones, 2010) mengatakan bahwa indeks koefisien  $r_{it}$  tersebut dapat diterima.

#### b) Analisis Abilitas Partisipan

Gambaran bukti level abilitas responden dianalisis dari kriteria *person measure*, *person fit order*, dan *item-person maps* menggunakan Rasch Model. Hasil pengujian *Person Measure* menampilkan responden nomor urut 419 (9PDIH) dengan + 7.29 logit cenderung menilai sangat relevan dan relevan deskriptor, sedangkan responden nomor urut 487 (7LDKI) dengan nilai -.46 logit cenderung menilai sangat tidak relevan dan tidak relevan deskriptor yang dinilai. *Person Fit Order* digunakan untuk menguji responden yang *fit* dan *misfit*. Kriteria yang digunakan sama dengan analisis *item fit order*. Berdasarkan kriteria ini, ditemukan 84 dari 517 responden yang *misfit*.

#### c) Analisis Kualitas Instrumen

Gambaran bukti level kualitas instrumen dianalisis dari kriteria *unidimensionality*, *rating scale*, dan *reliability (test, item, dan person)* menggunakan *Rasch Model v.3.73*

*Unidimensionality*. Aspek ini diukur menggunakan Rasch Model dengan analisis komponen utama (*Principle Component Analysis [PCA]*) dari residual yakni mengukur tingkat uniformitas dari instrumen dalam mengukur apa yang seharusnya diukur. Hasil pengukuran menampilkan *raw variance* sebesar 45.1% serta varians yang tidak dapat dijelaskan oleh instrumen sebesar 6.5%. Hal ini berarti bahwa persyaratan unidimensionalitas minimal 20% tidak terpenuhi dan varians yang tidak dapat dijelaskan oleh instrumen  $\leq 15\%$  terpenuhi (Bond & Fox, 2015; Dimitrov, 2012; Linacre, 2019; Sumintono & Widhiarso, 2014; 2015).

*Rating Scale*. Ini merupakan pengujian untuk memverifikasi pilihan peringkat yang digunakan dalam instrumen SAnKK dapat membingungkan partisipan atau tidak (Bond & Fox, 2015; Dimitrov, 2012; Linacre, 2019; Sumintono & Widhiarso, 2014; 2015). Hasil analisis Rasch Model menemukan bahwa rata-rata nilai pengamatan dimulai dari logit -3.17 untuk pilihan skor 1 (sangat tidak relevan), logit -1.54 untuk pilihan skor 2 (tidak relevan), logit -.17 untuk pilihan skor 3 (cukup relevan), logit +1.51 untuk pilihan skor 4 (relevan),

dan logit +3.54 untuk pilihan skor 5 (sangat relevan). Hasil senada juga ditampilkan oleh *Andrich Threshold* yang menguji ketepatan nilai politomi yang digunakan. Nilai tersebut menunjukkan pergerakan dari *NONE* ke negatif (-1.82, -1.04) dan terus mengarah ke positif (.53, 2.33) secara berurutan. Artinya, pilihan skala peringkat yang digunakan sudah valid dan tepat bagi partisipan.

*Reliability*. Tabel 3.14 *Ringkasan Statistik* menampilkan *person measure* = +1.72 logit. Nilai rata-rata yang lebih tinggi dari logit .0 menunjukkan kecenderungan partisipan yang lebih banyak menjawab penting pada setiap item. Nilai *Alpha Cronbach* untuk mengukur reliabilitas tes yakni interaksi antara *person* dan item secara keseluruhan menunjukkan .91 yang termasuk istimewa. Sementara itu, nilai *person reliability* sebesar .89 dan *item reliability* .99 sehingga dapat disimpulkan bahwa konsistensi jawaban partisipan termasuk baik dan seperangkat tes termasuk istimewa (Sumintono & Widhiarso, 2014; 2015). Hasil senada ditampilkan oleh *Infit* dan *Outfit Mnsq Mnsq* untuk *person* dan item, nilai rata-ratanya secara berurutan 1.03 dan .99 dan .99 dan .99 yang mendekati nilai ideal 1.00. Demikian juga *Infit* dan *Outfit Zstd* menunjukkan nilai rata-rata *person* dan item sebesar -.1 dan -.2 serta -.3 dan -.1 yang berarti baik karena mendekati .0.

