



UNIVERSITAS INDONESIA

**KONSTRUKSI ALAT UKUR INTEGRITAS DENGAN
PENDEKATAN *POLYTOMOUS ITEM RESPONSE THEORY*:
STUDI DALAM KONTEKS PEKERJAAN**

*(Construction of Integrity Scale
Using Polytomous Item Response Theory Model:
Study In Job Environment)*

TESIS

**M A S I T A H
1006742491**

**FAKULTAS PSIKOLOGI
PROGRAM PASCASARJANA
PROGRAM STUDI ILMU PSIKOLOGI
PEMINATAN TERAPAN PSIKOMETRI
DEPOK, JULI 2012**



UNIVERSITAS INDONESIA

**KONSTRUKSI ALAT UKUR INTEGRITAS DENGAN
PENDEKATAN *POLYTOMOUS ITEM RESPONSE THEORY*:
STUDI DALAM KONTEKS PEKERJAAN**

*(Construction of Integrity Scale
Using Polytomous Item Response Theory Model:
Study In Job Environment)*

TESIS

**Untuk memenuhi persyaratan memperoleh gelar Magister
Psikologi Terapan**

**M A S I T A H
1006742491**

**FAKULTAS PSIKOLOGI
PROGRAM PASCASARJANA
PROGRAM STUDI ILMU PSIKOLOGI
PEMINATAN TERAPAN PSIKOMETRI
DEPOK, JULI 2012**

HALAMAN PERNYATAAN ORISINALITAS

Tesis ini adalah hasil karya saya sendiri
dan semua sumber baik yang dikutip maupun dirujuk
telah saya nyatakan dengan benar.

Nama : Masitah

NPM : 1006742491

Tanda Tangan



Tanggal : 07 Juli 2012

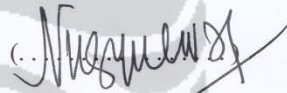
HALAMAN PENGESAHAN

Tesis ini diajukan oleh :
Nama : Masitah
NPM : 1006742491
Program Studi : Ilmu Psikologi
Peminatan : Terapan Psikometri
Judul Tesis : Konstruksi Alat Ukur Integritas Dengan Pendekatan
Polytomous Item Response Theory: Studi Dalam
Konteks Pekerjaan

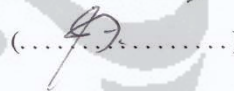
Telah berhasil dipertahankan di hadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Magister Terapan pada Program Studi Psikologi Terapan Psikometri, Fakultas Psikologi, Universitas Indonesia.

DEWAN PENGUJI

Pembimbing 1: Dr. Nugaan Yulia Wardani Siregar

(...  ...)

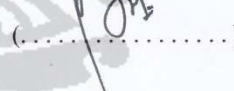
Pembimbing 2: Dr. Guritnaningsih Santoso

(...  ...)

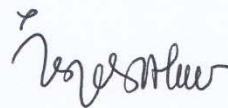
Penguji 1 : Jahja Umar, Ph.D.

(...  ...)

Penguji 2 : Dr. Ir. Bastari, M.A.

(...  ...)

Ketua Program Studi
Ilmu Psikologi Peminatan Terapan



Dr. Alice Salendu, MBA, M.Psi.
NUP. 0806050140

Dekan fakultas Psikologi
Universitas Indonesia





Dr. Witman Dahlan Mansoer., M. Org. Psy.
NIP. 19490403 197603 1 002

Ditetapkan di : Fakultas Psikologi Universitas Indonesia
Tanggal : 07 Juli 2012

UCAPAN TERIMA KASIH

Segala puji syukur saya panjatkan kepada Allah SWT yang telah memberi kesempatan menyelesaikan tesis ini. Penulisan tesis ini diselesaikan dalam rangka memenuhi persyaratan dalam menyelesaikan program Magister Psikologi Terapan Psikometri pada Fakultas Psikologi Universitas Indonesia. Saya menyadari banyak pihak yang telah memberi bantuan, bimbingan serta dukungan hingga tesis ini bisa selesai. Oleh karena itu, saya mengucapkan terima kasih kepada:

1. Bapak Dr. Wilman Dahlan Mansoer, M.Org.Psy. selaku Dekan Fakultas Psikologi UI yang telah mengizinkan saya menempuh jenjang studi di Magister Terapan Psikometri.
2. Ibu Dr. Nugaan Yulia Wardani Siregar, S.E., S.Psi., M.Psi. dan Ibu Dr. Guritnaningsih Santoso selaku dosen pembimbing yang telah bersedia menyediakan waktu, tenaga dan pikirannya dalam mengarahkan saya untuk menyusun tesis ini.
3. Guru-guru terbaik lainnya di Psikometri: Bapak Jahja Umar, Ph.D., Bapak Dr. Ir. Bastari, M.A., Bapak Hari Setiadi, Bpk. Agung P. Utomo, Ibu Wiwiek Arumwati, Mas Aries Yulianto, S.Psi., M.Si. dan Mas Andi Koentary, S.Psi., M.Si.
4. Istimewa untuk Mama, Prof. Dr. Julia Reveny, Apt., dan Papa Drs. M. Ali Musri S., M.Si. atas dukungan utama yang begitu berharga.
5. Mas Bina Ciptadi, S.Psi, M.Si. yang telah menyediakan waktu dan pikiran untuk membantu saya dalam proses penyelesaian tesis ini.
6. Rekan-rekan Magister Terapan Psikologi Universitas Indonesia 2010, khususnya Psikometri: Mas Eko Nugroho, S.Psi. dan Mas Barly Rahim, S.Psi.
7. Partisipan penelitian serta semua pihak yang telah memberikan dukungan dan perhatian namun tidak bisa disebutkan satu per satu.

Semoga yang tersaji dalam tesis ini bermanfaat bagi semua pihak, khususnya calon-calon *psychometrician* Indonesia di masa depan.

Depok, 07 Juli 2012

Penulis

masitah@ymail.com

**HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI
TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS**

Sebagai sivitas akademik Universitas Indonesia, saya yang bertanda tangan di bawah ini:

Nama : Masitah
NPM : 1006742491
Program Studi : Ilmu Psikologi
Peminatan : Terapan Psikometri
Fakultas : Psikologi
Jenis karya : Tesis

demikian pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Indonesia **Hak Bebas Royalti Noneksklusif (*Non-exclusive Royalty-Free Right*)** atas karya ilmiah saya yang berjudul :

**“KONSTRUKSI ALAT UKUR INTEGRITAS
DENGAN PENDEKATAN *POLYTOMOUS ITEM RESPONSE THEORY*:
STUDI DALAM KONTEKS PEKERJAAN”**

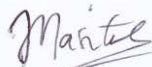
beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti Noneksklusif ini Universitas Indonesia berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan memublikasikan tugas akhir saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Depok

Pada tanggal : 07 Juli 2012

Yang menyatakan



(Masitah)

ABSTRAK

Nama : Masitah
Program Studi : Ilmu Psikologi
Peminatan : Terapan Psikometri
Judul : Konstruksi Alat Ukur Integritas Dengan Pendekatan
Polytomous Item Response Theory: Studi Dalam Konteks
Pekerjaan

Integritas merupakan kekuatan karakter yang mempengaruhi kesehatan mental, kesejahteraan psikologis dan keefektifan hubungan interpersonal. Integritas sangat dibutuhkan dalam dunia pekerjaan terutama dalam hal promosi. Namun, penelitian mengenai integritas masih kurang mendapat perhatian. Alat ukur integritas lebih banyak dikembangkan di luar negeri sehingga kurang sesuai dengan konteks masyarakat Indonesia. Selain itu, umumnya alat ukur integritas dikembangkan menggunakan pendekatan klasik yang memiliki beberapa kelemahan.

Penelitian ini mengembangkan alat ukur integritas menggunakan pendekatan *polytomous Item Response Theory* (IRT) dengan menerapkan *Rating Scale Model* (RSM). Alat ukur integritas yang dikembangkan dalam penelitian ini melibatkan 1210 pekerja di Indonesia. Hasil uji coba menunjukkan bahwa alat ukur integritas (26 *item*) terbukti reliabel ($\alpha=0.94$) dan valid. Hasil uji coba juga menunjukkan bahwa alat ukur integritas ini memenuhi asumsi unidimensionalitas.

Hasil pengujian dengan menerapkan RSM menunjukkan bahwa alat ukur integritas ini memiliki model yang *fit*. Dari 26 *item*, terdapat satu *item* yang tidak *fit*, sehingga *item* tersebut dikeluarkan. Hasil pengujian kembali terhadap 25 *item* menunjukkan bahwa model *fit*, dan seluruh *item fit* mengukur integritas. Analisis menggunakan *differential item functioning* (DIF) menunjukkan 1 *item* memiliki bias respon berdasarkan jenis kelamin. Dengan demikian, *item* yang dipertahankan dalam alat ukur integritas ini berjumlah 24 *item*.

Kata kunci :

Integritas, *Rating Scale Model*, *Polytomous Item Response Theory*

ABSTRACT

Name : Masitah
Study Program : Psychology
Specialization : Applied Psychometrics
Title : Construction of Integrity Scale Using Polytomous Item Response Theory Model: Study In Job Environment

Integrity is a strength of character that affects mental health, psychological well-being and improve interpersonal relationships. Various studies have shown that integrity is essential in the job environment, particularly with regard to their promotion issue. Unfortunately, research on integrity still received little attention and there is no standardized measurement for it. Integrity scale was developed overseas and has not adapted to the Indonesian cultural context. Moreover, the scale development is generally performed with classical theory approach, which has some drawbacks. Therefore, this study develops an integrity scale using polytomous Item Response Theory approach (IRT) by applying the Rating Scale Model (RSM). This study involving 1210 workers in Indonesia.

The pilot study results showed that the integrity scale (with 26 items) is a reliable measure ($\alpha = 0.94$) and valid. The pilot study results also showed that the integrity scale satisfies unidimensionality assumptions.

The test results using the RSM showed that the integrity scale had a fit model. Of the 26 items, there is one item that does not fit, so the item was issued. The second test results for the remaining 25 items showed that they fit the model and all the items were fit to measure integrity. Analysis using differential item functioning (DIF) showed one items have a response bias based on gender. Thus, there are 24 items remaining in the scale.

Key words :

Integrity, Rating Scale Model, Polytomous Item Response Theory

DAFTAR ISI

HALAMAN SAMPUL	
HALAMAN JUDUL	
HALAMAN PERNYATAAN ORISINALITAS	
LEMBAR PENGESAHAN	
UCAPAN TERIMA KASIH	
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI	
ABSTRAK	
<i>ABSTRACT</i>	
DAFTAR ISI	
DAFTAR TABEL	

BAB 1 PENDAHULUAN

1.1. Latar Belakang Masalah	1
1.2. Rumusan Masalah	7
1.3. Tujuan Penelitian	7
1.4. Manfaat Penelitian	8
1.5. Sistematika Penulisan	8

BAB 2 LANDASAN TEORI

2.1. Integritas	
2.1.1. Definisi Integritas	9
2.1.2. Proses Tercapainya Integritas	12
2.1.3. Faktor Pembentuk Integritas	13
2.1.4. Sejarah Perkembangan Tes Integritas	14
2.2. Pendekatan <i>Item Response Theory</i> (IRT)	
2.2.1. Kelemahan Pendekatan Klasik	17
2.2.2. Keunggulan Pendekatan IRT	18
2.2.3. Asumsi Dalam Pendekatan IRT	21
2.2.4. <i>Rating Scale Model</i> (RSM) Dalam Pendekatan IRT	22
2.2.5. Analisis <i>Differential Item Functioning</i> (DIF)	22

BAB 3 METODE PENELITIAN

3.1. Spesifikasi Alat Ukur	24
3.2. Partisipan Penelitian	26
3.3. Teknik Sampling	27
3.4. Prosedur Penelitian	
3.4.1. Tahap Persiapan	27
3.4.2. Tahap Uji Coba	28
3.4.3. Tahap Pelaksanaan	36
3.5. Teknik Analisis Data Dengan Pendekatan IRT	36

BAB 4 HASIL DAN ANALISA DATA	
4.1. Gambaran Umum Partisipan Penelitian	38
4.2. Hasil Pengujian Alat Ukur Integritas Menggunakan <i>Polytomous</i> IRT	
4.2.1. Pengujian <i>Person Fit</i>	39
4.2.2. Pengujian <i>Item Fit</i>	40
4.2.3. Pengujian Kembali <i>Person Fit</i> dan <i>Item Fit</i>	41
4.2.4. Analisis <i>Differential Item Functioning</i> (DIF)	43
4.2.5. Deskripsi Karakteristik Setiap <i>Band-Scale</i> Alat Ukur Integritas	45
BAB 5 KESIMPULAN, DISKUSI DAN SARAN	
5.1. Kesimpulan	48
5.2. Diskusi	50
5.3. Saran	52
DAFTAR PUSTAKA	54
LAMPIRAN	59



DAFTAR TABEL

Tabel 2.1.	Perbandingan Pendekatan Klasik dan Pendekatan Modern (IRT)	19
Tabel 3.1.	Cara Skoring Alat Ukur Integritas	24
Tabel 3.2.	Kisi-kisi Alat Ukur Integritas 40 <i>Item</i>	25
Tabel 3.3.	Gambaran Partisipan Elisitasi	28
Tabel 3.4.	Indeks Kecocokan <i>First Order</i> CFA 40 <i>Item</i>	30
Tabel 3.5.	Kisi-Kisi Alat Ukur Integritas 26 <i>Item</i>	31
Tabel 3.6.	Indeks Kecocokan <i>First Order</i> CFA 26 <i>Item</i>	32
Tabel 3.7.	Hasil <i>First Order</i> CFA Alat Ukur Integritas 26 <i>Item</i>	33
Tabel 3.8.	Indeks Kecocokan <i>Second Order</i> CFA 26 <i>Item</i>	34
Tabel 3.9.	Hasil <i>Second Order</i> CFA Alat Ukur Integritas 26 <i>Item</i>	35
Tabel 4.1.	Demografis Penyebaran Partisipan Berdasarkan Jenis Kelamin, Status Pekerjaan dan Jabatan	38
Tabel 4.2.	Demografis Penyebaran Partisipan Berdasarkan Usia	39
Tabel 4.3.	Nilai <i>Infit Mean Square</i> 26 <i>Item</i>	41
Tabel 4.4.	Nilai <i>Infit Mean Square</i> 25 <i>Item</i>	42
Tabel 4.5.	Hipotesis Pengujian	43
Tabel 4.6.	Perbandingan Estimasi <i>Item</i> Berdasarkan Jenis Kelamin	44
Tabel 4.7.	Deskripsi Karakteristik Setiap <i>Band-Scale</i> Alat Ukur Integritas	46

BAB 1

PENDAHULUAN

1.1. Latar Belakang Masalah

Integritas merupakan kekuatan karakter yang berlaku di segala aspek kehidupan seperti pendidikan, penelitian dan pekerjaan. Sebagaimana yang dikemukakan oleh Peterson dan Seligman (2004), integritas bersifat universal dan dibutuhkan di berbagai peran. Hal ini didukung oleh Schlenker, Miller dan Johnson (2009) yang menyatakan bahwa integritas telah dikembangkan dalam setiap lapisan masyarakat karena berdampak penting bagi hubungan sosial. Penjelasan Deci dan Ryan (2000) mengenai kontribusi integritas terhadap kesehatan mental, kesejahteraan psikologis serta keefektifan hubungan interpersonal juga memperkuat pernyataan tersebut. Khalil (2004) menambahkan bahwa Integritas bukan sifat bawaan, tetapi berkaitan dengan apa yang disetujui individu secara implisit atau eksplisit.

