

## **BAB III**

### **METODE PENELITIAN**

Bab ini memuat metode penelitian yang digunakan. Pada bab ini akan dipaparkan mengenai desain penelitian, populasi dan sampel penelitian, teknik pengambilan data, instrumen penelitian, prosedur penelitian, dan teknik analisis data.

#### **3.1. Desain Penelitian**

Penelitian ini menggunakan pendekatan kuantitatif dengan desain *cross-sectional survey*. Pendekatan kuantitatif yang digunakan dalam penelitian ini dimaksudkan untuk mengumpulkan data numerik dari sejumlah individu menggunakan instrumen yang pertanyaan dan tanggapannya telah ditentukan sebelumnya (Creswell, 2012). Kemudian, data yang dihasilkan akan dianalisis secara statistik untuk mengetahui properti psikometri dari instrumen yang diadaptasi, hal ini mencakup pengujian analisis item, nilai validitas, dan nilai reliabilitas. Adapun desain *cross-sectional survey* dipilih atas dasar efisiensi waktu sehingga pengumpulan data dalam penelitian ini hanya dilakukan pada satu titik waktu saja (Creswell, 2012).

#### **3.2. Populasi Penelitian**

Populasi yang dipilih pada penelitian ini adalah perempuan di Indonesia, dengan rentang usia 18-40 tahun, serta pernah mengalami kekerasan dalam hubungan romantis yang sedang atau pernah dijalaninya, dalam hal ini spesifik pada hubungan pacaran. Penentuan kriteria jenis kelamin dan usia didasarkan pada target partisipan yang dipilih oleh pengembang asli instrumen ini (Bhatti et al., 2021), serta diperkuat oleh data bahwa 47.94% korban kasus kekerasan terhadap perempuan di Indonesia dialami oleh individu dengan rentang usia tersebut dan sebagian besar terjadi dalam ranah personal yang bersifat intim, khususnya pada hubungan pacaran (Kementerian PPPA, 2024; Komnas Perempuan, 2024). Menurut ITC (2016), adanya kesesuaian antara target populasi yang dipilih dalam penelitian yang akan dilakukan dengan penelitian sebelumnya lebih direkomendasikan untuk meminimalisasi perbedaan hasil yang didapatkan. Adapun penentuan kriteria pernah mengalami kekerasan dalam hal ini meliputi kekerasan fisik, psikis, seksual,

dan/atau ekonomi. Hal ini didasarkan pada pendapat Mukhtar (2023) dan Sweet (2019) yang mengemukakan dalam artikelnya bahwa *gaslighting* sebagai salah satu jenis kekerasan psikis juga dapat terjadi secara bersamaan atau saling beriringan dengan tindakan kekerasan berbasis gender lainnya, seperti kekerasan fisik, seksual, dan ekonomi.

### 3.3. Sampel Penelitian

Teknik pengambilan sampel yang digunakan dalam penelitian ini yaitu teknik *nonprobability sampling* dengan pendekatan *purposive sampling*. Penentuan teknik tersebut didasarkan pada hadirnya karakteristik khusus yang ditentukan peneliti terhadap target individu yang akan diteliti sehingga data yang dihasilkan dapat lebih representatif (Cresswell, 2012). Karakteristik khusus tersebut yaitu individu yang pernah mengalami setidaknya satu dari empat bentuk kekerasan dalam hubungan, meliputi: kekerasan fisik, psikis, seksual, dan/atau ekonomi.

Ukuran sampel yang digunakan dalam penelitian ini terbagi ke dalam tiga jenis yang didasarkan pada tiga tahapan pengambilan data yang dilakukan, yaitu pengambilan data untuk uji keterbacaan, *pilot study 1*, dan *pilot study 2*. Pada tahap uji keterbacaan, ukuran sampel yang digunakan yaitu sebanyak 30-40 orang. Ukuran ini merupakan ukuran ideal yang disarankan oleh Beaton et al. (2000) untuk kebutuhan pengujian terkait. Selanjutnya, pada tahap *pilot study 1*, ukuran sampel yang ditargetkan berjumlah 140 orang. Penentuan ukuran ini didasarkan pada rasio 10:1 yang disarankan oleh Kline (2013) untuk model pengujian *Confirmatory Factor Analysis* (CFA), yang artinya terdapat setidaknya 10 responden dalam 1 item pertanyaan atau pernyataan yang diujikan. Terakhir, pada tahap *pilot study 2*, ukuran sampel minimal yang ditargetkan yaitu sebanyak 300 orang. Ukuran ini didasarkan pada kecukupan jumlah sampel yang dievaluasi menurut kriteria *rule of thumb* dari Comrey & Lee (1992; dalam Van Voorhis dan Morgan, 2007).

**Tabel 3.3.**

*Ukuran Sampel Berdasarkan Kriteria Rule of Thumb*

Jumlah Sampel	Kategori	Jumlah Sampel	Kategori
50	<i>Very Poor</i>	300	<i>Good</i>
100	<i>Poor</i>	500	<i>Very Good</i>
200	<i>Fair</i>	1000 atau lebih	<i>Excellent</i>

### 3.4. Teknik Pengambilan Data

Teknik pengambilan data dalam penelitian ini dilakukan secara *online* dengan menggunakan metode kuesioner. Kuesioner *online* tersebut disusun ke dalam *Google Form* lalu disebarluaskan melalui media sosial seperti *WhatsApp*, *Instagram*, *TikTok*, *Twitter*, dan lainnya. Secara keseluruhan, kuesioner *online* tersebut memuat beberapa bagian, mulai dari data demografis partisipan, alat ukur *Victim Gaslighting Questionnaire* (VGQ) sebagai alat ukur utama yang diadaptasi, serta tiga alat ukur lainnya yang digunakan sebagai alat ukur pengujian validitas konvergen, divergen, dan diskriminan.