**Tabel 3.14 Ringkasan Statistik SAnKK**

Kelayakan	Luaran	Indeks Hasil
Item	Reliabilitas item	.99
	Indeks pengelompokkan	12.85
	Strata pengelompokkan (H)	2.92
	Nilai logit tertinggi	+ 2.17 logit (X <sub>28</sub> )
	Nilai logit terendah	-1.35 logit (X <sub>6</sub> )
Person	Reliabilitas partisipan	.89
	Indeks pengelompokkan	2.85
	Strata pengelompokkan (H)	2.92
	Nilai logit tertinggi	+ 7.29 logit (419; 9PDIH)
	Nilai logit terendah	-.46 logit (487; 7LDK1)
Instrumen	<i>Alpha Cronbach</i>	.91
	<i>Raw variance explained by measures</i>	45.1
	<i>Unexplained variance in 1<sup>st</sup> contrast</i>	6.5
	<i>Unexecpected variance in 2<sup>nd</sup> contrast</i>	4.5
	<i>Unexecpected variance in 3<sup>rd</sup> contrast</i>	3.9

Nilai *separation* digunakan untuk mengelompokkan *person* dan item. Makin besar nilai *separation* menunjukkan makin bagus kualitas instrumen

karena dapat mengidentifikasi kelompok responden dan kelompok item. Persamaan yang digunakan yaitu:  $H = \frac{[(4 \times 1.94) + 1]}{3} = 2.92$  dibulatkan menjadi 3. Hasil ini bermakna terdapat tiga kelompok responden.

#### e. *The Working Alliance Inventory-Therapist*

##### 1) *Instrumen yang Diadaptasikan*

Data kualitas aliansi konseling calon konselor dikumpulkan menggunakan *Working Alliance Inventory-Therapist-Short Form (WAI-T-SF)* versi calon konselor Indonesia (Herdi, Kartadinata, & Taufiq, 2019e) hasil adaptasi dari *WAI-T* yang dikembangkan oleh Horvath (1981; (Horvath & Greenberg, 1986; 1989). *WAI-T* dipilih dan digunakan untuk mengukur kualitas aliansi konseling calon konselor didasarkan pada pertimbangan-pertimbangan berikut: (a) memiliki konstruk teoretis yang kuat; (b) memiliki properti psikometrik yang baik; (c) banyak dirujuk dan digunakan oleh peneliti lain; dan (d) ketersediaan dan kemudahan dalam perizinan untuk penggunaannya. (Mallinckrodt & Tekie, 2015) juga berpendapat bahwa seluruh versi *WAI* berkontribusi besar terhadap penelitian di bidang konseling dan psikoterapi.

*WAI-T* ditujukan untuk mengukur kualitas aliansi konseling berlandaskan pada teori tiga dimensi aliansi dari (Bordin, 1979). *WAI-T* versi asli dan lengkap berjumlah 36 item dimana setiap dimensi (*bond*, *tasks*, dan *goals*) berjumlah 12 item. *WAI-T* berbentuk laporan diri menggunakan skala Likert 7 jenjang dari 1 = tidak pernah sampai dengan 7 = selalu.

Penelitian menunjukkan *WAI-T* memiliki properti psikometrik yang baik, seperti: memiliki validitas konten yang baik menurut penimbangan pakar dan analisis menggunakan metode *multitrait-multimethod* (Horvath, 1994), berkorelasi positif signifikan dengan beragam konstruk luaran konseling (Horvath & Greenberg, 1986; 1989), serta konsistensi internal yang tinggi ( $r = .87$  secara keseluruhan dan  $r = .68 - .87$  untuk setiap dimensi *WAI-T*).

Kokotovic and Tracey (1989) mengembangkan *Working Alliance Inventory-Short Form (WAI-S)*. *WAI-S* berjumlah 12 item, dimana setiap dimensi

(*bond*, *goals*, dan *tasks*) berjumlah empat item. WAI-S berbentuk skala Likert 7 jenjang, mulai dari 1 = tidak pernah sampai dengan 7 = selalu. Hasil CFA menunjukkan WAI-S layak untuk mengukur satu dimensi aliansi maupun tiga dimensi (*bond*, *goals*, dan *tasks*) sesuai dengan teori dari Bordin dan WAI-T versi asli. Hatcher and Gillaspay (2006) mengembangkan *Working Alliance Inventory-Short Form Revised (WAI-SR)*. WAI-SR berjumlah 12 item, dimana setiap dimensi (*bond*, *goals*, dan *tasks*) berjumlah empat item. WAI-SR berbentuk skala Likert 5 jenjang, mulai dari 1 = tidak pernah sampai 5 = selalu. Hasil CFA menunjukkan WAI-SR terkonfirmasi layak untuk mengukur tiga dimensi aliansi konseling (*bond*, *goals*, dan *tasks*). WAI-SR konsistensi internal yang tinggi dengan  $\alpha = .91$  untuk keseluruhan dan .81 - .90 untuk setiap dimensinya (Hatcher & Gillaspay, 2006; Murder, Wilmers, Leonhart, Linster, & Barth, 2010; Herdi, Kartadinata, & Taufiq, 2019e; 2019h), reliabilitas *test-retest* yang tinggi dengan  $r = .93$  (Murder, Wilmers, Leonhart, Linster, & Barth, 2010; Hatcher & Gillaspay, 2006). Namun, kajian terkini dari Falkenstrom, Hatcher, Skjulsvik, Larsson, and Holmqvist (2015) terkonfirmasi bahwa WAI-SR versi 6 item hanya untuk 1 dimensi dan 2 dua dimensi (*bond* dan gabungan antara *goals* dan *tasks*). Karena itu, disarankan untuk menggunakan skor rerata WAI-SR daripada per subskala.