Integritas sendiri didefinisikan Rogers (1961) sebagai kondisi yang terjadi ketika individu mampu menerima serta bertanggung jawab terhadap perasaan, niat, komitmen dan perilaku, termasuk mampu mengakui kondisi itu kepada orang lain bila diperlukan. Carter (1996) memperkuat definisi tersebut dengan menyatakan bahwa individu yang memiliki integritas bersedia menanggung konsekuensi dari keyakinannya, meskipun hal itu sulit dilakukan, konsekuensinya tidak menyenangkan, bahkan tidak mendapat kerugian jika tidak mempertahankan integritasnya. Menurut Khalil (2004), integritas ditegaskan ketika individu memiliki pilihan untuk tidak menghormati atau melanggar komitmen dan janji yang ia buat sendiri karena pelanggaran terhadap komitmen/janji mendatangkan rasa malu terhadap dirinya sendiri.

Beberapa tahun terakhir banyak yang tertarik meneliti integritas sebagai kekuatan karakter. Deci dan Ryan (2000) menyatakan bahwa manusia memiliki kebutuhan psikologis terhadap integritas. Selain itu, Miller dan Schlenker (2011) menjelaskan bahwa integritas dalam diri individu berkaitan dengan pandangan hidup yang lebih positif, orientasi yang lebih positif terhadap orang lain, spiritualitas yang lebih tinggi serta minimnya tindakan irasional.

Dunia kerja adalah salah satu yang paling menuntut pentingnya integritas. Alat ukur integritas diperlukan dalam membina sumber daya manusia di dalam perusahaan karena sumber daya manusia dituntut dapat bekerja secara produktif untuk mendukung tujuan-tujuan yang ingin dicapai perusahaan. Namun kenyataan menunjukkan bahwa sumber daya manusia di perusahaan, yang disebut pekerja, seringkali justru berperilaku sebaliknya, yaitu perilaku kontraproduktif seperti korupsi dan mencuri waktu untuk keperluan pribadi. Perilaku kontraproduktif menunjukkan rendahnya kualitas pekerja pada suatu perusahaan. Hal ini diketahui berdasarkan *The Global Competitiveness Report 2011-2012* dari *World Economic Forum* yang merilis bahwa peringkat daya saing Indonesia menurun dari 44 menjadi 46 dari 142 negara yang di survei. Peringkat ini menunjukkan bahwa profesionalisme, kehadiran dan kuantitas pekerjaan pekerja Indonesia masih tergolong rendah. Untuk itu diperlukan unsur penting yang dapat meningkatkan efektivitas sikap serta kualitas dalam bekerja. Istilah yang paling menggambarkan unsur tersebut secara tepat adalah integritas (Impelman, 2006). Hal ini sesuai dengan pernyataan Marchus dan Schuler (2004) bahwa pekerja yang memiliki level integritas tinggi akan menghasilkan produktivitas kerja yang baik.

Untuk mendapatkan informasi mengenai integritas yang dimiliki pekerja, perusahaan perlu melakukan pengukuran. Waktu yang tepat mengukur integritas, menurut Mumford (2000) ditentukan berdasarkan tujuan yang diharapkan dari pengukuran. Pengukuran integritas pada waktu rekrutmen memungkinkan perusahaan menyeleksi calon karyawan dengan tingkat integritas yang baik. Hal ini penting karena karyawan yang memiliki integritas rendah dapat merusak keseluruhan kinerja perusahaan, *image* perusahaan, kepercayaan kolega dan pemegang saham, menurunkan performa finansial perusahaan serta mempengaruhi kondisi ekonomi masyarakat secara keseluruhan (Mount, Ilies & Johnson, 2006). Selain itu, pengukuran integritas juga dapat dilakukan pada waktu promosi menjadi pimpinan. Sebagai figur yang mengemban peran dan tanggung jawab lebih besar dalam perusahaan, tuntutan terhadap pemimpin yang memiliki integritas sangat besar (Erhard, Jensen dan Zaffron, 2011). Pimpinan yang memiliki integritas tinggi merupakan solusi mencegah perilaku kontraproduktif karyawan (Wanek, Sackett dan Ones, 2003). Jika karyawan menilai pimpinan

menunjukkan perilaku yang memiliki integritas, karyawan akan mengupayakan perilaku yang sama (Ones dan Viswesvaran, 2001), sehingga interaksi lingkungan kerja menjadi lebih efektif (Palanski dan Yammarino, 2007).

Sebagaimana umumnya variabel psikologis, integritas bersifat laten dan hanya dapat diamati melalui sampel perilaku. Merancang *item-item* yang mampu menggali integritas merupakan tantangan besar yang dihadapi dalam perancangan alat ukur integritas. Kebutuhan akan alat ukur integritas yang valid dan reliabel, khususnya dalam dunia kerja, sudah sedemikian besar dan akan semakin besar (Berry, Sackett dan Wiemann, 2007; Ones, Viswesvaran dan Schmidt, 2003; Sackett, Burris dan Callahan, 1989).

Selama beberapa dekade terakhir, penggunaan tes-tes integritas dalam konteks pekerjaan semakin berkembang di negara lain (Erhard, Jensen dan Zaffron, 2011; Egberink dan Veldkamp, 2007; Palanski dan Yammarino, 2007; Impelman, 2006; Barrett, 2001; Ones dan Viswesvaran, 2001; US Congressional Office of Technology Assessment, 1990; Martelli, 1988; Harris, 1987) serta dianggap sebagai salah satu alat ukur yang paling valid dan tidak memberi dampak yang merugikan (Berry, Sackett dan Wiemann, 2007; Ones, Viswesvaran dan Schmidt, 1995; Wanek, Sackett dan Ones, 2003; Hunter, in preparation). Hal ini didukung oleh Ones, Viswesvaran dan Schmidt (1995) yang menyatakan bahwa tes integritas dapat memprediksi performa kerja serta perilaku kontraproduktif dalam pekerjaan. Penelitian Ones (1993) juga menemukan bahwa tes integritas berkorelasi dengan usia dan *gender*. Individu yang berusia lebih muda cenderung lebih kontraproduktif karena kecerobohan dan keinginan coba-coba/eksperimen. Berkaitan dengan *gender*, rata-rata wanita dilaporkan memiliki skor integritas yang lebih tinggi daripada pria.

Integritas merupakan karakteristik positif yang memiliki sifat universal, pekerja di Amerika dan Indonesia sama-sama dituntut memiliki integritas dalam pekerjaannya. Namun kurang tepat jika alat ukur integritas yang dikembangkan pada latar belakang Amerika digunakan untuk mengukur integritas di Indonesia. Terlebih jika hasil pengukuran tersebut akan menjadi dasar keputusan-keputusan penting. Indonesia bahkan memiliki perbedaan mendasar dengan bangsa Asia lainnya. *Self* bagi bangsa Asia umumnya tidak ditekankan sebagai kesatuan yang

independen karena menganggap *self* ditentukan lingkungan dan konteks sosial (Markus, Kitayama, dan Heiman, 1996), hal ini sesuai dengan penelitian yang dilakukan Suh (2001) terhadap bangsa Korea. Hasil penelitian Iwao (1997) juga menambahkan bahwa partisipan yang berasal dari Jepang tidak mengungkapkan ketidaksetujuan yang mereka rasakan segamblang partisipan yang berasal dari Amerika. Temuan ini mengindikasikan bahwa mempresentasikan diri secara jujur tidak terlalu penting dalam kebudayaan Asia. Tidak berarti bangsa Asia memiliki tingkat integritas yang rendah atau mengkhianati nilai-nilai dan komitmen yang mereka anut, itu hanya indikasi perbedaan latar belakang budaya.

Perbedaan tersebut memunculkan tuntutan yang semakin besar terhadap alat ukur integritas yang dirancang berdasarkan kondisi Indonesia. Alasan tersebut juga menarik minat peneliti untuk merancang alat ukur integritas yang sesuai dengan konteks perusahaan di Indonesia. Alat ukur integritas yang dirancang dalam penelitian ini ditujukan bagi kepentingan promosi pimpinan, khususnya level manager pada perusahaan. Alat ukur ini tepat diberikan kepada partisipan yang telah memiliki pengalaman sebagai pekerja karena *item-item* alat ukur dirancang berdasarkan situasi yang dialami pekerja sehari-hari.

Di Indonesia masih sedikit yang mengembangkan penelitian mengenai alat ukur integritas, salah satunya Permatasari (2011). Tetapi menurut penulis, alat ukur integritas yang ia kembangkan memiliki kelemahan dari sisi teori dan psikometris. Kelemahan di sisi teori terjadi karena Permatasari (2011) merancang alat ukur integritas menggunakan *moral identity theory* dari Blasi (2004) yang tidak membahas integritas secara spesifik. Blasi (2004) adalah tokoh filsafat yang bahasan utamanya moral, bukan kondisi psikologis. Teori lain yang dikemukakan Rogers (1961), seorang psikolog humanistik, tentu lebih mampu menggambarkan integritas dari sisi psikologis. Oleh karena itu, peneliti merancang alat ukur berdasarkan konsep integritas yang diutarakan oleh Rogers (1961).

Kelemahan lain alat ukur integritas rancangan Permatasari (2011) dapat dijelaskan melalui sisi psikometris. Meskipun alat ukur dirancang berdasarkan kondisi Indonesia, pengembangannya menggunakan pendekatan klasik sehingga menghasilkan alat ukur yang bersifat *sample bound*. Hambleton, Swaminathan dan Rogers (1991) menjelaskan bahwa alat ukur yang tergolong *sample bound*

seolah-olah memiliki tingkat kesulitan yang tinggi ketika diberikan pada kelompok subjek yang memiliki kemampuan rendah dan seolah-olah memiliki tingkat kesulitan yang rendah ketika diberikan pada kelompok subjek yang memiliki kemampuan tinggi. Pengembangan alat ukur menggunakan pendekatan klasik juga menyebabkan tidak diketahuinya parameter *item*. Analisis kualitas *item* akan berubah tergantung siapa yang mengerjakan alat ukur tersebut. Alat ukur yang kualitas *item*nya tidak diketahui secara jelas tentunya tidak tepat untuk digunakan sebagai dasar pengambilan keputusan-keputusan penting.

Untuk mengatasi kelemahan psikometris tersebut, penelitian ini menggunakan pendekatan yang lebih modern, akurat dan bebas *sample bound* yaitu *Item Response Theory* (IRT). Kelebihan penggunaan IRT adalah diperoleh karakteristik *item* yang tidak tergantung pada kemampuan individu yang menempuhnya (*sampling invariant*). *Item-item* alat ukur dianalisis menggunakan pendekatan *Rasch Model* agar bisa dipakai berulang-ulang, ini juga menjadi alasan pentingnya memiliki *item-item* yang *sampling invariant*.

Alat ukur integritas yang dirancang dalam penelitian ini menggunakan pendekatan *Item Response Theory* (IRT). Hambleton, Swaminathan dan Rogers (1991) menyatakan bahwa pendekatan IRT dapat diterapkan apabila data memenuhi asumsi *unidimensionality* dan *local independence*. Asumsi *unidimensionality* akan memastikan bahwa hanya ada satu faktor dominan yang mempengaruhi skor partisipan. Pengujian *unidimensionality* diperlukan untuk membuktikan bahwa alat ukur yang dirancang dalam penelitian ini memang hanya mengukur integritas. Sedangkan asumsi *local independence* akan memastikan bahwa kemampuan integritas yang diukur merupakan satu-satunya faktor yang mempengaruhi respon partisipan.

Berkaitan dengan penerapan pendekatan IRT, alat ukur integritas yang dikembangkan dalam penelitian ini menggunakan model pengukuran *polytomous IRT*. Model ini ditentukan berdasarkan kategori respon serta jumlah parameter *item* yang terlibat. Mengingat alat ukur dirancang dalam bentuk skala *likert* yang memiliki empat kategori respon, model politomi yang paling tepat adalah *Rating Scale Model* (Embretson dan Reise, 2000). Selain itu, penerapan pendekatan IRT akan memberi informasi yang akurat mengenai kualitas *item-item* yang dirancang.

Keputusan yang diambil berdasarkan alat ukur integritas yang dirancang menggunakan pendekatan IRT menjadi lebih reliabel. Kelebihan IRT berikutnya adalah mampu mendeteksi *Different Item Functioning* (DIF) yang dapat menjelaskan apakah *item* yang sama akan memiliki kecenderungan berbeda ketika ditempuh oleh kelompok partisipan yang berbeda, namun memiliki tingkat integritas yang sama. Peneliti menguji DIF berdasarkan jenis kelamin karena Ones (1993) menemukan bahwa *gender* berpengaruh terhadap pengukuran integritas.