### 3.5. Instrumen Penelitian

#### 3.5.1. *Victim Gaslighting Questionnaire* (VGQ)

##### a. Spesifikasi Alat Ukur

Alat ukur utama yang digunakan dalam penelitian ini adalah *Victim Gaslighting Questionnaire* (VGQ) yang dikembangkan oleh Bhatti et al. (2021) dan diadaptasi dengan prosedur yang ketat oleh peneliti. VGQ digunakan dengan tujuan untuk mengukur tingkat viktimisasi *gaslighting* pada korban. VGQ terdiri atas 14 item dengan dua dimensi, yaitu “*peer disagreement*” dan “*loss of self-trust*”. VGQ diukur dengan menggunakan skala Likert lima poin. Seluruh item dalam alat ukur VGQ merupakan item *favorable*.

##### b. Pengisian Alat Ukur

Dalam pengisian alat ukur VGQ, partisipan diminta untuk memilih satu dari lima alternatif jawaban yang tersedia. Alternatif jawaban yang dapat dipilih oleh partisipan berkisar pada rentang 0 sampai 4, mulai dari Tidak Pernah (TP), Jarang (JR), Kadang-Kadang (KD), Sering (SR), hingga Selalu (SL).

##### c. Penyekoran Alat Ukur

**Tabel 3.5.1a.**  
*Penyekoran Alat Ukur VGQ*

Jenis Item	Nilai Item				
	TP	JR	KD	SR	SL
<i>Favorable</i>	0	1	2	3	4

## d. Kisi-Kisi Alat Ukur

**Tabel 3.5.1b.**  
*Kisi-Kisi Alat Ukur VGQ*

Dimensi	Nomor Item	Jumlah
<i>Peer disagreement</i>	1, 2, 3, 4, 5	5
<i>Loss of self-trust</i>	6, 7, 8, 9, 10, 11, 12, 13, 14	9
<b>Total Item</b>		<b>14</b>

3.5.2. *Gaslighting Behaviour Questionnaire (GBQ)*

## a. Spesifikasi Alat Ukur

Alat ukur *Gaslighting Behaviour Questionnaire (GBQ)* merupakan alat ukur yang dikembangkan Dickson, Ireland, dan Birch (2023) yang diadaptasi melalui prosedur yang tidak terlalu ketat oleh peneliti. GBQ terdiri atas 10 item dengan 2 dimensi, yaitu “*promoting self-doubt, confusion and blame*” dan “*altering recollection of behaviour*”. GBQ diukur dengan menggunakan tipe skala Likert lima poin. Seluruh item dalam alat ukur GBQ merupakan item *favorable*.

Dalam penelitian ini, GBQ digunakan sebagai alat ukur pengujian validitas konvergen. Hal ini didasarkan pada kesamaan konstruk yang diukur antara GBQ dengan VGQ. Selain itu, GBQ juga memiliki koefisien reliabilitas yang sangat baik yaitu berkisar antara  $\alpha = 0.86-0.95$  (Dickson, Ireland, & Birch, 2023).

## b. Pengisian Alat Ukur

Dalam pengisian alat ukur GBQ, partisipan diminta untuk memilih satu dari lima alternatif jawaban yang tersedia. Alternatif jawaban yang dapat dipilih oleh partisipan berkisar pada rentang 0 sampai 4, mulai dari Tidak Pernah (TP), Jarang (JR), Kadang-Kadang (KD), Sering (SR), hingga Selalu (SL).

## c. Penyekoran Alat Ukur

**Tabel 3.5.2a.**  
*Penyekoran Alat Ukur GBQ*

Jenis Item	Nilai Item				
	TP	JR	KD	SR	SL
<i>Favorable</i>	0	1	2	3	4

## d. Kisi-Kisi Alat Ukur

**Tabel 3.5.2b.**  
*Kisi-Kisi Alat Ukur GBQ*

Dimensi	Nomor Item	Jumlah
<i>Promoting self-doubt, confusion and blame</i>	1, 2, 3, 4, 5, 6, 7	7
<i>Altering recollection of behaviour</i>	8, 9, 10	3
<b>Total Item</b>		<b>10</b>

3.5.3. *The Rosenberg Self-Esteem Scale (RSES)*

## a. Spesifikasi Alat Ukur

Alat ukur *The Rosenberg Self-Esteem Scale (RSES)* merupakan alat ukur yang dikembangkan oleh Rosenberg (1965) dan telah diadaptasi ke dalam bahasa Indonesia oleh Azzahra (2023). RSES terdiri atas 10 item dengan 2 dimensi, yaitu “*penerimaan diri*” dan “*penghormatan diri*”. RSES diukur dengan menggunakan tipe skala Likert empat poin. RSES terdiri atas item *favorable* dan *unfavorable*. Dalam penelitian ini, RSES digunakan sebagai alat ukur pengujian validitas divergen. Hal ini didasarkan pada beberapa penelitian sebelumnya yang mengungkapkan bahwa pengalaman *gaslighting* berkaitan erat dengan penurunan *self-esteem* korban (Stern, 2007; Tjiu, 2022). Selain itu, RSES juga memiliki koefisien reliabilitas yang sangat baik yaitu  $\alpha = 0.82$  (Azzahra, 2023).