**Tabel 3.15 Kisi-kisi WAI-T dan WAI-T-SF**

Var	Dimensi	WAI-T			WAI-T-SF*)		
		+	-	$\Sigma$	+	-	$\Sigma$
Aliansi Konseling	<i>Bonds:</i>	5, 8,	1, 20,	12	17, 19,	-	4
	Kelekatan emosional antara konseli dan konselor	17, 19, 21, 23, 26, 28, 36	29		23, 26		
	<i>Tasks:</i>	2, 4,	7, 11,	12	4, 16,	-	4
	Kesepakatan antara konseli dan konselor dalam menetapkan tugas yang diperlukan untuk mencapai tujuan konseling	13, 16, 18, 24, 35	15, 31, 33		18, 24		
	<i>Goals:</i>	6, 14,	3, 9,	12	14, 22,	-	4
	Kesepakatan antara konseli dan konselor dalam menetapkan tujuan konseling	22, 25, 30, 32	10, 12, 27, 34		25, 30		
	Jumlah	22	14	36	12	0	12

\*) Nomor item WAI-T-SF diurutkan kembali menjadi B<sub>1</sub>-B<sub>4</sub>, T<sub>1</sub>-T<sub>4</sub>, dan G<sub>1</sub>-G<sub>4</sub> mengikuti urutan pada nomor aslinya.

Herdi, 2020

PEDAGOGIK UNTUK MENTRANSFORMASI KEARIFAN DAN ALIANSI KONSELING CALON KONSELOR  
Universitas Pendidikan Indonesia repository.upi.edu perpustakaan.upi.edu

Kajian lainnya menunjukkan bahwa *WAI-T* memiliki validitas konvergen yang baik dengan *Helping Alliance Questionnaire* dengan  $r = .64 - .74$  (Murder, Wilmers, Leonhart, Linster, & Barth, 2010; Hatcher & Gillaspy, 2006) dan *CALPAS* dengan  $r = .80$  (Hatcher & Gillaspy, 2006). *WAI-T* juga memiliki validitas prediktif yang baik pada kategori moderat ( $r = .24$ ) terhadap luaran konseling dari beragam adegan dan pendekatan konseling (Horvath, 1994; Horvath & Greenberg, 1989; Howard, Turner, Olkin, & Mohr, 2006; Martin, Garske, & Davis, 2000; Safran & Wallner, 1991). Kesimpulannya, *WAI-T-SF* merupakan instrumen aliansi yang valid dan reliabel, serta digunakan secara luas dalam penelitian di berbagai bidang, populasi, dan pendekatan konseling (Araujo, Filho, Ferreira, & Pinto, 2017; Falkenstrom, Hatcher, Skjulsvik, Larsson, and Holmqvist, 2015; Hatcher, Lindqvist, & Falkenstrom; 2019; Paap & Dijkstra, 2017; Paap, Schrier, & Dijkstra, 2019; Shu & Yu, 2017; Takasaki, Miki, & Hall, 2018). *WAI-T-SF* dan *WAI* versi lainnya juga telah diterjemahkan dan diadaptasi ke lebih dari 18 bahasa di dunia untuk berbagai kepentingan (<http://wai.profhorvath.com>, 2018).

## 2) *Prosedur Pengadaptasian*

Pengadaptasian *WAI-T* diawali dengan perizinan untuk menggunakannya kepada Prof. (Emiritus) Adam O. Horvath dari Simon Fraser University, Kanada sebagai pengembang dan pemilik hak ciptanya. Prosedur pengadaptasian *WAI-T* mengikuti pedoman umum pengadaptasian instrumen dari (Brislin, 1986; *International Test Commission*, 2016; Bartram, Hambleton, Gregoire, & Muniz, 2018) yaitu: “penerjemahan dan penerjemahan kembali oleh penerjemah profesional yang berbeda.” Prosedur pengadaptasian *WAI-T* secara rinci dalam penelitian ini adalah sebagai berikut.

Tahap pertama, menerjemahkan item. Penerjemahan item *WAI-T* dari bahasa Inggris ke bahasa Indonesia dilakukan oleh pakar Bahasa Inggris dan Linguistik.

Tahap kedua, menyunting dan mengintegrasikan hasil terjemahan. Hasil terjemahan dari kedua penerjemah disunting dan diintegrasikan ke dalam rumusan

saduran sementara. Kegiatan tahap ini dilakukan sendiri oleh peneliti. Tujuan tahap ini adalah mendapatkan rumusan terjemahan yang tepat dari dua penerjemah.