Mengingat masih sedikit alat ukur integritas yang dirancang berdasarkan konteks pekerjaan di Indonesia, maka penelitian ini berupaya merancang alat ukur integritas untuk kepentingan promosi level manager di perusahaan. Meskipun alat ukur dirancang menggunakan teori Rogers (1961) yang mampu menggambarkan integritas dari sisi psikologis, belum pernah ada yang melakukan analisis faktor untuk membuktikan bahwa teori integritas tersebut memiliki konstruk yang valid. Untuk itu, alat ukur integritas dalam penelitian ini dikembangkan menggunakan metode *confirmatory factor analysis* untuk membuktikan bahwa teori integritas yang dikemukakan Rogers (1961) memenuhi asumsi unidimensionalitas. Unidimensionalitas sangat penting diteliti untuk membuktikan bahwa alat ukur yang dirancang dalam penelitian ini memang hanya mengukur integritas.

Selain itu, dilakukan juga pengujian untuk membuktikan apakah alat ukur integritas yang dirancang dalam penelitian ini terdiri dari *item-item* yang *fit* karena *item-item* yang *fit* akan memberikan hasil ukur yang valid. Setelah asumsi unidimensionalitas terpenuhi dan alat ukur terbukti memiliki *item-item* yang *fit*, pengujian terakhir dilakukan untuk mendeteksi kemungkinan adanya *item* yang memiliki kecenderungan bias respon. Ini merupakan kelebihan penggunaan pendekatan IRT, tetapi semuanya tidak pernah diuji pada pendekatan klasik.

Terkait tujuan untuk merancang alat ukur integritas bagi kepentingan promosi level manager pada perusahaan, populasi penelitian mencakup pekerja yang bekerja pada perusahaan di Indonesia. 1210 pekerja yang berpartisipasi dalam penelitian ini berasal dari Pulau Sumatera dan Pulau Jawa, memiliki status pekerjaan PNS dan swasta serta mencakup jenis kelamin laki-laki dan perempuan.

1.2. Rumusan Masalah

Alat ukur yang mampu menggali integritas pekerja berdasarkan latar belakang kondisi Indonesia sangat dibutuhkan, terutama yang terkait tujuan promosi level manager pada perusahaan. Alat ukur integritas untuk pekerja Indonesia sudah pernah dirancang tetapi memiliki kelemahan di sisi konstruksi teori dan pengembangan alat ukur. Alat ukur tersebut tidak dirancang menggunakan teori yang membahas integritas dari sisi psikologis. Selain itu, alat ukur tersebut dikembangkan menggunakan pendekatan klasik yang bersifat *sample bound* dan tidak memberi informasi *item* secara akurat. Untuk mengatasi permasalahan tersebut, penelitian ini merancang alat ukur menggunakan teori yang membahas integritas dari sisi psikologis serta dikembangkan menggunakan pendekatan modern (IRT) yang mampu memberikan hasil ukur lebih akurat. Sebagai upaya menjawab permasalahan-permasalahan tersebut, pertanyaan dalam penelitian ini dirumuskan sebagai berikut:

- a. Apakah konstruk alat ukur integritas yang dirancang oleh Rogers (1961) sudah valid?
- b. Apakah alat ukur integritas ini memenuhi asumsi unidimensionalitas ketika menggunakan metode *confirmatory factor analysis*?
- c. Apakah alat ukur integritas ini terdiri dari *item-item* yang *fit* mengukur integritas ketika menggunakan penerapan model *polytomous IRT*?
- d. Apakah *item-item* dalam alat ukur integritas ini mengandung bias respon yang dapat dideteksi melalui *differential item functioning (DIF)*?

1.3. Tujuan Penelitian

Penelitian ini dilakukan dengan tujuan untuk memperoleh alat ukur integritas yang sesuai bagi kepentingan promosi level manager pada perusahaan di Indonesia.

1.4. Manfaat Penelitian

Sejalan dengan tujuan penelitian, tes yang terdiri dari *item-item* yang bebas dari bias respon menjadikannya mungkin untuk melakukan perancangan norma alat ukur integritas. Selain itu, penggunaan *polytomous IRT* dapat menjadi alternatif pengembangan alat ukur. Dengan demikian, penelitian ini memperkaya khasanah ilmu psikologi secara umum, psikometri secara khusus, serta psikologi industri-organisasi.

Penelitian ini juga memberi manfaat dalam bentuk pengembangan alat ukur integritas yang berguna bagi kepentingan perusahaan. Membantu perusahaan mendapatkan gambaran integritas sebagai pendukung pengambilan keputusan mengenai kinerja karena alat ukur ini sudah terbukti reliabel, valid dan terstandar. Perusahaan juga dapat menentukan kriteria integritas yang perlu diutamakan pada level kinerja tertentu. Hal ini tentu sangat bermanfaat dan memudahkan pelaksanaan promosi di perusahaan.

1.5. Sistematika Penulisan

Tesis ini terdiri dari lima bagian. Selanjutnya penulisan pada Bab 2 berisi tinjauan pustaka mengenai teori-teori yang mendukung penelitian. Pada Bab 3 dijelaskan metode penelitian yang mencakup spesifikasi alat ukur, partisipan penelitian, teknik sampling serta prosedur penelitian. Kemudian pada Bab 4 diuraikan analisis hasil dari penelitian yang dilakukan. Pada Bagian terakhir, yaitu Bab 5, dipaparkan kesimpulan dari penelitian yang telah dilakukan dalam upaya menjawab rumusan masalah serta beberapa hal yang perlu didiskusikan termasuk saran bagi penelitian selanjutnya terkait pengembangan alat ukur integritas.

BAB 2

LANDASAN TEORI

Bab ini menguraikan tinjauan teoritis sebagai acuan pembahasan masalah. Penjelasan diawali dengan studi literatur mengenai integritas serta memaparkan pendekatan yang digunakan untuk mengembangkan alat ukur integritas.

2.1. Integritas

Integritas diadaptasi dari bahasa Latin, *integritas*, yang berarti utuh, tak tersentuh, lengkap dan menyeluruh. Integritas merupakan salah satu karakter yang menyusun *strength of courage*. Integritas sebagai karakter disusun oleh *virtue honesty* dan *virtue authenticity* (Peterson dan Seligman, 2004). Jadi integritas bukan hanya kejujuran, meskipun ketika mendengar ungkapan ‘individu yang berintegritas’ umumnya langsung terpikir seseorang yang jujur (Carter, 1996). Konsep integritas dalam psikologi diutarakan psikolog humanistik Rogers (1961).

Peterson dan Seligman (2004) menambahkan bahwa integritas dapat dipahami menggunakan analisa psikobiografi. Individu yang memiliki integritas adalah individu yang memiliki kesempatan mengejar cita-cita berdasarkan ketertarikan pribadi. Kesempatan untuk memenuhi minat terdalam menjadikan individu memahami nilai-nilai dalam dirinya serta berhasil mencapai prestasi maksimal dalam hidupnya.

2.1.1. Definisi Integritas

Rogers (1961) mendefinisikan integritas sebagai perasaan yang dialami, disadari dan diakui individu serta mampu mengkomunikasikannya jika diperlukan. Ahli moral seperti Blasi (2004) serta Palanski dan Yamarino (2007) menambahkan definisi integritas sebagai konsistensi pikiran, emosi, kata-kata serta tindakan yang stabil sepanjang waktu dan situasi.

Penelitian ini mengembangkan alat ukur integritas menggunakan teori yang dikemukakan oleh Rogers (1961) karena memberi batasan definisi integritas secara lengkap dan jelas. Integritas sebagai kesatuan dalam *self* merupakan konsistensi pada komitmen yang telah dipilih individu. Komitmen mencerminkan

prinsip, motivasi, keyakinan, perasaan, logika, tindakan, identitas dan regulasi diri. Konsistensi menunjukkan kesatuan terhadap semua hal tersebut. Perilaku yang terlihat merupakan wujud tanggung jawab individu terhadap orang lain dan norma, serta bukan karena pertimbangan untung dan rugi.

Teori mengenai integritas berkembang akibat ketidakpuasan terhadap teori perkembangan moral Kohlberg yang kurang memberi tempat pada peran *self* dalam menerima nilai-nilai moral (Blasi, 2004). Kohlberg (1976) yang menyatakan bahwa pemahaman moral dituntun oleh perkembangan logika, tidak mampu menjawab mengapa pemahaman nilai moral tidak menjamin individu melakukan tindakan moral. Individu memiliki intuisi mengenai apa yang benar dan salah, namun sering tidak menyadari alasannya dan kesulitan menjelaskan tindakannya. Dalam hal inilah *self* mampu menjelaskan mengapa pengetahuan mengenai nilai moral tidak menjamin seseorang melakukan tindakan moral. *Self* merupakan kesatuan pengalaman, nilai-nilai, *beliefs* dan *trait* yang aktif menyaring dan menyesuaikan nilai-nilai moral untuk dicocokkan dalam diri individu. Bila pemahaman nilai moral tidak sesuai dengan *self*, individu tidak akan melakukan tindakan moral tertentu. Lapsey dan Hill (2009) menggambarkan individu yang memiliki integritas sebagai orang yang dinilai memiliki karakter baik dan dijadikan contoh moral seperti Galileo Galilei, Confusius, Mahatma Gandhi, dan Muhammad Hatta.

Integritas mencakup tanggung jawab dan identitas moral (Puka, 2004). Tanggung jawab berisi hasrat, komitmen serta perasaan individu terhadap serangkaian norma dan hubungan dengan orang lain. Identitas moral diartikan sebagai kesatuan nilai-nilai dan komitmen yang menyatu dalam *self*. Keduanya menciptakan konsistensi antara perasaan dan pikiran dengan tindakan serta konsistensi ucapan dan perilaku di segala situasi. Jadi integritas muncul apabila nilai-nilai moral telah terintegrasi dalam *self* sehingga terjadi koherensi antara nilai-nilai moral yang dipahami dengan *beliefs*, tindakan, komitmen dan perkataan. Tidak ada pedoman kapan atau pada usia berapa integrasi dipastikan terjadi. Namun ketika terjadi akan menetap dalam diri individu karena sudah menjadi identitas diri yang menuntun setiap perasaan, perkataan dan perbuatan.

Ada kalanya timbul pertentangan *beliefs*, contohnya ketika mengetahui teman melakukan pencurian. Terjadi pertentangan antara keyakinan harus berkata benar dengan keyakinan harus menolong teman ketika diminta memberi kesaksian apakah temannya mencuri. Individu yang memiliki integritas dipercaya mampu mempertimbangkan tindakan yang paling tepat. Karena bila individu melakukan tindakan yang bertentangan dengan *beliefs* dasar, muncul perasaan tidak nyaman yang merusak identitas dirinya (Puka, 2004).

Integritas digerakkan oleh diri sendiri (*self-directed*) (Puka, 2004). Blasi (2004) menyebut tidak ada pertimbangan untung dan rugi terhadap diri sendiri sebagai komitmen moral tak bersyarat. Hal ini merupakan inti memahami integritas. Individu yang memiliki integritas punya serangkaian tujuan, aturan serta standar tentang benar dan salah yang mengarahkan sikap dan alasannya melakukan tindakan (Peterson dan Seligman, 2004). Tujuan hidup serta standar tentang benar dan salah menjadikan individu memiliki kejujuran dan keteguhan.

Schlenker, Miller dan Johnson (2009) menyatakan bahwa derajat kejujuran individu bisa berbeda apabila dihadapkan pada kelompok individu yang berbeda, tergantung tujuan ingin menampilkan diri seperti apa (*self-presentational goal*). Namun hal itu tidak terjadi pada individu yang memiliki integritas tinggi, karena ia merasa tidak memiliki kepentingan apapun (contohnya ingin dipuji) dan hanya melakukan sesuatu yang benar berdasarkan standar dan batasannya sendiri.

Peterson dan Seligman (2004) serta Schlenker, Miller dan Johnson (2009) juga menjelaskan bahwa individu yang memiliki integritas memperlakukan orang lain seperti ia ingin diperlakukan, tetapi tidak mengharapkan timbal balik. Menolong orang lain dilakukan karena alasan *altruistic* (fokus pada kesejahteraan orang lain), bukan untuk alasan egois (penghargaan pribadi seperti meningkatkan karir, pengakuan sosial atau menghindari hukuman).

Integritas bukanlah sifat bawaan yang sudah dimiliki individu sejak lahir. Berikut adalah penjelasan mengenai proses tercapainya Integritas.

2.1.2. Proses Tercapainya Integritas

Proses tercapainya integritas menurut Petrick dan Quinn (2000) terjadi melalui tahapan-tahapan sebagai berikut :

a. *Moral Awareness* (kesadaran moral)

Moral awareness adalah kemampuan memahami dan kepekaan terhadap isu-isu etis yang relevan yang patut menjadi pertimbangan dalam membuat pilihan yang akan memiliki dampak signifikan terhadap orang lain. *Moral awareness* dibentuk melalui persepsi dan sensitivitas terhadap etika. Yang dimaksud dengan persepsi adalah kemampuan untuk melihat, mengenali atau menemukan fitur etis dari suatu situasi. Sedangkan sensitivitas terhadap etika adalah kemampuan untuk menilai kepentingan yang terkait dengan fitur etika dari sebuah situasi.

b. *Moral Deliberation* (pertimbangan moral)

Moral deliberation adalah kemampuan melakukan analisa kritis terhadap faktor penyebab dan pilihan moral yang dimiliki untuk mendapatkan keputusan yang masuk akal bagi standar yang penting di masa depan. *Moral deliberation* terdiri dari analisis etika dan resolusi etika. Analisis etika merupakan langkah rasional berdasarkan argumentasi moral yang dirancang untuk mengidentifikasi, menafsirkan dan mempertimbangkan penyebab utama masalah moral dan sumber daya kunci untuk penyelesaian masalah etika. Sedangkan resolusi etika adalah langkah rasional membuat suatu keputusan.

c. *Moral Character* (karakter moral)

Moral character adalah kemampuan untuk siap bertindak etis. *Moral character* dikuatkan dengan melatih nilai-nilai seperti moral, sosial, emosional dan politik.

d. *Moral Conduct* (tindakan moral)

Moral conduct berarti individu melakukan tindakan yang dapat dipertanggungjawabkan secara berkelanjutan. Jadi untuk mencapai integritas, individu harus melatihnya dengan sengaja.