## b. Pengisian Alat Ukur

Dalam pengisian alat ukur RSES, partisipan diminta untuk memilih satu dari empat alternatif jawaban yang tersedia. Alternatif jawaban yang dapat dipilih oleh partisipan berkisar pada rentang 1 sampai 4, mulai dari Sangat Tidak Setuju (STS), Tidak Setuju (TS), Setuju (S), hingga Sangat Setuju (SS) untuk item *favorable*, dan sebaliknya untuk item *unfavorable*.

## c. Penyekoran Alat Ukur

**Tabel 3.5.3a.**  
*Penyekoran Alat Ukur RSES*

Jenis Item	Nilai Item			
	STS	TS	S	SS
<i>Favorable</i>	1	2	3	4
<i>Unfavorable</i>	4	3	2	1

## d. Kisi-Kisi Alat Ukur

**Tabel 3.5.3b.**  
*Kisi-Kisi Alat Ukur RSES*

Dimensi	Nomor Item		Jumlah
	<i>Favorable</i>	<i>Unfavorable</i>	
Penerimaan Diri	7, 10	2, 6, 8, 9	6
Penghormatan Diri	1, 3, 4	5	4
<b>Total Item</b>			<b>10</b>

3.5.4. *Big Five Inventory* (BFI)

## a. Spesifikasi Alat Ukur

*Big Five Inventory* (BFI) merupakan alat ukur yang dikembangkan oleh John, Naumann, dan Soto (2008) dan telah diadaptasi kedalam bahasa Indonesia oleh Reza (2015). BFI terdiri atas 44 item dengan 5 dimensi, yaitu “*extraversion*”, “*agreeableness*”, “*conscientiousness*”, “*neuroticism*”, dan “*openness*”. BFI diukur dengan menggunakan tipe skala Likert lima poin. Alat ukur BFI terdiri atas item *favorable* dan *unfavorable*.

Dalam penelitian ini, BFI digunakan sebagai alat ukur pengujian validitas diskriminan. Hal ini didasarkan pada penelitian adaptasi sebelumnya yang telah menggunakan alat ukur ini sebagai bagian dari pengujian validitas diskriminan, khususnya pada dimensi “*agreeableness*” dan “*conscientiousness*”, yang kemudian hasilnya terbukti menunjukkan bahwa kedua dimensi tersebut memiliki korelasi negatif yang sangat rendah dengan pengalaman *gaslighting* (Hassan, Iqbal, & Hassan, 2022). Selain itu, BFI juga memiliki koefisien reliabilitas yang cukup baik di setiap dimensinya, yaitu: *extraversion* ( $\alpha = 0.659$ ), *agreeableness* ( $\alpha = 0.691$ ), *conscientiousness* ( $\alpha = 0.772$ ), *neuroticism* ( $\alpha = 0.812$ ), dan *openness* ( $\alpha = 0.709$ ; Azzahra, 2023).

## b. Pengisian Alat Ukur

Dalam pengisian alat ukur BFI, partisipan diminta untuk memilih satu dari lima alternatif jawaban yang tersedia. Alternatif jawaban yang dapat dipilih oleh partisipan berkisar pada rentang 1 sampai 5, mulai dari Sangat Tidak Setuju (STS), Tidak Setuju (TS), Netral (N), Setuju (S),

hingga Sangat Setuju (SS) untuk item *favorable*, dan sebaliknya untuk item *unfavorable*.

c. Penyekoran Alat Ukur

**Tabel 3.5.4a.**  
*Penyekoran Alat Ukur BFI*

Jenis Item	Nilai Item				
	STS	TS	N	S	SS
<i>Favorable</i>	1	2	3	4	5
<i>Unfavorable</i>	5	4	3	2	1

d. Kisi-Kisi Alat Ukur

**Tabel 3.5.4b.**  
*Kisi-Kisi Alat Ukur BFI*

Dimensi	Nomor Item		Jumlah
	<i>Favorable</i>	<i>Unfavorable</i>	
<i>Extraversion</i>	1, 11, 16, 26, 36	6, 21, 31	8
<i>Agreeableness</i>	7, 17, 22, 32, 42	2, 12, 27, 37	9
<i>Conscientiousness</i>	3, 13, 28, 33, 38	8, 18, 23, 43	9
<i>Neuroticism</i>	4, 14, 19, 29, 39	9, 24, 34	8
<i>Openness</i>	5, 10, 15, 20, 25, 30, 40, 44	35, 41	10
<b>Total Item</b>			<b>44</b>

### 3.6. Prosedur Penelitian

Secara umum, prosedur pelaksanaan adaptasi alat ukur yang dilakukan dalam penelitian ini mengacu pada panduan *cross-cultural adaptation of psychological instruments* milik Borsa, Damásio, dan Bandeira (2012). Panduan *cross-cultural adaptation* milik Beaton et al. (2000) dan *translating and adapting tests* milik ITC (2016) juga turut disertakan sebagai informasi pelengkap dalam setiap tahapan prosedur.