Tahap ketiga, menerjemahkan hasil saduran sementara. Hasil saduran pada tahap ketiga, diterjemahkan kembali ke dalam bahasa Inggris. Penerjemahan kembali dimaksudkan untuk menguji kecocokan isi antara hasil terjemahan dengan bahasa aslinya. Proses penerjemahan hasil saduran dilakukan oleh pakar bahasa Inggris dan Linguistik yang berbeda dari penerjemah sebelumnya dari Balai Bahasa UPI.

Tahap keempat, menyunting dan mengintegrasikan hasil terjemahan pertama dan kedua. Tahap ini dilakukan sendiri oleh peneliti. *WAI-T* yang digunakan dalam penelitian ini sedikit mengalami beberapa perubahan istilah dari versi aslinya. Perubahan yang dimaksud sebagai berikut. Pertama, istilah “klien” dan “terapis” diubah menjadi “konseli” dan “konselor”. Kedua, istilah “terapi” diubah menjadi “konseling”. Ketiga, pada *WAI-T* kolom pengisian nama konseli ditiadakan dan diganti dengan konseli. Hal ini dilakukan agar waktu yang digunakan dalam mengisi lebih efektif. Misalnya, pada pernyataan di *WAI-T* versi asli “*Saya merasa tidak nyaman dengan \_\_\_\_\_.*” diganti menjadi “*Saya merasa tidak nyaman dengan konseli yang saya tangani.*” Hasil saduran akhir disiapkan untuk tahap selanjutnya.

Tahap kelima, mengujicobakan *WAI-T*. Kegiatan uji coba dilakukan terhadap 517 calon konselor Indonesia. Analisis data hasil uji coba berlandaskan pendekatan *Item Respon Theory (IRT)* menggunakan *Rasch Model v. 3.73* dan *Classical Test Theory (CTT)* sebagai pembanding.

### **3) Pengujian Properti Psikometrik**

#### **a) Analisis Kelayakan Item**

Gambaran bukti level kelayakan item *WAI-T* dianalisis menggunakan kriteria *item measure*, *item fit order*, dan *DIF*.

*Item Measure* digunakan untuk mengukur tingkat kesulitan item. Hasil pengujian menunjukkan item item G<sub>3</sub> dengan + 1.11 logit menunjukkan item yang

paling sukar disetujui oleh partisipan, sedangkan item B<sub>8</sub> dengan nilai  $-.56$  logit merupakan item yang paling mudah disetujui.

*Item Fit Order* digunakan untuk menguji item yang *fit* dan *misfit*. Hasil penelitian menemukan bahwa terdapat 28 dari 36 item yang *fit*. Keputusan ini didasarkan pada kriteria bahwa item dikatakan *fit* jika nilai *Infit* dan *Outfit Mnsq* berada pada rentang nilai yang dapat diterima ( $.5 Mnsq < 1.5$ ), *Outfit Zstd* ( $-2 < Zstd < +2$ ), *Pt-MCorr* ( $.4 < pt-MCorr .85$ ) (Bond & Fox, 2015; Dimitrov, 2012; Linacre, 2019; Sumintono & Widhiarso, 2014; Sumintono & Widhiarso, 2015; Boone, Staver, & Yale, 2014). Dalam penelitian ini, kriteria *Outfit Zstd* tidak digunakan karena ukuran sampel besar  $> 500$  orang (Sumintono & Widhiarso, 2014; 2015). Item-item yang *misfit*, yaitu: B<sub>1</sub>, B<sub>6</sub>, T<sub>6</sub>, T<sub>11</sub>, G<sub>1</sub>, G<sub>3</sub>, G<sub>4</sub>, dan G<sub>9</sub>.

*DIF* digunakan untuk mendeteksi bias item dalam kategori partisipan tertentu. Item yang terdeteksi bias diketahui berdasarkan nilai probabilitas yang kurang dari 5% atau  $.05$  (Bond & Fox, 2015; Dimitrov, 2012; Linacre, 2019; Sumintono & Widhiarso, 2014; 2015). Sembilan item mengalami bias terhadap jenis kelamin tertentu, yaitu: B<sub>1</sub> (.027), B<sub>3</sub> (.004), B<sub>4</sub> (.004), B<sub>10</sub> (.002), B<sub>11</sub> (.001), T<sub>3</sub> (.006), T<sub>6</sub> (.000), G<sub>2</sub> (.017), dan G<sub>12</sub> (.004). Satu item mengalami bias terhadap usia tertentu, yaitu: T<sub>11</sub> (.003). Dua item mengalami bias terhadap afiliasi agama tertentu, yaitu: G<sub>3</sub> (.000), G<sub>7</sub> (.04), dan T<sub>11</sub> (.000).