2.1.3. Faktor Pembentuk Integritas

Faktor pembentuk Integritas yang menjadi landasan pengembangan alat ukur integritas dalam penelitian ini berasal dari konsep yang dikemukakan oleh psikolog humanistik Rogers (1961). Faktor-faktor tersebut adalah :

1. Jujur

Jujur berarti tidak mengingkari hati nurani, berbicara dan bertindak sesuai nilai-nilai pribadi yang dipegang teguh serta menjaga komitmen terhadap orang lain (Puka, 2004). Individu dikatakan jujur apabila menerima dan mampu bertanggung jawab atas perasaan serta perilaku sebagaimana adanya. Meski memegang erat prinsip kejujuran, namun dalam situasi yang penuh tipu muslihat dan harus menghadapi orang yang tidak jujur, individu yang memiliki integritas tinggi akan bertindak dan menegur dengan mempertimbangkan berbagai hal serta tidak menyakiti. Karenanya individu yang memiliki integritas lebih dihormati daripada disukai (Schlenker, Miller dan Johnson, 2009).

2. Teguh

Teguh artinya tidak menyalahi prinsip dalam menjalankan kewajiban, tidak dapat disuap atau diajak melakukan perbuatan curang meskipun ada godaan materi atau dorongan dari orang lain. Peterson dan Seligman (2004) menyatakan ada dua situasi yang membuat individu dikatakan memiliki keteguhan. Pertama ketika harus menghadapi situasi yang tidak menguntungkan seperti pertentangan serta ketidakpercayaan dan yang kedua ketika harus menghadapi kesulitan atau keadaan bahaya. Keteguhan yang memiliki integritas dinilai muncul dalam situasi pertama karena integritas melibatkan suatu pilihan antara beberapa tindakan atau cara. Keteguhan menjalankan prinsip berbeda dengan kefanatikan yang rela membunuh orang tak bersalah guna mencapai tujuan, yang baginya merupakan prinsip hidup. Individu dengan integritas tinggi memiliki kebijakan yang ditujukan bukan hanya untuk kelompok atau golongannya, tetapi kepentingan manusiawi yang lebih besar (Schlenker, Miller dan Johnson, 2009).

3. Memiliki *self-control* yang kuat

Self-control didefinisikan sebagai kemampuan individu dalam mengontrol atau memantau respon agar sesuai dengan tujuan hidup dan standar moral yang dimiliki. Untuk bisa memperlakukan orang lain, bahkan orang yang sesungguhnya tidak disukai secara baik, individu harus memiliki *self-control* yang kuat. Kemampuan individu mengontrol atau memantau respon, selain penting untuk menjaga agar perilaku tetap sesuai dengan tujuan hidup dan standar moral, juga penting untuk berhubungan dengan orang lain (Peterson dan Seligman, 2004). Individu yang memiliki *self-control* kuat tidak mudah memperlihatkan reaksi emosional lewat ucapan maupun sikap badan. Individu yang memiliki *self-control* terlihat tenang bila dihadapkan pada stimulus yang memancing emosi, hal ini menjadikan orang lain lebih nyaman berhubungan dengan mereka.

4. Memiliki *self-esteem* yang tinggi

Self-esteem adalah kepercayaan bahwa individu mampu berperilaku sesuai nilai moral yang diyakini. Blasi (dalam Power, 2004) menyebut *self-esteem* sebagai perasaan positif individu bahwa dirinya bermoral dan mampu menjalankan prinsip-prinsip moral. Karena berasal dari *beliefs*, Mecca, Smelser dan Vasconcellos (dalam Power, 2004) meyakini bahwa harga diri mampu meningkatkan perilaku yang baik dan keteguhan.

Faktor merupakan suatu kesatuan utuh (koherensi) yang artinya seluruh faktor tersebut tidak bisa dipecah-pecah karena saling terkait satu sama lain. Individu dikatakan memiliki integritas apabila memiliki seluruh faktor tersebut.

2.1.4. Sejarah Perkembangan Tes Integritas

Penelitian seputar integritas telah dimulai sekitar tahun 1900 namun masih diragukan oleh banyak orang. Salah satunya adalah Gough (1990) yang mengembangkan tes psikologis berbentuk *paper and pencil* untuk mengetahui potensi integritas yang dimiliki karyawan, instrumen itu dinamakan *Personnel Reaction Blank*. Alat ukur tersebut dikembangkan dari *Delinquency Scale*, yang

merupakan bagian dari *California Psychological Inventory*. Kemudian *Delinquency Scale* diubah nama menjadi *The Socialization Scale* (Ones, 1993).

Alasan utama lahirnya ketertarikan untuk mengembangkan tes integritas dalam konteks pekerjaan adalah tindakan kontraproduktif yang dilakukan para pekerja. *American Management Association* melaporkan bahwa kerugian akibat tindakan kontraproduktif karyawan besarnya 10 kali lipat dibandingkan kerugian akibat pencurian dan kejahatan jalanan lain (Ones, 1993). Hal tersebut dapat dikendalikan dengan memberlakukan supervisi yang ketat agar dapat mendeteksi terjadinya perilaku kontraproduktif. Alternatifnya dilakukan dengan menyeleksi pekerja yang akan diterima karena perusahaan mementingkan calon pekerja yang jujur dan dapat diandalkan. Harris dan Sackett (1987) bahkan menerapkan metode analisis faktor dan IRT dalam penelitian mengenai salah satu tes kejujuran.

Awalnya kejujuran dites menggunakan detektor kebohongan yang dinamakan poligraf. Namun sejak tahun 1988, *The Polygraph Protection Act* hanya mengizinkan penggunaan poligraf pada situasi yang benar-benar khusus. Program pengujian kejujuran yang dikembangkan setelahnya juga menuai larangan karena menilai seseorang tidak jujur berdasarkan hasil tes dianggap mengandung sensitivitas nilai moral yang berbeda dibandingkan dengan menguji kemampuan yang dimiliki seseorang (Lasson, 1992; Guastello dan Rieke, 1991). Fakta-fakta tersebut menjadikan minat terhadap alat ukur integritas berkembang sangat pesat. Tuntutan perusahaan terhadap kejujuran pekerja telah melahirkan industri pengukuran integritas bernilai jutaan dolar (O'Bannon dkk., 1989).

Penelitian mengenai integritas dalam konteks pekerjaan semakin diminati. Diawali oleh Craig dan Smith (2000) yang meneliti integritas untuk membedakan perilaku jujur dan tidak jujur para pekerja. Selanjutnya Impelman (2006) meneliti kaitan integritas terhadap perilaku kerja kontraproduktif. Kemudian Schlenker, Miller dan Johnson (2009) mengembangkan skala integritas dengan meminta responden mendefinisikan apakah prinsip-prinsip moral yang dimilikinya benar atau salah. Penelitian sejenis terus berkembang sangat pesat hingga saat ini.

The Integrity Scale merupakan salah satu alat ukur integritas yang banyak digunakan (Schlenker dan Forsyth, 1977). Namun alat ukur integritas seperti ini tidak tepat diaplikasikan di Indonesia. Hal ini terlihat dari bunyi salah satu *item*

yaitu “*Integrity is more important than financial gain*”. Jika langsung ditanyakan apakah seseorang memiliki integritas, makna pertanyaannya kurang operasional mengingat konsep integritas secara umum belum dipahami dengan tepat. Lagipula integritas seharusnya menanyakan keputusan yang diambil dalam sebuah situasi.

Contoh alat ukur integritas lainnya adalah *Integrity Scale Morton* yang setiap *item*-nya dirancang agar subjek menentukan derajat pilihan respon yang paling sesuai dengan dirinya di antara dua pernyataan.

The South African Integrity Scale (SAIS) juga merupakan contoh skala integritas yang khusus dirancang bagi sampel multikultural di Afrika Selatan. SAIS terdiri dari 36 *item* dengan waktu tempuh 10-15 menit. SAIS biasa digunakan dalam proses seleksi kerja pada tingkat pendidikan minimal *grade 10* untuk level non-manajerial. SAIS memasukkan *honest, norm abiding, punitive towards rule breaking, responsible* dan *trustworthy* sebagai komponen integritas.

Sejauh ini, hampir seluruh penggunaan alat ukur integritas bertujuan mengukur kejujuran. Padahal integritas tidak hanya identik dengan kejujuran. Individu yang jujur belum tentu memiliki integritas tinggi. Selain tidak khusus dirancang untuk mengukur integritas, melainkan lebih sebagai tes kejujuran dan tes perilaku kontraproduktif, alat ukur tersebut juga tidak bersifat unidimensi (tidak dapat dipastikan apakah memang mengukur integritas) karena dikembangkan menggunakan pendekatan tes klasik.

Selain alat ukur integritas dalam bentuk *paper and pencil* seperti yang dipaparkan, Egberink dan Veldkamp (2007) juga telah berupaya merancang pengembangan *computerized adaptive testing* bagi alat ukur integritas.

Di Indonesia, penelitian mengenai integritas belum berkembang meski kebutuhannya sangat disadari. Salah satu contoh adalah penelitian yang dilakukan Permatasari (2011) tentang kaitan gaya berpikir, integritas dan usia pada perilaku kerja kontraproduktif terhadap profesi wartawan. Penelitiannya menggunakan alat ukur integritas yang dirancang berdasarkan *moral identity theory* yang dikemukakan oleh Blasi (2004). Namun teori integritas tersebut dinilai tidak sesuai jika digunakan untuk mengukur integritas pekerja sebagai variabel psikologis. Kelemahan lain dari alat ukur integritas yang dirancang Permatasari (2011) adalah tidak dapat dipastikan apakah alat ukur tersebut memang benar

mengukur, dan hanya mengukur, integritas karena dirancang menggunakan pendekatan klasik.

Isu lain yang dinilai relevan dengan alat ukur integritas adalah *candidate faking*. Individu yang menempuh tes integritas memiliki kemungkinan memberi respon yang tidak sesuai dengan kondisi diri yang sebenarnya. Ellingson, Sackett dan Hough (1999) menyatakan hal tersebut bisa mengakibatkan hasil tes individu menjadi lebih tinggi atau malah lebih rendah. Pernyataan tersebut dibuktikan dengan hasil berbagai pengukuran mengenai *faking and socially desirable responses* yang dilakukan terhadap beberapa alat ukur integritas (Morgeson et al., 2007; Ones dan Viswesveran, 1998). Namun hal-hal tersebut dinyatakan tidak mempengaruhi validitas karena partisipan yang menempuh tes integritas diminta memberi respon diantara beberapa pilihan keputusan yang dianggap paling sesuai untuk dirinya. Sehingga disimpulkan bahwa *faking* dan *social desirability* tidak berdampak terhadap pengukuran integritas (Hough et al., 1990; Morgeson et al., 2007; Ones & Viswesvaran, 1998; Ones, Viswesvaran & Reiss, 1996).

Alat ukur integritas dalam penelitian ini dikembangkan menggunakan teori moderen yang dikenal sebagai *Item Response Theory*.

2.2. Pendekatan *Item Respon Theory* (IRT)

2.2.1. Kelemahan Pendekatan Klasik

Selama ini alat ukur dikembangkan menggunakan pendekatan klasik. Namun pendekatan ini memiliki kelemahan karena menghasilkan alat ukur yang terikat kepada sampel (*sample bound*), yang diistilahkan sebagai *group dependent* (Hambleton, Swaminathan dan Rogers, 1991). Alat ukur yang tergolong *sample bound* seolah-olah memiliki tingkat kesulitan yang tinggi ketika diberikan pada kelompok subjek yang memiliki kemampuan rendah dan seolah-olah memiliki tingkat kesulitan yang rendah ketika diberikan pada kelompok subjek yang memiliki kemampuan tinggi. Selain itu, dalam teori klasik sulit untuk menyeleksi soal-soal dengan tingkat kesulitan yang sesuai dengan kemampuan individu yang akan diukur. Ketika tes diberikan kepada kelompok individu dengan kemampuan tinggi, tingkat kesulitan *item* terlihat mudah karena sebagian besar menjawab benar. Tetapi ketika tes diberikan kepada kelompok individu dengan kemampuan

rendah, *item* terlihat sulit karena sebagian besar tidak bisa mengerjakan. Tingkat kesulitan *item* dalam teori tes klasik bisa berubah tergantung tingkat kemampuan individu yang menempuh tes.

Selain itu, kemampuan individu yang terukur dipengaruhi oleh kemampuan *item*. Individu terlihat memiliki kemampuan tinggi jika hanya mampu mengerjakan tes yang mudah dan terlihat memiliki kemampuan rendah jika mampu mengerjakan tes yang sulit. Inilah kelemahan lain pendekatan klasik yang disebut Embretson dan Reise (2000) sebagai *test-dependent*, yaitu kemampuan individu dipengaruhi oleh karakteristik *item* dalam sebuah tes.

Berdasarkan penjelasan di atas diketahui bahwa pada pendekatan klasik, karakteristik *item* dipengaruhi kemampuan individu dan kemampuan individu dipengaruhi karakteristik *item*. Sehingga sulit membandingkan kemampuan antara individu yang mengerjakan tes berbeda, serta membandingkan karakteristik *item* yang dikerjakan oleh kelompok individu yang berbeda.

2.2.2. Keunggulan Pendekatan IRT

Pendekatan IRT berupaya mengatasi kelemahan-kelemahan pendekatan klasik seperti *item dependent*, *sample dependent*, *test oriented* (sebaiknya *item oriented*) dan pemberlakuan *measurement error* yang sama untuk semua penempuh tes (Hambleton, Swaminathan dan Rogers, 1991).