#### Tahap 1: *Pre-Condition*

Tahapan ini dilakukan sebelum proses adaptasi dimulai. Langkah yang dilakukan pada tahapan ini adalah permohonan izin kepada pemilik alat ukur asli bahwa akan dilakukan adaptasi terhadap alat ukur yang dikembangkannya. Dalam hal ini, peneliti menghubungi penulis korespondensi pengembangan alat ukur

*Victim Gaslighting Questionnaire* (VGQ; Bhatti et al., 2021) melalui email. Langkah ini didasarkan pada panduan ITC (2016) yang menyebutkan bahwa pemilik alat ukur asli sebagai pemegang hak kekayaan intelektual berhak mengetahui dan memberikan izin pelaksanaan adaptasi terhadap karya ilmiahnya.

## **Tahap 2: *Cross-Cultural Adaptation***

Tahapan ini memuat rincian penjelasan mengenai setiap langkah yang dilakukan dalam proses penerjemahan dan pengadaptasian alat ukur, yaitu sebagai berikut:

### 1. *Forward Translation*

*Forward translation* merupakan langkah pertama dalam proses adaptasi alat ukur. Pada langkah ini, terdapat dua penerjemah yang menerjemahkan alat ukur secara independen dari bahasa asli (bahasa Inggris) ke dalam bahasa sasaran (bahasa Indonesia). Menurut Beaton et al. (2000), kedua penerjemah tersebut harus berasal dari latar belakang yang berbeda, yaitu sebagai berikut:

- a. Penerjemah pertama (T1) adalah ahli dwibahasa yang memahami konstruk yang sedang diperiksa dari alat ukur yang sedang diterjemahkan. Kriteria seperti ini dapat memberikan kemungkinan dihasilkannya terjemahan dengan kesepadanan yang lebih dapat diandalkan dari perspektif psikometrik karena cenderung memberikan kesamaan ilmiah yang lebih tinggi dari alat ukur (Beaton et al., 2000). Adapun penerjemah pertama yang dilibatkan pada langkah ini adalah seorang mahasiswa psikologi di Universitas Pendidikan Indonesia (UPI) yang memiliki nilai sertifikasi IELTS sebesar 8.0.
- b. Penerjemah kedua (T2) adalah ahli dwibahasa yang tidak memahami konstruk yang sedang diperiksa dan tidak berasal dari latar belakang keilmuan yang relevan dengan alat ukur yang sedang diterjemahkan. Kriteria seperti ini dapat memberikan kemungkinan dihasilkannya versi terjemahan yang cenderung lebih mencerminkan bahasa yang digunakan oleh populasi sasaran karena tidak terlalu dipengaruhi oleh tujuan akademis dari penerjemahan



(Beaton et al., 2000). Adapun penerjemah kedua yang dilibatkan pada langkah ini adalah seorang sarjana sastra Inggris dari Universitas Gadjah Mada (UGM).

Penggunaan standar minimal dua penerjemah ini dilakukan dengan tujuan untuk meminimalkan risiko bias bahasa, psikologis, budaya, dan pemahaman teoretis maupun praktis (Borsa, Damásio, & Bandeira, 2012). Dalam proses pelaksanaannya, peneliti juga menginstruksikan kepada masing-masing penerjemah untuk membuat laporan tertulis mengenai hasil terjemahannya serta menyertakan komentar tambahan apabila terdapat kendala yang dialami. Melalui cara ini, hasil terjemahan dapat dibandingkan sehingga kata-kata yang dianggap ambigu atau kurang tepat dalam proses penerjemahan dapat diidentifikasi dan didiskusikan lebih lanjut pada langkah sintesis hasil terjemahan (Beaton et al., 2000).

## 2. *Synthesis of the Translated Versions*

Setelah melalui proses *forward translation*, peneliti mendapatkan dua versi hasil terjemahan alat ukur. Pada langkah ini, peneliti dan seorang juri (*judge*) berdiskusi untuk membandingkan kedua versi hasil terjemahan tersebut dan menilai kesesuaiannya dengan versi aslinya guna dihasilkannya satu versi terjemahan final. Juri adalah seorang ahli yang memiliki pengetahuan khusus tentang apa yang dinilai oleh alat ukur (Borsa, Damásio, & Bandeira, 2012). Dalam penelitian ini, juri yang dipilih adalah seorang dosen psikologi gender di Universitas Pendidikan Indonesia (UPI) yang memiliki kepedulian terhadap isu kekerasan pada perempuan serta mempunyai pengalaman empiris dalam menangani korban *gaslighting*.

Dalam pelaksanaannya, proses sintesis terjemahan alat ukur ini dilakukan dengan memperhatikan kesesuaian antara alat ukur versi terjemahan dengan versi aslinya dalam empat aspek, yaitu: (1) kesepadanan semantik, yang bertujuan untuk menilai kesamaan dan kejelasan makna serta ketepatan tata bahasa dari alat ukur yang telah diterjemahkan; (2) kesepadanan idiomatik, yang bertujuan untuk menilai kesetaraan ekspresi dari item yang sulit untuk diterjemahkan; (3) kesepadanan pengalaman, yang bertujuan untuk menilai kesetaraan budaya dari item yang

diterjemahkan; dan (4) kesepadanan konseptual, yang bertujuan untuk menilai ketepatan penggunaan istilah dalam menggambarkan aspek/dimensi yang hendak diukur (Borsa, Damásio, & Bandeira, 2012). Jika versi terjemahan memiliki kekurangan pada satu atau lebih aspek tersebut, peneliti dan juri berdiskusi untuk saling mengusulkan dan memilih terjemahan baru yang lebih sesuai dengan karakteristik alat ukur dan konteks budaya di Indonesia. Hasil sintesis terjemahan dibuat dalam sebuah laporan tertulis yang mendokumentasikan proses sintesis, isu yang dibahas, dan bagaimana isu-isu tersebut diselesaikan hingga mendapatkan keputusan yang paling tepat.