Sebagai pembandingan, kelayakan item juga diuji dengan *item-total correlation* ( $r_{it}$ ) dari pendekatan *Classical Test Theory* (CTT) menggunakan *IBM SPSS 25.0 for Windows*. Hasil pengujian menemukan 32 dari 36 item yang memiliki dan memenuhi persyaratan indeks koefisien minimal  $r_{it} \geq .300$  (Azwar, 2011; Drummond & Jones, 2010). Adapun keempat item, yaitu: G<sub>1</sub>, G<sub>3</sub>, G<sub>4</sub>, dan G<sub>9</sub> tidak memenuhi kriteria minimal yang dipersyaratkan oleh pakar psikometrik tersebut. Hasil analisis kelayakan item *3D-WS* menggunakan *Rasch Model 3.73* dan *IBM SPSS v. 25.0* terlampir pada Lampiran D.

#### b) Analisis Abilitas Partisipan

Gambaran bukti level abilitas responden dianalisis dari kriteria *person measure* dan *person fit order* menggunakan Rasch Model. Hasil pengujian *Person Measure* menampilkan responden nomor urut 436 (6PDIH) dengan + 2.40 logit cenderung memiliki aliansi konseling yang tinggi karena lebih banyak menjawab sesuai dan sangat sesuai, sedangkan responden nomor urut 227 (7PDIC) dengan nilai -.81 logit cenderung memiliki kearifan pribadi yang rendah karena lebih banyak menjawab ke arah tidak sesuai dari item-item *WAI-T*.

#### c) Analisis Kualitas Instrumen

Gambaran bukti level kualitas instrumen dianalisis dari kriteria *unidimensionality*, *rating scale*, dan *reliability (test, item, dan person)* menggunakan *Rasch Model v. 3.73*.

*Unidimensionality* diukur menggunakan analisis komponen utama (*Principle Component Analysis [PCA]*) dari residual yakni mengukur tingkat uniformitas dari instrumen dalam mengukur apa yang seharusnya diukur (Linacre, 2019; Sumintono & Widhiarso, 2014; 2015). Hasil pengukuran menampilkan *raw variance* sebesar 30.8% serta varians tertinggi yang tidak dapat dijelaskan oleh instrumen sebesar 12.3%. Hal ini berarti bahwa persyaratan unidimensionalitas minimal 20% terpenuhi dan tidak ada varians yang tidak dapat dijelaskan oleh instrumen  $\geq 15\%$  (Bond & Fox, 2015; Dimitrov, 2012; Linacre, 2019; Sumintono & Widhiarso, 2014; 2015).

*Rating Scale*. Ini merupakan pengujian untuk memverifikasi pilihan peringkat yang digunakan dalam instrumen dapat membingungkan partisipan atau tidak (Bond & Fox, 2015; Dimitrov, 2012; Linacre, 2019; Sumintono & Widhiarso, 2014; 2015). Hasil analisis Rasch Model menemukan bahwa rata-rata nilai pengamatan dimulai dari logit -2.29 untuk pilihan skor 1 (tidak pernah), logit -1.17 untuk pilihan skor 2 (jarang), logit -.56 untuk pilihan skor 3 (agak jarang), logit -.07 untuk pilihan skor 4 (agak sering), dan logit +.46 untuk pilihan skor 5 (sering), logit +1.21 untuk pilihan 6 (sangat sering), dan +2.52 untuk pilihan 7 (selalu). Hasil senada juga ditampilkan oleh *Andrich Threshold* yang menguji

ketepatan nilai politomi yang digunakan. Nilai tersebut menunjukkan pergerakan dari *NONE* ke negatif (-.70, -.73, -.43) dan terus mengarah ke positif (.05, .78, 1.03) secara berurutan. Artinya, pilihan skala peringkat yang digunakan sudah valid dan tepat bagi partisipan.

*Reliability*. Tabel 3.16 Ringkasan Statistik menampilkan *person measure* = +.55 logit. Nilai rata-rata yang lebih tinggi dari logit .0 menunjukkan kecenderungan partisipan yang lebih banyak menjawab penting pada setiap item. Nilai *Alpha Cronbach* untuk mengukur reliabilitas tes yakni interaksi antara *person* dan item secara keseluruhan menunjukkan .90 yang termasuk bagus sekali. Sementara itu, nilai *person reliability* sebesar .87 dan *item reliability* .99 sehingga dapat disimpulkan bahwa konsistensi jawaban partisipan termasuk bagus dan aitem-atem dalam instrumen *WAI-T* termasuk istimewa (Sumintono & Widhiarso, 2014; 2015). Hasil senada ditampilkan oleh *Infit Mnsq* dan *Outfit Mnsq* untuk tabel *person* dan item, nilai rata-ratanya secara berurutan 1.11 dan 1.06 serta 1.01 dan 1.06 yang mendekati nilai ideal 1.00. Demikian juga *Infit Zstd* dan *Outfit Zstd* menunjukkan nilai rata-rata pada tabel *person* dan item sebesar -.1 dan -.2 serta -.6 dan .0 yang berarti baik karena mendekati .0.