Keunggulan IRT dinamakan sifat parameter *item* dan parameter kemampuan yang invarian (*invariance property*), yaitu karakteristik *item* (atau tingkat kesulitan soal) yang tidak bergantung pada kelompok peserta tes yang berasal dari populasi yang sama. Demikian pula sebaliknya, estimasi kemampuan peserta (*ability*) tidak tergantung pada karakteristik tes yang diberikan. Sehingga dapat dilakukan perbandingan antar individu penempuh tes serta perbandingan antar *item* tes. Keunggulan lain IRT adalah bermanfaat untuk pengujian dan pengembangan tes, pembuatan bank soal serta *Computerized Adaptive Testing* (CAT). Embretson dan Reise (2000) menjelaskan perbedaan mendasar pendekatan klasik dan modern seperti terlihat dalam tabel berikut:

Tabel 2.1. Perbandingan Pendekatan Klasik dan Pendekatan Modern (IRT)

No	Pendekatan Klasik	No	Pendekatan Modern (IRT)
1	<i>Standard error of measurement</i> berlaku untuk semua skor. (<i>The standard error of measurement applies to all scores in a particular population</i>)	1	Setiap skor (<i>tetha</i>) memiliki <i>standard error of measurement</i> berbeda. (<i>The standard error of measurement differs across scores but generalizes across populations</i>)
2	Semakin banyak jumlah <i>item</i> , semakin reliabel suatu tes. (<i>Longer tests are more reliable than shorter tests</i>)	2	Sedikit <i>item</i> berkualitas baik bisa lebih reliabel. (<i>Shorter tests can be more reliable than longer tests</i>)
3	Perbandingan skor tes beberapa <i>form</i> optimal apabila <i>form-form</i> tersebut paralel. (<i>Comparing test scores across multiple forms is optimal when test forms are parallel</i>)	3	Perbandingan skor tes beberapa <i>form</i> akan optimal bila derajat kesulitan (<i>threshold</i>) bervariasi. (<i>Comparing test scores across multiple forms is optimal when test difficulty levels vary between persons</i>)
4	<i>Unbiased estimates</i> dari karakteristik <i>item</i> bergantung dari sampel yang representatif. (<i>Unbiased assessment of item properties depends on having representative sampel</i>)	4	<i>Unbiased estimates</i> dari karakteristik <i>item</i> dapat diperoleh dari sampel yang tidak representatif. (<i>Unbiased assessment of item properties may be obtained from unrepresentative sampel</i>)
5	Skor tes diinterpretasi berdasarkan posisi dalam kelompok normatif. (<i>Tes scores obtain meaning by comparing their position in a norm group</i>)	5	Skor tes diinterpretasi berdasarkan isi <i>item</i> . (<i>Tes scores obtain meaning by comparing their distance from items</i>)
6	Skala interval diperoleh dari distribusi skor normal. (<i>Interval scale properties are achieved by obtaining normal score distributions</i>)	6	Skala interval diperoleh dengan memilih model pengukuran yang tepat. (<i>Interval scale properties are achieved by applying justifiable measurement models</i>)
7	Jenis <i>item</i> campuran memberi pengaruh tidak seimbang pada skor. (<i>Mixed item formats leads to unbalanced impact on test total scores</i>)	7	Jenis <i>item</i> campuran dapat mengarahkan skor tes secara optimal. (<i>Mixed item formats can yield optimal test scores</i>)
8	Perubahan skor tidak dapat	8	Perubahan skor dapat dibandingkan

	dibandingkan secara berarti ketika tingkat skor awal berbeda. (<i>Change scores cannot be meaningfully compared when initial score levels differ</i>)		secara berarti meskipun tingkat skor awal berbeda. (<i>Change scores can be meaningfully compared when initial score levels differ</i>)
9	Analisis faktor terhadap <i>item</i> biner menghasilkan objek, bukan faktor. (<i>Factor Analysis on binary items produces artifacts rather than factors</i>)	9	Analisis faktor terhadap data mentah membuat informasi analisis faktor secara maksimal. (<i>Factor Analysis on raw item data yields a full information factor analysis</i>)
10	Karakteristik <i>item</i> tidaklah penting untuk properti psikometri. (<i>Item stimulus features are unimportant compared to psychometric properties</i>)	10	Karakteristik <i>item</i> dapat berhubungan langsung dengan properti psikometri. (<i>Item stimulus features can be directly related to psychometric properties</i>)

Sumber: Embretson, S.E. & Reise, S.P. (2000). *Item response theory for psychologists*. New Jersey: Lawrence Erlbaum Associates, Publishers, pp. 15.

Keunggulan lain IRT adalah probabilitas peserta tes untuk menjawab benar suatu *item* benar-benar tergantung pada kemampuan peserta tes. Karakteristik *item* (seperti tingkat kesulitan soal) dan kemampuan peserta (*ability*) berada pada satu skala (dimensi) sehingga memungkinkan untuk melakukan perbandingan antar individu penempuh tes maupun antar *item* tes. Estimasi kesalahan pengukuran (*error*) bervariasi antar skor, namun berlaku umum dan dapat digeneralisasi antar populasi. Keunggulan-keunggulan tersebut menjadikan pendekatan IRT sangat bermanfaat untuk mendesain sebuah tes, melakukan seleksi *item*, mengoptimalkan sebuah desain tes serta mendeteksi *differential item functioning* (DIF), yaitu bias respon yang disebabkan oleh perbedaan fungsi *item* pada kelompok tertentu (Embretson dan Reise, 2000).

Hambleton, Swaminathan dan Rogers (1991) menyebutkan bahwa IRT memiliki dua hipotesis, yaitu performa penempuh tes pada suatu *item* dapat diprediksi oleh satu set faktor yang disebut dengan kemampuan/*trait*, dan hubungan antara performa penempuh tes pada suatu *item* dan kemampuan/*trait* yang mendasarinya dapat dijelaskan dengan suatu fungsi yang disebut dengan *item characteristic function*, yaitu semakin meningkat kemampuan/*trait*,

probabilitas jawaban atau persetujuan terhadap suatu *item* akan semakin besar pula. Sehingga diasumsikan bahwa penempuh tes dengan skor lebih tinggi dalam kemampuan/*trait* tertentu, memiliki probabilitas lebih besar untuk menjawab suatu *item* dengan benar atau ke arah persetujuan.

2.2.3. Asumsi Dalam Pendekatan IRT

Pendekatan IRT menggunakan model matematis dimana peluang individu menjawab benar terkait kemampuan individu dan karakteristik *item*. Hambleton, Swaminathan dan Rogers (1991) menyatakan pendekatan IRT dapat diterapkan ketika data memenuhi asumsi *unidimensionality* dan *local independence*.

a. *Unidimensionality* (unidimensionalitas)

Unidimensionalitas berarti ada satu faktor dominan yang mempengaruhi skor individu. Asumsi ini terkadang sulit dipenuhi karena adanya faktor-faktor kognitif, kepribadian dan faktor-faktor lain yang ditemui saat administrasi tes seperti kecemasan, motivasi dan sebagainya yang dapat mempengaruhi skor atau performa tes seseorang. Namun yang terpenting dalam asumsi tersebut adalah adanya satu komponen yang dianggap paling dominan dalam menentukan performa peserta tes. Faktor dominan tersebutlah yang terkait dengan apa yang diukur oleh tes.

b. *Local Independence*

Asumsi *local independence* bermakna tidak ada hubungan antara respon subjek terhadap *item-item* yang berbeda, yang artinya kemampuan yang diukur merupakan satu-satunya faktor yang mempengaruhi respon subjek. Jika kemampuan yang mempengaruhi performa tes dibuat konstan, maka respon penempuh tes pada serangkaian *item* manapun tidak terkait secara statistik.

2.2.4. *Rating Scale Model (RSM)* Dalam Pendekatan IRT

Hambleton, Swaminathan dan Rogers (1991) menyatakan bahwa pendekatan IRT menggunakan model yang dapat cocok ataupun tidak pada tes yang dianalisis (*falsifiable model*). Ada kemungkinan model IRT yang digunakan tidak dapat menjelaskan data, sehingga perlu dilakukan analisis kecocokan model

(*model fit*) terhadap data. Jika terdapat ketidakcocokan antara data dengan model, artinya model IRT yang digunakan tidak dapat diterapkan pada data tersebut.

Di dalam IRT, model pengukuran dibedakan berdasarkan kategori respon dan banyaknya parameter *item* yang dilibatkan dalam model. Berdasarkan kategori respon, model pengukuran IRT dibedakan menjadi model IRT dikotomi dan model IRT politomi. Model IRT dikotomi digunakan pada tes yang *item*-nya memiliki dua kategori respon. Sedangkan model IRT politomi digunakan pada tes yang *item*-nya memiliki lebih dari dua kategori respon. Penelitian ini menggunakan model IRT politomi karena *item-item* dalam alat ukur dirancang berbentuk skala *likert* yang memiliki empat kategori respon. Model ini biasa digunakan pada *typical performance test*. Model pengukuran IRT politomi yang digunakan dalam penelitian ini adalah *Rating Scale Model* (RSM). RSM yang dikembangkan oleh Andrich tahun (1999) merupakan kelompok *Model Rasch*. *Item* dengan format *rating scale* memiliki jarak tingkat kesulitan atau tingkat kecenderungan persetujuan *item* yang hampir sama pada setiap *item*. RSM lebih tepat digunakan karena alat ukur integritas dalam penelitian ini berbentuk skala likert dengan opsi 1 = *Sangat Setuju*, 2 = *Setuju*, 3 = *Tidak Setuju*, dan 4 = *Sangat Tidak Setuju*. Pada skala likert, perbedaan tingkat kesulitan antara setiap opsi dalam sebuah *item* diharapkan tidak jauh berbeda (Embretson dan Riese, 2000).

2.2.5. Analisis *Differential Item Functioning* (DIF)

Analisis *item* dalam IRT dapat dilakukan menggunakan DIF untuk mendeteksi adanya bias respon yang disebabkan perbedaan karakteristik antara suatu kelompok dengan kelompok yang lain. *Item* dikatakan terdeteksi DIF apabila peserta tes dengan karakteristik sama, namun dari kelompok berbeda, memiliki peluang yang berbeda dalam menjawab atau merespon *item* tersebut.

DIF dibedakan menjadi *uniform DIF* dan *non uniform DIF*. *Uniform DIF* terjadi ketika kedua kelompok peserta tes memiliki peluang yang sama (*uniform*) untuk menjawab atau merespon *item* pada setiap tingkat kemampuan atau *trait*. Contoh *uniform DIF* adalah pada *item* tertentu, laki-laki dan perempuan yang memiliki kemampuan sebesar x cenderung memberi persetujuan berbeda, dimana laki-laki memberikan kecenderungan persetujuan yang lebih tinggi dibanding

perempuan. Begitu pula laki-laki dan perempuan yang memiliki kemampuan sebesar y cenderung memberikan persetujuan yang berbeda, dimana laki-laki tetap memberi kecenderungan persetujuan yang lebih tinggi dibanding perempuan.

Sementara *non uniform DIF* terjadi ketika kedua kelompok peserta tes tidak memiliki peluang yang sama (*non uniform*) untuk menjawab atau merespon *item* pada setiap tingkat kemampuan atau trait. Contoh *non uniform DIF* adalah pada *item* tertentu, pada kelompok laki-laki dan perempuan yang memiliki kemampuan sebesar x , laki-laki cenderung memberikan persetujuan lebih tinggi dibanding perempuan. Namun pada kelompok laki-laki dan perempuan yang memiliki kemampuan sebesar y , laki-laki cenderung memberikan persetujuan lebih rendah dibanding perempuan (Embretson dan Riese, 2000).



BAB 3

METODE PENELITIAN

Dalam bab ini dijelaskan metode penelitian yang meliputi spesifikasi alat ukur, partisipan penelitian, teknik sampling, prosedur penelitian, teknik analisis data dan pelaksanaan uji coba.

3.1. Spesifikasi Alat Ukur

Alat ukur yang dikonstruksi dalam penelitian ini adalah alat ukur integritas. Berdasarkan tipe tingkah laku yang diukur, alat ukur tergolong *typical performance test* karena mengukur *trait*/kepribadian yang tidak mengandung jawaban benar dan salah (Friedenberg, 2011). Berdasarkan respon jawaban yang diminta, alat ukur termasuk jenis skala *likert* karena selain memilih jawaban sesuai-tidak sesuai, partisipan juga diminta memberi kepastian derajat kesesuaian dari pilihan jawaban (DeVellis, 2003). Derajat kesesuaian antar pilihan jawaban disusun berdasarkan interval yang diasumsikan sama sehingga partisipan dapat menentukan pilihan dengan menyesuaikan karakteristik yang ada pada dirinya.

Skala dalam alat ukur ini terdiri dari *item-item* yang menyediakan 4 pilihan respon mulai dari “Sangat Sesuai” sampai “Sangat Tidak Sesuai”. Skala tersebut dipilih agar partisipan memberikan respon yang lebih bervariasi dalam rentang tertentu dan dapat mengevaluasi pernyataan yang ada sesuai dengan kondisi dirinya. Skala 1-4 dipilih untuk menghindari kecenderungan partisipan menjawab netral dan memilih respon yang cenderung ke arah tertentu. Berikut cara skoring alat ukur integritas yang dibuat.

Tabel 3.1. Cara Skoring Alat Ukur Integritas

Skala	Skala <i>Item Favorable</i>	Skala <i>Item Unfavorable</i>
SS (Sangat Sesuai)	4	1
S (Sesuai)	3	2
TS (Tidak Sesuai)	2	3
STS (Sangat Tidak Sesuai)	1	4

Alat ukur dirancang berdasarkan teori integritas yang dikemukakan oleh Rogers (1961). Awalnya alat ukur memiliki 40 *item* yang ditetapkan berdasarkan hasil wawancara dengan beberapa *expert*. Sebelumnya telah dilakukan studi pendahuluan dan analisis statistik untuk menjadi acuan memperbaiki *item-item* yang kurang baik. Kemudian peneliti menguji coba alat ukur integritas kepada 100 partisipan yang hasilnya diolah kembali untuk mendapatkan hasil uji psikometrinya. Alat ukur integritas yang telah melalui proses uji coba selanjutnya melewati proses elisitasi untuk merevisi *item-item* yang kurang sesuai sebelum digunakan untuk pengambilan data. Berikut kisi-kisi 40 *item* alat ukur integritas.