### 3. *Evaluation of the Synthesized Version by Experts*

Setelah melakukan sintesis terjemahan, langkah yang dilakukan selanjutnya adalah mengevaluasi hasil sintesis terjemahan tersebut oleh para ahli. Pada umumnya, langkah ini lebih sering dikenal dengan sebutan *expert judgement*. Pada langkah ini, peneliti meminta para ahli di bidang psikologi untuk memberikan evaluasi secara kualitatif dan penilaian secara kuantitatif terhadap beberapa aspek yang terkandung dalam alat ukur yang sedang diadaptasi secara keseluruhan untuk mendapatkan nilai validitas isi. Beberapa aspek tersebut meliputi: (a) aspek relevansi dan representasi komponen-komponen dalam alat ukur, (b) aspek kesesuaian item dengan konstruk yang diukur, serta (c) aspek ketepatan tata bahasa dan kesepadanan budaya (Ihsan, 2015).

Kriteria yang digunakan dalam pemilihan ahli yaitu individu yang memiliki kemampuan untuk melakukan evaluasi psikologis serta memiliki pengetahuan yang cukup mengenai apa yang sedang dinilai oleh alat ukur. Hal ini didasarkan pada rekomendasi yang diberikan oleh Borsa, Damásio, dan Bandeira (2012). Adapun ahli yang dipilih dalam penelitian ini terdiri atas dua orang dosen psikologi di Universitas Pendidikan Indonesia sebagai perwakilan ahli dalam bidang akademisi dan psikometri, serta satu orang konselor di UPTD PPA Kota Bandung sebagai perwakilan ahli dalam bidang praktisi. Penggunaan beberapa ahli dari latar belakang yang berbeda

dilakukan untuk memperkaya sudut pandang terhadap penilaian yang dilakukan.

4. *Evaluation by the Target Population*

Langkah selanjutnya yaitu mengevaluasi pemahaman target populasi terhadap seluruh item dan instruksi yang telah dievaluasi oleh peneliti dan para ahli pada langkah sebelumnya. Pada umumnya, langkah ini lebih sering dikenal dengan sebutan uji keterbacaan. Langkah ini dilakukan untuk memastikan bahwa istilah yang digunakan pada alat ukur yang diadaptasi sudah jelas dan dapat dipahami oleh target populasi. Partisipan yang terlibat pada langkah ini berjumlah 30-40 orang dengan kriteria yang mencerminkan karakteristik target populasi. Setiap partisipan diminta untuk mengisi sebuah kuesioner yang berisi daftar evaluasi pemahaman mereka terhadap setiap item yang disajikan. Jika terdapat item yang maknanya kurang dapat dipahami dengan jelas, mereka diberikan kesempatan untuk memberikan tanggapan dan saran perbaikan di akhir pengisian kuesioner.

5. *Back Translation*

*Back translation* mengacu pada penerjemahan versi alat ukur yang telah disintesis dan dievaluasi dari bahasa sasaran (bahasa Indonesia) ke dalam bahasa asli (bahasa Inggris). Tujuan dilakukannya *back translation* adalah untuk mengevaluasi sejauh mana versi terjemahan mencerminkan isi item dari versi aslinya, mengidentifikasi kata-kata yang tidak jelas dalam bahasa sasaran, dan mengidentifikasi ketidakkonsistenan atau kesalahan konseptual dalam versi akhir terjemahan (Beaton et al., 2000). Menurut Borsa, Damásio, dan Bandeira (2012), *back translation* tidak menyiratkan bahwa suatu item harus tetap identik secara harfiah dengan versi aslinya, melainkan harus mempertahankan kesetaraan konseptual dan mempertimbangkan makna item dalam konteks budaya yang sesuai.

Beaton et al. (2000) menyarankan agar *back translation* setidaknya dilakukan oleh dua orang penerjemah yang berbeda dengan penerjemah pada langkah *forward translation*. Perlu dipastikan bahwa kedua penerjemah tersebut tidak mengetahui versi alat ukur yang asli, tidak

memiliki pengetahuan terkait konsep-konsep yang dieksplorasi, serta tidak berasal dari latar belakang pendidikan yang relevan dengan alat ukur yang diterjemahkan agar diperoleh hasil terjemahan yang bebas bias (Beaton et al., 2000). Dalam penelitian ini, penerjemah yang terlibat adalah dua orang ahli dwibahasa yang merupakan penerjemah independen dengan latar belakang keilmuan non-psikologi.

#### 6. *Synthesized Translated Version*

Pada langkah ini, peneliti meminta seorang ahli dwibahasa dengan latar belakang keilmuan psikologi untuk melakukan sintesis terhadap dua versi *back translation* yang telah didapatkan guna dihasilkannya satu versi terjemahan final. Ahli dwibahasa tersebut merupakan seorang mahasiswa psikologi di Universitas Pendidikan Indonesia (UPI) yang memiliki nilai sertifikasi DET sebesar 140, serta memiliki pengetahuan yang cukup mengenai konstruk yang diukur.

Setelah itu, peneliti bersama dengan seorang juri (*judge*) berdiskusi untuk menilai kesesuaian makna antara versi sintesis *back translation* dengan versi aslinya. Juri yang terlibat dalam tahapan ini sama seperti pada tahapan sintesis *forward translation*. Berdasarkan pelaksanaan diskusi tersebut, didapatkan hasil bahwa versi sintesis *back translation* dari alat ukur yang sedang diadaptasi sudah memiliki makna yang sesuai dengan versi aslinya, meskipun secara bahasa terdapat perbedaan.