**Tabel 3.16 Ringkasan Statistik *WAI-T* dan *WAI-T-SF***

Kelayakan	Luaran	Indeks Hasil	
		<i>WAI-T</i>	<i>WAI-T-SF</i>
Item	Reliabilitas item	.99	.96
	Indeks pengelompokkan	8.69	4.80
	Strata pengelompokkan (H)	3.8	3.15
	Nilai logit tertinggi	+ 1.11 logit (G <sub>3</sub> )	+ .45 logit (G <sub>4</sub> )
	Nilai logit terendah	-.56 logit (B <sub>8</sub> )	-.48 logit (B <sub>3</sub> )
Person	Reliabilitas partisipan	.87	.88
	Indeks pengelompokkan	2.60	2.67
	Strata pengelompokkan (H)	3.8	
	Nilai logit tertinggi	+ 2.40 (436; 6PDIH)	+6.16 (130; 0LDIG)
Instrumen	Nilai logit terendah	-.81 (227; 7PDIC)	-1.66 (514; 4LDPJ)
	<i>Alpha Cronbach</i>	.90	.93
	<i>Raw variance explained by measures</i>	30.8	55.4
	<i>Unexplained variance in 1<sup>st</sup> contrast</i>	12.3	6.2
	<i>Unexecpected variance in 2<sup>nd</sup> contrast</i>	6.2	6.1
	<i>Unexecpected variance in 3<sup>rd</sup> contrast</i>	5.0	5.5
	<i>Unexecpected variance in 4<sup>th</sup> contrast</i>	3.6	4.2
	<i>Unexecpected variance in 5<sup>th</sup> contrast</i>	2.9	4.0

Nilai *separation* digunakan untuk mengelompokkan *person* dan item. Makin besar nilai *separation* menunjukkan makin bagus kualitas instrumen karena dapat mengidentifikasi kelompok responden dan kelompok item. Persamaan yang digunakan yaitu:  $H = \frac{[(4 \times 2,60) + 1]}{3} = 3,8$  dibulatkan menjadi 4. Hasil ini bermakna terdapat empat kelompok partisipan.

#### d) Bentuk Akhir *WAI-T*

Dengan mempertimbangkan hasil pengujian kelayakan, maka *WAI-T-SF* versi calon konselor Indonesia dipilih dan digunakan dalam penelitian ini. *WAI-T-SF* diharapkan dapat menjadi alat asesmen aliansi konseling bagi calon konselor yang lebih efisien, praktis, dan valid. Instrumen ini mengukur tiga dimensi aliansi konseling (*bond*, *tasks*, dan *goals*) yang setiap dimensinya berjumlah 4 item. *WAI-T-SF* merupakan alat ukur laporan diri berbentuk *summated rating* dengan skala Likert 7 jenjang yang ketentuannya disajikan pada Tabel 3.17 Kisi-kisi *WAI-T-SF* versi calon konselor Indonesia disajikan pada Tabel 3.15 dan instrumen lengkapnya dapat diperiksa pada Lampiran D.

**Tabel 3.17 Pedoman Penskoran *WAI-T-SF***

Aitem	Skor Alternatif Jawaban						
	Selalu	Sangat Sering	Sering	Agak Sering	Agak Jarang	Jarang	Tidak Pernah
Positif	7	6	5	4	3	2	1
Negatif	1	2	3	4	5	6	7

## 2. Instrumen Penelitian *Delphi*

Skala Penilaian PEKA (SP-PEKA) digunakan untuk menilai kerangka pikir integratif komponen-komponen inti landasan filosofis-teoretis-praktik PEKA hasil kajian meta-analisis, yaitu: (a) kaidah nilai utama hakikat manusia; (b) kaidah nilai utama konselor; (c) kaidah nilai utama program/kurikulum; (d) kaidah nilai utama praktik; (e) kaidah nilai utama penilaian (Chenault, 1971; Kartadinata, 2011); (f) kaidah nilai utama hakikat dan tujuan pendidikan (Bruya & Ardel, 2018a; Brown, 2004; Sternberg & Hagen, 2019; Sternberg, Jarvin, & Grigorenko, 2009); (g) kaidah nilai utama metode pendidikan; (h) kaidah nilai utama prinsip, prosedur, proses pendidikan (Bruya & Ardel, 2018a; Brown, 2004;

Sternberg & Hagen, 2019; Sternberg, Jarvin, & Grigorenko, 2009; Bassett, 2011b); (i) kaidah nilai utama lingkungan pendidikan (Brown, 2004; DeMichelis, Ferrari, & Rozin, 2015; Ferrari & Kim, 2019; Osterlund, 2016); dan (j) kaidah nilai utama supervisi pendidikan (BTW; Sternberg, 2001a; 2004; 2013; Sternberg & Hagen, 2019; Sternberg, Reznitskaya, & Jarvin, 2007; Sternberg, Jarvin, & Alina, 2008; Sternberg, Jarvin, & Grigorenko, 2009).