Tabel 3.2. Kisi-Kisi Alat Ukur Integritas 40 Item

Faktor Integritas	Indikator	Nomor Item	
		Favorable	Unfavorable
1. Jujur	a. konsistensi antara ucapan dan perilaku sehingga sangat bisa dipercaya	1, 9, 17	21, 27
	b. bertindak tanpa tujuan tersembunyi	5, 13	24, 29
2. teguh	a. Taat dalam prinsip	18, 25	2, 10
	b. tidak dapat disuap atau diajak melakukan perbuatan curang	22, 28	6, 14
	c. menjalankan kewajiban dengan tekun	30, 34, 38	32, 36
	d. tahan dalam menghadapi pertentangan dan ketidakpercayaan	31, 35, 39	33, 37, 40
3. Memiliki <i>self-control</i> yang kuat	Mampu mengontrol atau memantau tindakan dan pikirannya berdasar prinsip dan beliefs dirinya	3, 11, 19, 26	7, 15, 23
4. Memiliki <i>self-esteem</i> yang besar	mampu berperilaku sesuai nilai moral yang diyakini	4, 12, 20	8, 16
Total Item		22	18

3.2. Partisipan Penelitian

Untuk menguji validitas alat ukur integritas dalam konteks pekerjaan di Indonesia, peneliti memilih pekerja sebagai partisipan penelitian. Alat ukur diberikan kepada individu yang menduduki level manager, supervisor dan staf dengan masa kerja minimal satu tahun di perusahaan tempat bekerja saat ini. Persyaratan masa kerja minimal satu tahun digunakan agar pekerja memiliki pengetahuan dan pengalaman bekerja di perusahaan. Partisipan juga disyaratkan memiliki tingkat pendidikan minimal strata satu (S1) dalam rentang usia 24–56 tahun. Patokan itu merujuk pada usia pensiun yang ditetapkan Undang-Undang Ketenagakerjaan. Usia tersebut merupakan tahap formal-operasional yang ditandai adanya kemampuan berpikir abstrak sehingga partisipan diharapkan mampu melakukan evaluasi dan penilaian terhadap hidupnya (Papalia, 2004). Individu yang mencapai tahap ini dianggap memiliki kemampuan mengintegrasikan *self* agar terjadi koherensi antara nilai-nilai moral yang dipahami, *beliefs*, tindakan, komitmen serta perkataannya. Tahapan usia tersebut juga menjadikan individu lebih mampu menjalani proses-proses untuk tercapainya integritas seperti yang dikemukakan oleh Petrick dan Quinn (2000). Partisipan penelitian mencakup 1210 pekerja yang berdomisili di wilayah Pulau Sumatera dan Jawa yang merupakan pulau dengan tingkat kepadatan penduduk serta jumlah perusahaan terbesar di Indonesia. Partisipan dari Pulau Sumatera berasal dari Sumatera Utara (Medan, Binjai, Rantau Prapat, Pematang Siantar dan Tebing Tinggi), Sumatera Barat (Padang), Sumatera Selatan (Palembang), Pekanbaru (Riau), Lampung, Jambi dan Bengkulu. Partisipan dari Pulau Jawa berasal dari Jakarta, Jawa Barat (Bandung), Jawa Tengah (Semarang), Jawa Timur (Surabaya) dan Yogyakarta. Berdasarkan paparan tersebut, kriteria partisipan penelitian ini adalah :

1. Pekerja dengan masa kerja minimal satu tahun di perusahaan tempat bekerja saat ini. Pekerja yang dimaksud menduduki posisi manager, supervisor dan staf.
2. Tingkat pendidikan minimal strata satu (S1).
3. Rentang usia 24 – 56 tahun.
4. Berjenis kelamin laki-laki atau perempuan.
5. Bekerja dan berdomisili di Pulau Sumatera atau Pulau Jawa.

3.3. Teknik Sampling

Populasi partisipan sangat besar sehingga sulit diidentifikasi satu per satu, maka peneliti menggunakan teknik *nonprobability accidental/incidental sampling* yang dipilih dengan mempertimbangkan kemudahan mengakses sampel (Kumar, 1996; Guilford dan Fruchter, 1978). Selain itu, teknik *accidental sampling (convenience sampling)* memungkinkan peneliti mendapatkan jumlah sampel relatif besar dalam waktu singkat (Mitchell dan Janina, 2012).

Peneliti memberi kuesioner kepada pekerja yang menyatakan kesediaan berpartisipasi yang sebelumnya telah dipilih berdasarkan kriteria yang ditetapkan. Partisipan dicari berdasarkan ketersediaan dan kemudahan yang ditemui peneliti. Menurut Kumar (1996) teknik ini paling mudah untuk menyeleksi responden dan menjamin diperolehnya karakteristik responden yang dibutuhkan.

3.4. Prosedur Penelitian

Penelitian ini dilakukan mengikuti beberapa prosedur yang harus dipenuhi peneliti. Berikut adalah penjelasan setiap tahapan penelitian yang dilakukan.

3.4.1. Tahap Persiapan

Peneliti melakukan kajian literatur mengenai integritas untuk melengkapi teori serta mempersiapkan alat ukur yang digunakan dalam penelitian. Peneliti juga melakukan elisitasi untuk mendapatkan gambaran spesifik mengenai indikator-indikator integritas secara langsung dari partisipan. Pertanyaan yang diajukan dalam proses elisitasi dirancang berbentuk wawancara semi terstruktur dan *focus group discussion (FGD)* kepada 10 pekerja yang terdiri dari enam orang manager, dua orang supervisor dan dua orang staf. Dua dari tiga manager yang bekerja sebagai PNS berasal dari Pulau Jawa, sedangkan lainnya dari Pulau Sumatera. Dua dari tiga manager yang bekerja di perusahaan swasta berasal dari Pulau Sumatera, sedangkan lainnya berasal dari Pulau Jawa. Seluruh manager telah memiliki pengalaman kerja minimal lima tahun. Kedua supervisor berasal dari Pulau Jawa tetapi yang satu bekerja sebagai PNS dan lainnya bekerja di perusahaan swasta. Kedua staf juga berasal dari Pulau Jawa dan masing-masing bekerja di jenis perusahaan yang berbeda (swasta dan PNS), tetapi yang satu

memiliki status sebagai pegawai kontrak dan yang lainnya sudah memiliki status sebagai pegawai tetap. Seluruh supervisor dan staf telah memiliki masa kerja lebih dari satu tahun pada perusahaan tempat bekerja saat ini.

Tabel 3.3. Gambaran Partisipan Elisitasi

Jabatan	Status Pekerjaan dan Asal Domisili	
	PNS	Swasta
Manager	2 orang (Pulau Jawa) 1 orang (Pulau Sumatera)	1 orang (Pulau Jawa) 2 orang (Pulau Sumatera)
Supervisor	1 orang (Pulau Jawa)	1 orang (Pulau Jawa)
Staff	1 orang (Pulau Jawa)	1 orang (Pulau Jawa)

Pertanyaan-pertanyaan yang diajukan dalam proses elisitasi mencakup pemahaman, pemaknaan dan kebutuhan praktis integritas di perusahaan. Hasil elisitasi memberikan informasi tambahan mengenai situasi-situasi yang dinilai dapat menggali integritas dalam diri individu terkait konteks pekerja di Indonesia.

3.4.2. Tahap Uji Coba

Pada tahap ini dilakukan pengujian validitas, analisis *item* dan uji asumsi unidimensionalitas menggunakan *confirmatory factor analysis* (CFA) yang harus dipenuhi dalam penggunaan *Item Response Theory*. Jika hasil CFA menunjukkan bahwa alat ukur memenuhi asumsi unidimensionalitas, maka dapat dilakukan perhitungan menggunakan IRT untuk menentukan apakah model alat ukur integritas sudah *fit* dan terdiri dari *item-item* yang *fit*.

Uji coba dilakukan pada tanggal 14–21 April 2012 melibatkan 100 pekerja yang representatif terhadap partisipan penelitian. Uji coba untuk mendapatkan validitas alat ukur integritas diawali dengan meminta *expert judgement* dan dilanjutkan dengan pengujian menggunakan metode *confirmatory factor analysis*.

a. Uji Validitas Isi dan *Expert Judgement*

Awalnya peneliti melakukan pengujian validitas isi (*content validity*). Alat ukur integritas didiskusikan dengan beberapa pakar psikologi industri-organisasi

yang memahami konstruk integritas untuk mendapatkan *expert judgement*. Narasumber terdiri dari praktisi di beberapa perusahaan serta akademisi dari beberapa universitas di Pulau Sumatera dan Pulau Jawa. Peneliti memperbaiki alat ukur integritas berdasarkan masukan para pakar tersebut.

Selanjutnya peneliti melakukan uji keterbacaan terhadap calon partisipan untuk mendapatkan masukan mengenai penulisan *item*, tata bahasa, pilihan respon serta tampilan kuesioner. Uji keterbacaan termasuk analisis *item* secara kualitatif yang hasilnya menjadi dasar memperbaiki teknis penulisan. Setelah *item-item* direvisi, alat ukur menjalani proses uji coba lebih lanjut untuk menguji validitas serta analisis *item* secara kuantitatif. Hasil ujicoba akan memberi informasi apakah alat ukur valid, reliabel dan terdiri dari *item-item* yang baik.

b. Pengujian Menggunakan *Confirmatory Factor Analysis*

Pengujian validitas konstruk dilakukan menggunakan metode analisis faktor yaitu dengan *confirmatory factor analysis* (CFA) yang dilakukan untuk menguji dimensionalitas sebagai acuan menguji asumsi unidimensionalitas alat ukur. CFA juga memberi informasi apakah model *fit* untuk mengukur integritas menggunakan alat ukur ini. Pada tahap ini, analisis dilakukan menggunakan program LISREL versi 8.7 (Joreskog & Sorbom, 1996).

Untuk menilai apakah model pengukuran benar-benar *fit* dengan data, perlu diperhatikan nilai indeks *fit*. Indeks *fit* yang dihasilkan dari analisis menggunakan metode CFA ada berbagai macam. Suatu indeks yang menunjukkan model tersebut *fit* tidak memberikan jaminan bahwa model benar-benar *fit*, begitu pula sebaliknya. Maka peneliti tidak hanya bergantung pada salah satu indeks *fit* untuk menguji model tersebut. Penelitian ini menggunakan kriteria indeks *fit* berupa *Chi-Square* dan *Root Mean Square Error of Approximation* (RMSEA).

1. *Chi-square*

Nilai *chi-square* menunjukkan penyimpangan antara *sample covariance matrix* dengan *model covariance matrix* (Joreskog dan Sorbom, 1993). *Chi-square* merupakan ukuran mengenai buruknya *fit* suatu model. Probabilitas *chi-square* dalam pengujian ini diharapkan tidak signifikan ($p > 0.05$) agar model dikatakan *fit* dengan data.

2. *Root Mean Square Error of Approximation* (RMSEA)

RMSEA diperkenalkan oleh Steiger dan Lind tahun 1980. RMSEA mengukur penyimpangan nilai parameter pada suatu model dengan matriks kovarians populasinya. Nilai $RMSEA < 0.05$ mengindikasikan model *fit* dan nilai RMSEA yang berkisar antara 0.08 menyatakan bahwa model memiliki perkiraan kesalahan yang *reasonable*. Sementara nilai RMSEA yang berkisar antara 0.08 sampai dengan 0.10 menunjukkan model memiliki *fit* yang cukup. Sedangkan nilai $RMSEA > 0.1$ menunjukkan model *fit* yang sangat buruk. Kesimpulannya, model dikatakan *fit* dengan data ketika nilai $RMSEA < 0.10$.

Dalam penelitian ini, *first order* CFA diuji terhadap 100 sampel yang dipilih untuk memenuhi ketentuan teknik estimasi melakukan CFA menggunakan LISREL. Berikut hasil pengujian *first order* CFA.

1. Pengujian *First Order* CFA Integritas 40 item pada 100 partisipan

Pengujian *first order* CFA dilakukan dengan membuat model pengukuran untuk menggambarkan sebaik apa indikator-indikator dalam alat ukur dapat digunakan sebagai instrumen pengukuran variabel laten. Dalam pengujian ini, peneliti membuat model pengukuran dengan satu variabel laten (integritas) dan 40 itemnya dijadikan sebagai indikator yang diukur secara langsung. Pengujian ini membuktikan bahwa seluruh *item* mengukur satu variabel laten yaitu integritas.

Hasil pengujian terhadap model pengukuran *first order* CFA menghasilkan nilai $p\text{-value} = 0.016$ ($p < 0.05$) dan $RMSEA = 0.034$ ($RMSEA < 0.1$). Hanya kriteria RMSEA yang berhasil terpenuhi sehingga peneliti memutuskan model ini tidak *fit* dengan data. *Path* diagram dari pengujian dapat dilihat pada lampiran 4.

Tabel 3.4. Indeks Kecocokan *First Order* CFA 40 Item

	Indeks Kecocokan	Keterangan
<i>Chi-Square</i>	825.07; $df=740$; $p=0.01582$	Model tidak <i>fit</i>
RMSEA	0.034	Model <i>fit</i>

Peneliti lalu mengeliminasi beberapa *item* yang memberi sumbangan kecil pada alat ukur. Pada faktor JUJUR, yang dieliminasi adalah *item* 17, 27 dan 29.

Pada faktor TEGUH, yang dieliminasi adalah *item* 10, 14, 32, 35, 37 dan 40. Pada faktor *SELF-CONTROL*, yang dieliminasi adalah *item* 11, 15 dan 26. Pada faktor *SELF-ESTEEM*, yang dieliminasi adalah *item* 8 dan 16. Total diperoleh 26 *item* integritas.

Uji coba alat ukur integritas 26 *item* kembali dilakukan dengan melibatkan 100 pekerja lain yang masih representatif terhadap partisipan penelitian. Total 26 *item* yang lulus hasil pengujian *first order* CFA dipaparkan dalam tabel berikut.