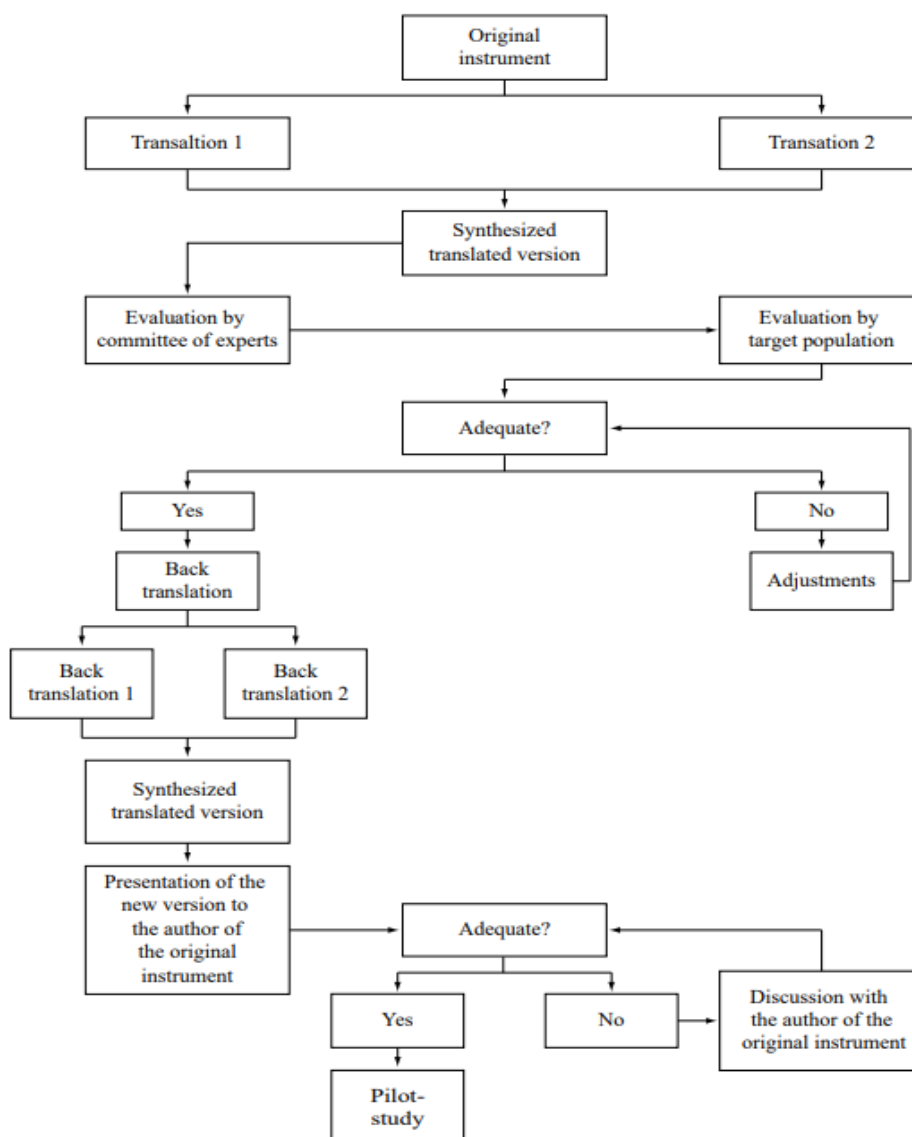
#### 7. *Pilot Study*

Studi percontohan (*pilot study*) mengacu pada pengaplikasian alat ukur hasil adaptasi ke dalam sampel yang mencerminkan karakteristik target populasi (Gudmundsson, 2009). Tujuan dilakukan *pilot study* adalah untuk menilai apakah alat ukur yang diadaptasi sudah siap untuk digunakan. Dalam penelitian ini, *pilot study* dilakukan sebanyak dua kali, yang meliputi *pilot study 1* dan *pilot study 2*. *Pilot study 1* atau *try out* alat ukur dilakukan pada sampel kecil untuk menguji kesesuaian model antara alat ukur versi adaptasi dengan versi aslinya menggunakan metode *Confirmatory Factor Analysis* (CFA). Adapun *pilot study 2* dilakukan untuk menguji nilai

validitas dan reliabilitas dari alat ukur yang diadaptasi pada ukuran sampel yang lebih besar.

Berikut merupakan gambaran garis besar dari prosedur *cross-cultural adaptation of psychological instruments* milik Borsa, Damásio, dan Bandeira (2012).

**Gambar 3.6.**  
*Prosedur Adaptasi Alat Ukur*



### 3.7. Teknik Analisis Data

Analisis data kuantitatif yang dilakukan pada penelitian ini akan terbagi ke dalam beberapa tahapan. Keseluruhan tahapan analisis data dilakukan dengan menggunakan bantuan *software* IBM SPSS *Statistics* 24 dan JASP 0.19. Berikut penjelasan lebih lanjut mengenai setiap tahapan analisis data yang dilakukan:

## 1. Analisis Item

Analisis item dilakukan untuk menguji kualitas dari setiap item dalam alat ukur yang diadaptasi. Indeks yang digunakan untuk melakukan analisis item dalam penelitian ini adalah daya diskriminasi item. Daya diskriminasi item dapat dilihat berdasarkan koefisien *corrected item-total correlations* (Coaley, 2010). Menurut Boateng et al. (2018), item dengan kualitas yang baik memiliki skor *corrected item-total correlations* di atas 0.30. Item dengan skor kurang dari ketentuan tersebut perlu dibuang karena dianggap memiliki daya diskriminasi yang rendah dan dapat berpengaruh pada reliabilitas alat ukur (Azwar, 2021).