SP-PEKA berbentuk Skala *Rating* 5 jenjang dari 1 = Sangat Tidak Penting/Sangat Tidak Sesuai/Sangat Tidak Mungkin; 2 = Tidak Penting/Tidak Sesuai/Tidak Mungkin; 3 = Cukup Penting/Cukup Sesuai/Cukup Mungkin; 4 = Penting/Sesuai/Mungkin; 5 = Sangat Penting/Sangat Sesuai/Sangat Mungkin; atau Instrumen SP-PEKA disajikan pada Lampiran C.7.3.

## C. Prosedur Penelitian

### 1. Prosedur Penelitian Kuantitatif

Penelitian kuantitatif dilakukan dengan prosedur berikut. Pertama, memohon izin penelitian kepada 11 Ketua Prodi BK jenjang S1 di Indonesia via pos, e-mail, dan telepon. Kedua, menginformasikan dan mengkoordinasikan jadwal pelaksanaan pengumpulan data kepada ketua Prodi BK, dosen pengampu mata kuliah, dan calon konselor semester keenam dari 11 Prodi BK. Ketiga, pelaksanaan pengumpulan data dilakukan secara tradisional menggunakan *paper and pencil questionnaire* dengan mendatangi langsung calon konselor semester keenam yang menjadi sampel penelitian di 11 Prodi BK di Indonesia sesuai jadwal yang telah disepakati. Calon konselor diberikan penjelasan tentang tujuan penelitian, jaminan kerahasiaan data, serta petunjuk teknis sebelum pengisian instrumen dilakukan. Pengumpulan data dilaksanakan pada bulan Maret s.d Juli 2016. Setelah data penelitian terkumpul, langkah terakhir yaitu melakukan verifikasi, pengolahan, dan analisis data penelitian untuk menjawab pertanyaan dan hipotesis penelitian yang diajukan.

## 2. Prosedur Penelitian *Delphi*

Prosedur studi *delphi* untuk mengembangkan dan memvalidasi PEKA dirancang antara 1-3 putaran sampai mencapai konsensus. Penetapan ini didasarkan pada kajian literatur bahwa 1-3 putaran *delphi* memadai untuk mengumpulkan data yang dibutuhkan dan mencapai konsensus (Custer, Scarcella, & Stewart, 1999; Day & Bobeva, 2005; Davidson, 2013; Delbecq, Van de Ven, & Gustafson, 1975; Kim & Yeo, 2018).

Pada putaran pertama, proses *delphi* dimulai dengan kuesioner semi terstruktur untuk mengumpulkan informasi/masukan spesifik dan menilai PEKA dari pandangan pakar. Kuesioner semi terstruktur dipilih daripada kuesioner terbuka karena informasi dasar dari isu yang ditargetkan tersedia dan dapat digunakan (Kerlinger, 1973). Selanjutnya, data yang terkumpul dari para pakar dianalisis dan dibuat laporannya sebagai bahan untuk putaran kedua jika belum mencapai konsensus.

Pada putaran kedua, para pakar menerima kuesioner kedua dan diminta untuk meninjau ringkasan informasi hasil penelitian pada putaran pertama jika belum mencapai konsensus. Selanjutnya, pakar diminta untuk menilai kembali PEKA. Selain itu, pakar juga diminta untuk menyampaikan alasan penilaiannya. Hasil yang ditargetkan dari putaran kedua yaitu teridentifikasi tanggapan aktual, kesepakatan dan pertentangan, serta terbentuk konsensus awal PEKA dari para pakar.

Pada putaran ketiga, pakar menerima kuesioner dan ringkasan hasil penelitian pada putaran kedua jika belum mencapai konsensus. Selanjutnya, pakar diminta untuk mengkaji ulang dan merevisi penilaiannya atau memberikan alasan khusus untuk tetap berada di luar konsensus. Putaran ketiga memberikan kesempatan kepada pakar untuk mengklarifikasi dan menimbang kembali informasi/substansi materi yang penting.

Dalam pelaksanaannya, penelitian ini dilakukan satu putaran *delphi*. Hal ini karena pada putaran pertama sudah diperoleh konsensus pakar sesuai dengan ketentuan yang telah ditetapkan. Proses pengadministrasian mulai dari

perencanaan, pengumpulan, pengolahan, analisis, dan pelaporan putaran *delphi* dilaksanakan selama satu minggu.

#### D. Analisis Data

##### 1. Analisis Data Kuantitatif

Teknik analisis data untuk menjawab pertanyaan tentang dimensi kearifan konselor dan anteseden kearifan konselor menurut perspektif calon konselor menggunakan *EFA*.