Tabel 3.5. Kisi-Kisi Alat Ukur Integritas 26 Item

Faktor Integritas	Subdefinisi	Nomor Item	
		Favorabel	Unfavorabel
1 Jujur	Konsistensi antara ucapan dan perilaku sehingga sangat bisa dipercaya	1, 9	21
	Bertindak tanpa tujuan tersembunyi,	5, 13	24
2 Teguh	Taat dalam prinsip	18, 25	2
	Tidak dapat disuap atau diajak melakukan perbuatan curang,	22, 28	6
	Menjalankan kewajiban dengan tekun	30, 34, 38	36
	Tahan dalam menghadapi pertentangan dan ketidakpercayaan	31, 39	33
3 Memiliki <i>self-control</i> yang kuat	Mampu mengontrol atau memantau tindakan dan pikirannya berdasar prinsip dan beliefs dirinya	3, 19	7, 23
4 Memiliki <i>self-esteem</i> yang tinggi	Mampu berperilaku sesuai nilai moral yang diyakini	4, 12, 20	-
Total Item		18	8

26 *item* yang telah lulus pengujian *first order* CFA tersebut kemudian kembali menjalani proses pengujian *first order* CFA.

2. Pengujian kembali *first order* CFA

Hasil pengujian terhadap model pengukuran *first order* CFA dengan 26 *item* menghasilkan nilai $p\text{-value} = 0.48$ ($p > 0.05$) dan $RMSEA = 0.00$ ($RMSEA < 0.1$). Berdasarkan data, $p\text{-value}$ dan $RMSEA$ berhasil terpenuhi sehingga disimpulkan model ini benar-benar *fit* dengan data.

Tabel 3.6. Indeks Kecocokan *First Order* CFA 26 Item

	Indeks Kecocokan	Keterangan
<i>Chi-Square</i>	299.39; df=299; p=0.48	Model <i>fit</i>
RMSEA	0.00	Model <i>closed fit</i>

Dengan kata lain, 26 *item* tersebut merupakan indikator yang valid bagi pengukuran konstruk integritas. Hasil juga menunjukkan bahwa 26 *item* tersebut mengukur satu variabel laten, yaitu integritas. Maka alat ukur integritas ini dapat dikatakan telah memenuhi asumsi unidimensionalitas sehingga penerapan model dengan pendekatan IRT bisa dilakukan. Hasil *path* diagram dari pengujian tersebut dapat dilihat pada bagian lampiran 6.

Selain mengetahui model fit dari suatu pengukuran, CFA juga memberi informasi mengenai indeks fit dari setiap *item*. Berikut adalah tabel *fit item first order* CFA alat ukur integritas.

Tabel 3.7. Hasil *First Order* CFA Alat Ukur Integritas 26 Item

Item	Loading Factor	t-Value	R²	Keterangan
Faktor: JUJUR				
<i>Item 1</i>	0.091	8.08	0.51	<i>Item Fit</i>
<i>Item 5</i>	0.089	7.52	0.46	<i>Item Fit</i>
<i>Item 9</i>	0.080	7.19	0.43	<i>Item Fit</i>
<i>Item 13</i>	0.10	8.06	0.51	<i>Item Fit</i>
<i>Item 21</i>	0.094	6.90	0.40	<i>Item Fit</i>
<i>Item 24</i>	0.080	7.17	0.42	<i>Item Fit</i>
Faktor: TEGUH				
<i>Item 2</i>	0.084	7.78	0.48	<i>Item Fit</i>
<i>Item 6</i>	0.093	6.95	0.40	<i>Item Fit</i>
<i>Item 18</i>	0.090	7.03	0.41	<i>Item Fit</i>
<i>Item 22</i>	0.093	7.24	0.43	<i>Item Fit</i>
<i>Item 25</i>	0.098	7.68	0.47	<i>Item Fit</i>
<i>Item 28</i>	0.092	7.45	0.45	<i>Item Fit</i>
<i>Item 30</i>	0.10	7.08	0.42	<i>Item Fit</i>
<i>Item 31</i>	0.093	6.88	0.40	<i>Item Fit</i>
<i>Item 33</i>	0.10	6.95	0.40	<i>Item Fit</i>
<i>Item 34</i>	0.096	7.06	0.41	<i>Item Fit</i>
<i>Item 36</i>	0.10	7.61	0.47	<i>Item Fit</i>
<i>Item 38</i>	0.10	7.26	0.43	<i>Item Fit</i>
<i>Item 39</i>	0.093	7.60	0.46	<i>Item Fit</i>
Faktor: SELF CONTROL				
<i>Item 3</i>	0.076	7.99	0.50	<i>Item Fit</i>
<i>Item 7</i>	0.10	7.31	0.44	<i>Item Fit</i>
<i>Item 19</i>	0.095	7.27	0.43	<i>Item Fit</i>
<i>Item 23</i>	0.094	7.29	0.44	<i>Item Fit</i>
Faktor: SELF ESTEEM				
<i>Item 4</i>	0.100	7.53	0.46	<i>Item Fit</i>
<i>Item 12</i>	0.095	7.32	0.44	<i>Item Fit</i>
<i>Item 20</i>	0.086	6.93	0.40	<i>Item Fit</i>

Tabel menunjukkan bahwa seluruh *item fit* dalam mengukur integritas. Hal ini terlihat dari indeks *t-value* yang lebih besar dari 1.96. *Item* yang memberikan kontribusi terbesar untuk alat ukur integritas adalah *item 4, 7, 13, 30, 33, 36 dan 38* dengan muatan faktor sebesar 0.10 sementara *item* yang memberikan kontribusi terkecil adalah *item 3* dengan muatan faktor sebesar 0.076.

3. Pengujian *second order* CFA

Hasil pengujian yang dilakukan terhadap model pengukuran *second order* CFA pada 26 *item* menghasilkan nilai *p-value* = 0.47 ($p > 0.05$), dan RMSEA = 0.00 (RMSEA < 0.1). Berdasarkan data, *p-value* dan RMSEA berhasil terpenuhi

sehingga dapat disimpulkan model ini benar-benar *fit* dengan data. Nilai RMSEA sebesar 0.00 menunjukkan bahwa model ini *closed fit* atau sangat *fit*.

Tabel 3.8. Indeks Kecocokan *Second Order CFA 26 Item*

	Indeks Kecocokan	Keterangan
<i>Chi-Square</i>	297.36; df=296; p=0.47	Model <i>fit</i>
RMSEA	0.00	Model <i>closed fit</i>

Dengan kata lain 26 *item* tersebut merupakan indikator yang valid bagi pengukuran konstruk integritas. Hasil juga menunjukkan bahwa 26 *item* tersebut mengukur satu variabel laten, yaitu integritas. Maka dapat dikatakan bahwa alat ukur integritas ini telah memenuhi asumsi unidimensionalitas sehingga penerapan model dengan pendekatan IRT bisa dilakukan. Hasil *path* diagram dari pengujian tersebut dapat dilihat pada lampiran 8. Berikut adalah tabel *fit* seluruh *item* dari hasil pengujian *second order CFA* alat ukur integritas.

Tabel 3.9. Hasil *Second Order CFA* Alat Ukur Integritas 26 *Item*

<i>Item</i>	<i>Loading Factor</i>	<i>t-value</i>	<i>R²</i>	<i>Keterangan</i>
Faktor: JUJUR				
Item 1	-	-	0.50	Reference Variable
<i>Item 5</i>	0.14	6.53	0.44	<i>Item Fit</i>
<i>Item 9</i>	0.12	6.35	0.42	<i>Item Fit</i>
<i>Item 13</i>	0.16	6.95	0.50	<i>Item Fit</i>
<i>Item 21</i>	0.14	6.12	0.39	<i>Item Fit</i>
<i>Item 24</i>	0.12	6.30	0.41	<i>Item Fit</i>
Faktor: TEGUH				
Item 2	-	-	0.48	Reference Variable
<i>Item 6</i>	0.16	6.09	0.40	<i>Item Fit</i>
<i>Item 18</i>	0.14	7.27	0.41	<i>Item Fit</i>
<i>Item 22</i>	0.17	6.26	0.43	<i>Item Fit</i>
<i>Item 25</i>	0.18	6.54	0.47	<i>Item Fit</i>
<i>Item 28</i>	0.16	6.41	0.45	<i>Item Fit</i>
<i>Item 30</i>	0.18	6.16	0.41	<i>Item Fit</i>
<i>Item 31</i>	0.16	6.06	0.40	<i>Item Fit</i>
<i>Item 33</i>	0.18	6.07	0.40	<i>Item Fit</i>
<i>Item 34</i>	0.17	6.15	0.41	<i>Item Fit</i>
<i>Item 36</i>	0.18	6.51	0.46	<i>Item Fit</i>
<i>Item 38</i>	0.18	6.27	0.43	<i>Item Fit</i>
<i>Item 39</i>	0.17	6.50	0.46	<i>Item Fit</i>
Faktor: SELF CONTROL				
Item 3	-	-	0.50	Reference Variable
<i>Item 7</i>	0.19	6.41	0.44	<i>Item Fit</i>
<i>Item 19</i>	0.18	6.42	0.44	<i>Item Fit</i>
<i>Item 23</i>	0.18	6.39	0.44	<i>Item Fit</i>
Faktor: SELF ESTEEM				
Item 4	-	-	0.47	Reference Variable
<i>Item 12</i>	0.15	6.24	0.46	<i>Item Fit</i>
<i>Item 20</i>	0.13	6.09	0.44	<i>Item Fit</i>

Berdasarkan *t-value* dari pengujian *second order* CFA diketahui seluruh *item fit* mengukur integritas karena seluruh *t-value* lebih besar dari 1.96. Dari tabel diketahui juga bahwa *item 7* memberikan kontribusi terbesar untuk alat ukur dengan muatan faktor sebesar 0.19, sementara *item 9* dan 24 memberikan kontribusi terkecil dengan muatan faktor sebesar 0.12.

3.4.3. Tahap Pelaksanaan

Alat ukur yang digunakan dalam penelitian terdiri dari 26 *item* integritas yang lolos proses uji coba. Peneliti mengumpulkan data lengkap calon partisipan dan menghubunginya secara personal untuk meminta kesediaan berpartisipasi.

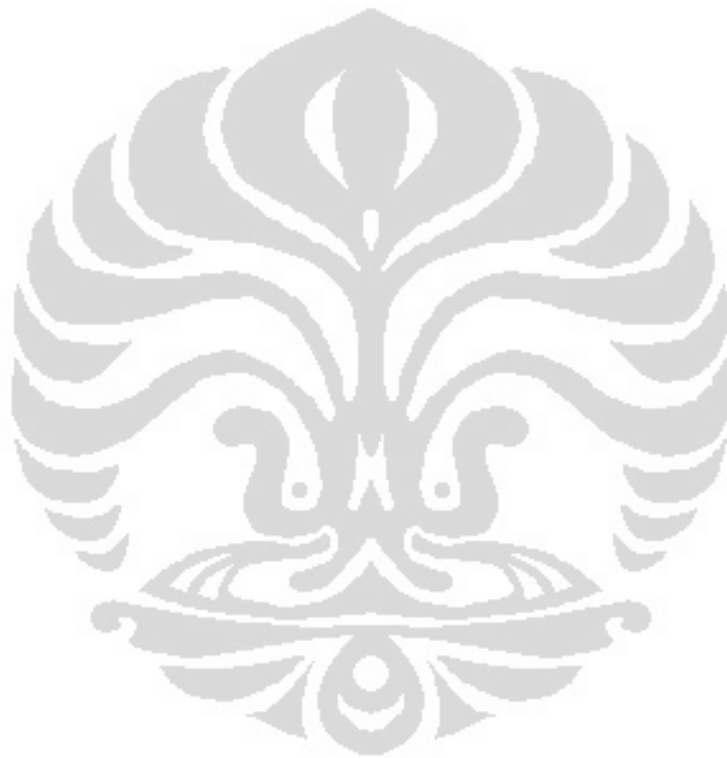
Pengambilan data dilakukan pada tanggal 28 April 2012 sampai dengan tanggal 25 Mei 2012. Kuesioner disebar kepada 1798 pekerja di Pulau Sumatera dan Pulau Jawa yang memenuhi kriteria yang telah ditetapkan sebelumnya. Dari 1798 kuesioner yang disebar, hasil yang dapat diolah hanya diperoleh dari 1000 kuesioner. Data dari 798 kuesioner sisanya tidak dapat diolah karena alasan pengisian yang tidak lengkap atau tidak dikembalikan.

3.5. Teknik Analisis Data dengan Pendekatan IRT

Analisis hasil pengukuran dilakukan menggunakan model pengukuran RSM (*Rating Scale Model*) untuk menguji model *polytomous* IRT. Untuk dapat menerapkan pendekatan IRT, kedua asumsi IRT yaitu unidimensionalitas dan *local independence* harus terpenuhi (Hambleton, Swaminathan dan Rogers, 1991). Alat ukur integritas memenuhi asumsi unidimensionalitas berdasarkan pengujian CFA karena terbukti mengukur satu faktor, yaitu integritas. Terkait dengan asumsi *local independence*, respon partisipan pada suatu *item* alat ukur integritas tidak terkait atau dipengaruhi oleh *item-item* lain, sehingga asumsi *local independence* juga terpenuhi. Oleh karena itu, penerapan model IRT dapat dilakukan.

Model pada *Item Response Theory* (IRT) dipilih berdasarkan bentuk matematik fungsi karakteristik *item* dan jumlah parameter yang dilibatkan dalam model. Model yang sesuai (*fit*) dengan alat ukur (skala) tertentu, belum tentu *fit* dengan skala yang lain. Peneliti menggunakan model *polytomous IRT* karena alat ukur integritas ini dirancang menggunakan skala tipe *likert* dengan empat pilihan respon jawaban. Model politomi yang diterapkan adalah *rating scale model*, yaitu suatu model IRT politomi yang menggunakan satu parameter *item*. Model ini sejalan dengan *Rasch Model* pada model IRT dikotomi. Model tersebut dipilih karena penelitian ini berupaya untuk mengembangkan alat ukur integritas yang bisa dipakai berulang-ulang, sehingga diperlukan pengukuran yang bebas sampel (*sampling invariant*). Untuk alasan tersebut, peneliti menggunakan pendekatan *model measurement* dan menggunakan satu parameter *item*.