## 2. Validitas Isi

Pengujian validitas isi dilakukan untuk menilai sejauh mana elemen-elemen dalam alat ukur relevan dan merepresentasikan konstruk yang sesuai dengan tujuan pengukuran (Haynes, Richard, & Kubany, 1995). Terdapat beberapa metode yang dapat digunakan untuk menilai validitas isi, salah satunya yaitu melalui koefisien V Aiken (Aiken, 1985). Koefisien V Aiken adalah metrik yang digunakan untuk mengevaluasi tanggapan yang diberikan oleh penilai atau ahli terhadap item-item dalam suatu alat ukur (Nurjanah et al., 2023). Koefisien V Aiken yang tinggi mengindikasikan bahwa item tersebut memiliki validitas isi yang tinggi (Aiken, 1985). Koefisien V Aiken diperoleh berdasarkan penilaian beberapa *rater* dalam prosedur *expert judgement* dan uji keterbacaan. Perhitungan koefisien V Aiken dilakukan melalui rumus berikut:

$$V = \frac{\sum s}{n(c-1)}$$

Keterangan:

- $V$  = Indeks Aiken
- $s$  =  $r - l_o$
- $r$  = Skor yang diberikan penilai
- $l_o$  = Skor penilaian terendah
- $n$  = Jumlah penilai
- $c$  = Skor penilaian tertinggi

Koefisien V Aiken yang tinggi mengindikasikan bahwa item tersebut memiliki validitas isi yang tinggi (Aiken, 1985). Adapun kategorisasi dari indeks validitas berdasarkan koefisien V Aiken (Retnawati, 2016) dapat dilihat pada tabel berikut:

**Tabel 3.7a.**  
*Kategorisasi Indeks Validitas*

<b>Indeks Validitas</b>	<b>Kategori</b>
$0 \leq V \leq 0.4$	Tidak Valid
$0.4 < V \leq 0.8$	Validitas Sedang
$0.8 < V \leq 1$	Validitas Tinggi

### 3. Validitas Konstruk

Penelitian ini menguji validitas konstruk dengan dua pendekatan, yaitu pendekatan analisis faktor dan *multitrait-multimethod*. Berikut merupakan penjelasan lebih lanjut dari masing-masing pengujian validitas konstruk yang dilakukan:

#### a. Pendekatan Analisis Faktor (*Factor Analysis*)

Analisis faktor untuk pengujian validitas faktorial yang digunakan pada penelitian ini adalah *Confirmatory Factor Analysis* (CFA). CFA dilakukan untuk menetapkan indeks kecocokan model serta untuk mengidentifikasi muatan faktor dari setiap item pada versi terjemahan skala (Hassan, Iqbal, Hassan, 2022). Dalam pengujian CFA, indeks kecocokan (*fit index*) yang digunakan terdiri dari beberapa parameter. Parameter utama yang digunakan yaitu statistik *chi-square* ( $\chi^2$ ). Adapun parameter lainnya berupa dua indeks kecocokan absolut, yaitu *Root Mean Square Error of Approximation* (RMSEA) dan *Standardized Root Mean Square Residual* (SRMR), serta dua indeks kecocokan inkremental, yaitu *Comparative Fit Index* (CFI) dan *Tucker–Lewis index* (TLI). Keseluruhan parameter di atas merupakan parameter utama yang paling banyak digunakan dan disarankan untuk menilai kecocokan model pada penelitian-penelitian sebelumnya (misalnya, Furr, 2011; Leach et al., 2008; Sarmiento & Costa, 2019; Xia & Yang 2019).

- 1) *Chi-square* ( $\chi^2$ ) merupakan statistik uji kecocokan model faktor dengan cara membandingkan matriks kovarians yang diamati

dengan matriks kovarians yang diusulkan secara teoritis (Boateng et al., 2018). Nilai *chi-square* yang mendekati nol mengindikasikan kecocokan yang lebih baik, yaitu perbedaan yang lebih kecil antara matriks kovarians yang diharapkan dan yang diamati (Sarmiento & Costa, 2019), yang artinya memberikan dukungan terhadap model pengukuran yang dihipotesiskan (Furr, 2011). Sebaliknya, nilai *chi-square* yang besar dan signifikan dengan *p-value* yang kecil ( $p < 0.05$ ) mengindikasikan ketidaksesuaian yang besar dan memberikan bukti yang menentang terhadap model yang dihipotesiskan (Furr, 2011; Sarmiento & Costa, 2019). Namun, nilai *chi-square* sangat sensitif terhadap jumlah sampel dan biasanya mengindikasikan ketidaksesuaian yang signifikan bahkan untuk model yang cocok untuk tujuan praktis (Cook, Kallen, & Amtmann, 2009).