Teknik analisis data untuk menjawab pertanyaan penelitian tentang profil kearifan pribadi calon konselor menggunakan rata-rata dan persentase. Klasifikasi profil kearifan, baik total *3D-WS* maupun setiap dimensinya (kognitif, reflektif, dan afektif) menggunakan kriteria:  $\bar{X} < 3$  = rendah,  $\geq 3 \bar{X} < 4$  = sedang, dan  $\bar{X} \geq 4$  = tinggi (Ardelt, 2003; 2009; Bergsma & Ardel, 2012; Ferrari, Kahn, Benayon, & Nero, 2011). Tinggi atau rendahnya menggambarkan tingkat kearifan pribadi calon konselor.

Teknik analisis data untuk menjawab pertanyaan penelitian tentang profil kearifan multibudaya menggunakan statistik deskriptif. Penskoran terentang dari 22 - 110 untuk total *MC-WiSe*, 12 - 60 untuk dimensi kognitif dan 10 - 50 untuk dimensi afektif dan kesadaran, dan 2 - 10 untuk setiap indikator. Selain itu, dapat juga menggunakan pedoman penskoran lima dimensi: 8 - 40 untuk dimensi metakognitif, 5 - 25 untuk dimensi mengelola emosi, 4 - 20 untuk dimensi keterampilan *coping*, 3 - 15 untuk dimensi empati, dan 2 - 10 untuk dimensi penalaran dan toleransi terhadap ambiguitas. Tinggi atau rendahnya skor merefleksikan tingkat kearifan umum calon konselor multibudaya. Klasifikasi profil kearifan multibudaya, baik secara keseluruhan, per dimensi maupun indikator menggunakan kriteria:  $\bar{X} < 3$  = rendah,  $\geq 3 \bar{X} < 4$  = sedang, dan  $\bar{X} \geq 4$  = tinggi. Tinggi atau rendahnya skor menggambarkan tingkat kearifan multibudaya.

Teknik analisis data untuk menjawab pertanyaan penelitian tentang profil aliansi konseling menggunakan rata-rata dan persentase. Penskoran terentang dari 4 - 28 untuk setiap dimensi (*bond*, *tasks*, dan *goals*) dan 12 - 84 untuk total *WAI-T-12*. Klasifikasi aliansi konseling calon konselor, baik secara keseluruhan maupun per dimensi menggunakan kriteria, yaitu:  $\bar{X} < 3$  = kurang baik/rendah,  $\leq 3 \bar{X} < 6$  = cukup baik/sedang, dan  $\bar{X} \geq 6$  = baik/tinggi. Skor yang tinggi

menggambarkan tingkat aliansi konseling yang lebih positif/baik, begitupun sebaliknya.

Teknik analisis data untuk menjawab pertanyaan penelitian tentang pengaruh variabel etno-sosio-demografi (institusi pendidikan, status institusi pendidikan, status akreditasi program studi, jenis kelamin, usia, agama, etnis, pengalaman hidup kritis, dan frekuensi konseling) terhadap kearifan pribadi (keseluruhan serta dimensi kognitif, reflektif, dan afektif) dan kearifan multibudaya (keseluruhan serta dimensi kognitif dan afektif dan kesadaran) calon konselor menggunakan *General Linear Model Multivariate Analysis of Variance (GLM MANOVA)*.

Teknik analisis data untuk menjawab pertanyaan penelitian tentang pengaruh variabel etno-sosio-demografi (institusi pendidikan, status institusi pendidikan, status akreditasi program studi, jenis kelamin, usia, agama, etnis, pengalaman hidup kritis, dan frekuensi konseling), kearifan pribadi (keseluruhan serta dimensi kognitif, reflektif, dan afektif) dan kearifan multibudaya (keseluruhan serta dimensi kognitif dan afektif dan kesadaran) terhadap aliansi konseling calon konselor menggunakan *General Linear Model Multivariate Analysis of CoVariance (GLM MANCOVA)*.

Pengolahan data penelitian secara operasional menggunakan perangkat lunak *Microsoft Excel 2010, IBM SPSS v. 25.0 for Windows, dan Rasch Model v. 3.73*.

## 2. Analisis Data Penelitian *Delphi*

Teknik analisis data hasil studi *delphi* tentang PEKA menggunakan rata-rata dan standar deviasi. Teknik ini digunakan relevan dengan anjuran pakar dan peneliti sebelumnya bahwa rata-rata dan standar deviasi dapat digunakan untuk menentukan konsensus pakar dalam studi *delphi* (Keeney, Hasson, & McKenna, 2011; Kim & Yeo, 2018).

Penetapan konsensus pakar terhadap PEKA didasarkan pada saran dari Green (Hsu & Stanford, 2007) bahwa minimal 70% pakar menilai  $\geq 3$  pada Skala Likert 5 jenjang dengan  $\bar{X} \geq 3.2$ . Klasifikasi konsensus pakar ditetapkan menjadi tiga kategori, yaitu:  $\bar{X} < 3.2$  = kurang sepakat;  $\geq 3.2$   $\bar{X} < 4.0$  = cukup sepakat; dan  $\geq 4.0$  = sepakat.