- 2) *Root Mean Square Error of Approximation* (RMSEA) merupakan ukuran yang mengestimasi perbedaan antara populasi dan matriks kovarians populasi yang diimplikasikan model per derajat kebebasan (Boateng et al., 2018). Nilai RMSEA mengestimasi perbedaan antara model yang diperiksa dan model hipotesis di mana setiap komponen dalam model terkait dengan setiap komponen lainnya (Cook, Kallen, & Amtmann, 2009). Secara sederhana, RMSEA menilai seberapa jauh model yang dihipotesiskan dari model yang sempurna (Xia & Yang, 2019). Browne dan Cudeck (1992) merekomendasikan  $RMSEA \leq 0,05$  sebagai indikasi kecocokan yang sangat fit atau “*close fit*”,  $0,05 \leq RMSEA \leq 0,08$  sebagai indikasi kecocokan yang cukup fit atau “*fair fit*”, dan nilai  $> 0,10$  sebagai indikasi kecocokan yang buruk atau “*poor fit*”.
- 3) *Standardized Root Mean Square Residual* (SRMR) merupakan perbedaan rata-rata antara matriks korelasi dari sampel yang diamati dan model yang dihipotesiskan (Cook, Kallen, & Amtmann, 2009). Boateng et al. (2018) menyebutkan bahwa ambang batas untuk kecocokan model yang dapat diterima atau “*acceptable fit*” adalah  $SRMR \leq 0.08$ .



- 4) *Comparative Fit Index* (CFI) merupakan indeks kecocokan relatif inkremental yang mengukur peningkatan relatif dalam kecocokan model peneliti dibandingkan dengan model *baseline* (Boateng et al., 2018). Indeks ini dianggap sangat baik atau “*very good*” jika  $\geq 0.95$ , baik atau “*good*” jika di antara 0.9 dan 0.95, kurang baik atau “*suffering*” jika di antara 0.8 dan 0.9, dan buruk atau “*bad*” jika  $< 0.8$  (Sarmiento & Costa, 2019).
- 5) *Tucker–Lewis index* (TLI) atau yang dikenal juga dengan *Non-Normed Fit Index* (NNFI) merupakan kombinasi dari ukuran parsimoni dengan indeks komparatif antara model yang diusulkan dan model nol (Sarmiento & Costa, 2019). Sama seperti CFI, Indeks ini dianggap sangat baik atau “*very good*” jika  $\geq 0.95$ , baik atau “*good*” jika di antara 0.9 dan 0.95, kurang baik atau “*suffering*” jika di antara 0.8 dan 0.9, dan buruk atau “*bad*” jika  $< 0.8$  (Sarmiento & Costa, 2019).

b. Pendekatan *Multitrait-Multimethod*

Pendekatan *multitrait-multimethod* yang digunakan untuk menguji validitas konstruk pada adaptasi alat ukur VGQ yaitu validitas konvergen dan diskriminan.

- 1) Validitas konvergen terpenuhi jika terdapat korelasi positif yang tinggi di antara dua atau lebih alat ukur yang mengukur konstruk yang sama (Widaman, 2018), atau terdapat korelasi negatif yang tinggi di antara dua atau lebih alat ukur yang mengukur konstruk yang berbeda (divergen; Furr, 2011). Dalam penelitian ini, peneliti akan menggunakan alat ukur *Gaslighting Behaviour Questionnaire* (GBQ; Dickson, Ireland, & Birch, 2023) sebagai alat ukur pengujian validitas konvergen, dan alat ukur *The Rosenberg Self-Esteem Scale* (RSES; Rosenberg, 1965) sebagai alat ukur pengujian validitas divergen.
- 2) Validitas diskriminan terpenuhi jika tidak terdapat korelasi di antara dua atau lebih alat ukur yang mengukur konstruk yang berbeda (Furr, 2011; Widaman, 2018). Dalam penelitian ini, peneliti akan

menggunakan alat ukur *Big Five Inventory* (BFI; John, Naumann & Soto, 2008) sebagai alat ukur pengujian validitas diskriminan.

Berdasarkan penjelasan di atas, diketahui bahwa pengujian validitas konvergen dan diskriminan didasarkan pada aspek korelasi antara dua alat ukur yang berbeda. Oleh karena itu, berikut disajikan interpretasi koefisien korelasi yang dijadikan sebagai acuan tinggi rendahnya korelasi dalam penelitian ini (Schober, Boer, & Schwarte, 2018).

**Tabel 3.7b.**  
*Kategorisasi Koefisien Korelasi*

Koefisien Korelasi	Interpretasi
0.00-0.10	<i>Negligible Correlation</i>
0.10-0.39	<i>Weak Correlation</i>
0.40-0.69	<i>Moderate Correlation</i>
0.70-0.89	<i>Strong Correlation</i>
0.90-1.00	<i>Very Strong Correlation</i>

#### 4. Reliabilitas

Pengujian reliabilitas dalam penelitian ini akan menggunakan metode konsistensi internal yang didasarkan pada koefisien korelasi antar item dalam satu alat ukur dengan pendekatan *Cronbach's alpha*. Metode ini merupakan metode pengujian reliabilitas yang paling umum digunakan (Furr, 2011), dan dianggap sesuai untuk menguji alat ukur yang memiliki respons item berupa skala Likert (Gliem & Gliem, 2003). Guilford (1956) menginterpretasikan nilai koefisien reliabilitas sebagai berikut:

**Tabel 3.7c.**  
*Kategorisasi Koefisien Reliabilitas*

Koefisien Reliabilitas	Interpretasi
0.80 – 1.00	Sangat Tinggi
0.60 – 0.80	Tinggi
0.40 – 0.60	Sedang
0.20 – 0.40	Rendah
0.00 – 0.20	Sangat Rendah