Analisis dilakukan menggunakan *software* QUEST. *Software* ini memungkinkan peneliti mengetahui indeks nilai ambang (*delta*) dari setiap pilihan jawaban untuk setiap *item*, yang biasa disebut derajat kesulitan *item*, atau dalam hal ini merupakan tingkat kecenderungan persetujuan *item*. Perhitungan *item-fit statistics* berdasarkan nilai *infit* juga dilakukan untuk melihat apakah *item* baik atau buruk. Selanjutnya dilakukan pengujian *differential item functioning* (DIF) untuk melihat kemungkinan adanya *item-item* yang memiliki bias respon terhadap jenis kelamin. Analisis DIF juga menggunakan program QUEST.



BAB 4

HASIL DAN ANALISIS HASIL

Bagian ini menguraikan hasil dari penelitian yang diperoleh dari kuesioner yang telah diisi oleh partisipan penelitian.

4.1. Gambaran Umum Partisipan Penelitian

Gambaran demografis penyebaran partisipan penelitian dihitung berdasarkan jenis kelamin, status pekerjaan dan jabatan. Berikut gambaran demografis berdasarkan distribusi frekuensi.

**Tabel 4.1. Demografis Penyebaran Partisipan
Berdasarkan Jenis Kelamin, Status Pekerjaan dan Jabatan**

Jenis Kelamin	Status Pekerjaan	Jabatan	Frekuensi	Persentase (%)
Laki-Laki	PNS	Manager	119	11.9
		Supervisor/Staf	113	11.3
	Swasta	Manager	168	16.8
		Supervisor/Staf	100	10
Perempuan	PNS	Manager	148	14.8
		Supervisor/Staf	120	12
	Swasta	Manager	115	11.5
		Supervisor/Staf	117	11.7
Jumlah			1000	100

Berdasarkan tabel diketahui bahwa partisipan perempuan dalam penelitian ini berjumlah 500 orang (50%) Sementara partisipan laki-laki dalam penelitian ini berjumlah 500 orang (50%).

Partisipan yang memiliki status pekerjaan sebagai PNS berjumlah 500 orang (50%), sementara partisipan yang memiliki status pekerjaan sebagai swasta berjumlah 500 orang (50%).

Partisipan yang memiliki jabatan sebagai manager berjumlah 550 orang (55%), sementara partisipan yang memiliki jabatan sebagai supervisor/staf berjumlah 450 orang (45%).

Tabel 4.2. Demografis Penyebaran Partisipan Berdasarkan Usia

Usia	Frekuensi	Persentase (%)
> 50 tahun	99	9.9
40 – 50 tahun	351	35.1
30 – 39 tahun	415	41.5
< 30 tahun	135	13.5
Jumlah	1000	100

Berdasarkan tabel diketahui bahwa penelitian ini melibatkan 135 (13.5%) partisipan yang berusia di bawah 30 tahun, 415 (41.5%) partisipan yang berusia antara 30-39 tahun, 351 (35.1%) partisipan yang berusia antara 40-50 tahun, serta 99 (9.9%) partisipan yang berusia di atas 50 tahun.

4.2. Hasil Pengujian Alat Ukur Integritas Menggunakan Pendekatan Polytomous *Item Response Theory* (IRT)

Dalam pengujian ini diperoleh hasil pengujian *case fit*, pengujian terhadap model *fit*, *item fit* dan estimasi parameter setiap *item*. Berikut hasil pengujian model *fit*, *item fit* dan estimasi parameter setiap *item*. Pengujian *model fit* dalam IRT terdiri dari 2 bentuk yaitu pengujian *person fit* dan pengujian *item fit*.

4.2.1. Pengujian *Person fit*

Dari 1000 partisipan, seluruh kasus bisa digunakan karena tidak ada kasus yang memiliki nilai salah semua (*zero*) atau benar semua (*perfect*). Selain itu analisis terhadap nilai *infit mean square* pada estimasi kasus menunjukkan bahwa model dari seluruh kasus *fit* dengan data, yaitu sebesar 0.96 dan berada pada rentang yang dapat diterima. Peneliti juga melakukan pengujian model *fit*. Dalam pengujian *case fit* menggunakan QUEST, terdapat kriteria yang menunjukkan seberapa baik model tersebut dapat menjelaskan hasil tes yang diperoleh. Data dianggap sesuai dengan model apabila nilai *mean square* mendekati 1,0 (baik *infit* maupun *outfit*) dan nilai uji t mendekati 0 (baik *infit* maupun *outfit*). Adams dan Khoo (1993) menyarankan penggunaan *fit mean square* karena lebih berguna untuk melihat kesesuaian antara model dengan data, dibandingkan uji t yang sangat peka dengan jumlah sampel. *Infit mean square* dianggap lebih baik karena berkaitan dengan diskriminasi *item*, dibandingkan dengan *outfit mean square*,

karena nilai *outfit* sensitif terhadap data yang menyimpang (*outlier*). Selain itu, nilai *infit* lebih kuat dan stabil dibanding nilai *outfit*. Meskipun biasanya nilai *infit mean square* dan *outfit mean square* tidak jauh berbeda.

Kriteria pengujian *item fit* dalam pengujian alat ukur integritas ini adalah nilai *infit mean square*. Menurut Adams dan Khoo (1993), model dikatakan *fit* dengan data jika nilai *infit mean square* berada pada rentang 0,77 dan 1,30. Hasil pengujian *item fit* pada alat ukur integritas memperoleh nilai *infit mean square* sebesar 0.96. Nilai tersebut berada dalam rentang *infit mean square* yang dapat diterima, sehingga dapat disimpulkan bahwa alat ukur integritas memiliki model yang *fit* dengan data.

4.2.2. Pengujian Item Fit

Suatu *item* dikatakan *fit* jika nilai *infit mean square* berada pada rentang 0,77 dan 1,30 (Adams dan Khoo, 1993). Pengujian *item fit* memberi informasi mengenai *item-item* mana yang tidak *fit*, sehingga sebaiknya tidak digunakan dalam analisis lebih lanjut. Berikut merupakan tabel yang menggambarkan pengujian *item fit* terhadap 26 *item* integritas.

Tabel 4.3. Nilai *Infit Mean Square* 26 Item

<i>Item</i>	<i>Infit Mean Square</i>	<i>Keterangan</i>
<i>Item 1</i>	0.83	<i>Item Fit</i>
<i>Item 2</i>	0.85	<i>Item Fit</i>
<i>Item 3</i>	0.80	<i>Item Fit</i>
<i>Item 4</i>	0.89	<i>Item Fit</i>
<i>Item 5</i>	1.18	<i>Item Fit</i>
<i>Item 6</i>	0.97	<i>Item Fit</i>
<i>Item 7</i>	0.89	<i>Item Fit</i>
<i>Item 9</i>	0.92	<i>Item Fit</i>
<i>Item 12</i>	1.10	<i>Item Fit</i>
<i>Item 13</i>	1.13	<i>Item Fit</i>
<i>Item 18</i>	1.04	<i>Item Fit</i>
<i>Item 19</i>	0.87	<i>Item Fit</i>
<i>Item 20</i>	0.98	<i>Item Fit</i>
<i>Item 21</i>	1.00	<i>Item Fit</i>
<i>Item 22</i>	0.71	<i>Item Tidak Fit</i>
<i>Item 23</i>	1.02	<i>Item Fit</i>
<i>Item 24</i>	1.23	<i>Item Fit</i>
<i>Item 25</i>	0.81	<i>Item Fit</i>
<i>Item 28</i>	0.81	<i>Item Fit</i>
<i>Item 30</i>	0.88	<i>Item Fit</i>
<i>Item 31</i>	1.05	<i>Item Fit</i>
<i>Item 33</i>	1.00	<i>Item Fit</i>
<i>Item 34</i>	0.96	<i>Item Fit</i>
<i>Item 36</i>	0.88	<i>Item Fit</i>
<i>Item 38</i>	1.03	<i>Item Fit</i>
<i>Item 39</i>	1.15	<i>Item Fit</i>

Berdasarkan tabel diketahui terdapat satu *item* yang tidak *fit* karena berada di luar rentang *infit mean square* yang dapat diterima, yaitu *item 22* yang berbunyi “Sebesar apapun resiko yang saya hadapi, saya tidak akan melakukan kecurangan”. Maka dapat disimpulkan bahwa dari 26 *item*, terdapat 25 *item* yang *fit* mengukur integritas.

4.2.3. Pengujian Kembali *Person Fit* dan *Item Fit*

Berdasarkan pengujian *item fit* diketahui terdapat satu *item* yang tidak *fit* yaitu *item 22*. Langkah selanjutnya adalah mengeliminasi *item* tersebut dan menguji kembali *person fit* dan *item fit* untuk *item* yang tersisa.

4.2.3.1. Pengujian Kembali *Person Fit*

Setelah melakukan pengujian kembali menguji model *fit* dengan mengeluarkan *item* 22, diperoleh nilai *infit mean square* untuk alat ukur integritas sebesar 0,97. Berdasarkan penjelasan sebelumnya telah diketahui bahwa model dikatakan *fit* apabila memiliki nilai *infit mean square* pada rentang 0,77 dan 1,30 (Adams dan Khoo, 1993). Untuk itu, dapat dikatakan bahwa alat ukur integritas dengan 25 *item* memiliki model yang *fit* dengan data.

4.2.3.2. Pengujian Kembali *Item Fit*

Setelah menguji model *fit* untuk 25 *item*, peneliti melakukan pengujian *item fit* untuk setiap *item* alat ukur integritas. *Item* dikatakan *fit* jika nilai *infit mean square* berada pada rentang 0,77 dan 1,30 (Adams & Khoo, 1993). Berikut tabel yang menggambarkan pengujian *item fit* terhadap 25 *item* integritas.

Tabel 4.4. Tabel Nilai *Infit Mean Square* 25 *Item*

<i>Item</i>	<i>Infit Mean Square</i>	Keterangan
<i>Item</i> 1	0.82	<i>Item Fit</i>
<i>Item</i> 2	0.85	<i>Item Fit</i>
<i>Item</i> 3	0.79	<i>Item Fit</i>
<i>Item</i> 4	0.89	<i>Item Fit</i>
<i>Item</i> 5	1.17	<i>Item Fit</i>
<i>Item</i> 6	0.96	<i>Item Fit</i>
<i>Item</i> 7	0.89	<i>Item Fit</i>
<i>Item</i> 9	0.92	<i>Item Fit</i>
<i>Item</i> 18	1.03	<i>Item Fit</i>
<i>Item</i> 19	0.87	<i>Item Fit</i>
<i>Item</i> 20	0.97	<i>Item Fit</i>
<i>Item</i> 21	0.99	<i>Item Fit</i>
<i>Item</i> 23	1.01	<i>Item Fit</i>
<i>Item</i> 24	1.23	<i>Item Fit</i>
<i>Item</i> 25	0.81	<i>Item Fit</i>
<i>Item</i> 28	0.81	<i>Item Fit</i>
<i>Item</i> 30	0.88	<i>Item Fit</i>
<i>Item</i> 31	1.04	<i>Item Fit</i>
<i>Item</i> 12	1.08	<i>Item Fit</i>
<i>Item</i> 13	1.11	<i>Item Fit</i>
<i>Item</i> 33	0.99	<i>Item Fit</i>
<i>Item</i> 34	0.95	<i>Item Fit</i>
<i>Item</i> 36	0.88	<i>Item Fit</i>
<i>Item</i> 38	1.03	<i>Item Fit</i>
<i>Item</i> 39	1.14	<i>Item Fit</i>

Berdasarkan tabel diketahui bahwa setelah mengeluarkan *item 22*, seluruh *item* alat ukur integritas berada dalam rentang *infit mean square* yang dapat diterima sehingga dapat disimpulkan bahwa seluruh *item fit* dalam mengukur integritas. Gambaran pengujian *item fit* dapat dilihat secara jelas pada lampiran 9 yang menggambarkan plot *infit mean square* dari setiap *item*.

4.2.5. Analisis *Differential Item Functioning* (DIF)

Peneliti melakukan analisis DIF untuk mendeteksi kemungkinan terjadinya bias respon terhadap *item-item* dalam alat ukur integritas. Pendeteksian DIF dilakukan menggunakan program QUEST. *Item* dikatakan terdeteksi DIF apabila memiliki tingkat kecenderungan persetujuan yang berbeda secara signifikan antara kelompok yang diperbandingkan (Adams dan Khoo, 1993). Pengujian DIF terhadap *item-item* alat ukur integritas dalam penelitian ini membandingkan kelompok berdasarkan jenis kelamin. Kelompok berdasarkan jenis kelamin terdiri dari laki-laki dan perempuan.

Tabel 4.5. Hipotesis Pengujian

Hipotesis
Ho : Tidak ada perbedaan kecenderungan persetujuan antar kelompok
Ha : Ada perbedaan kecenderungan persetujuan antar kelompok

Statistik dalam pengujian ini adalah *chi-square* dengan hipotesis *null* tidak terdapat perbedaan kecenderungan persetujuan antar kelompok dan hipotesis alternatif terdapat perbedaan kecenderungan persetujuan antar kelompok.

4.2.5.1. DIF berdasarkan jenis kelamin

Analisis DIF terhadap kelompok jenis kelamin dilakukan berdasarkan hasil penelitian Ones (1993) yang melaporkan terjadinya kemungkinan bias respon antara laki-laki dan perempuan dalam menjawab *item-item* yang mengukur integritas. Perbandingan estimasi kecenderungan persetujuan untuk kelompok responden laki-laki dan perempuan disajikan dalam tabel berikut:

Tabel 4.6. Perbandingan Estimasi *Item* Berdasarkan Jenis Kelamin

<i>Item</i>	<i>delta</i>		<i>difference</i> L-P	<i>chi-sq</i>	<i>p</i>	Keterangan
	Laki-laki	Perempuan				
<i>Item1</i>	0.10	0.13	-0.03	0.10	0.75	Tidak DIF
<i>Item2</i>	0.01	0.01	0.00	0.00	0.99	Tidak DIF
<i>Item3</i>	0.11	-0.02	0.13	1.91	0.17	Tidak DIF
<i>Item4</i>	0.76	0.68	0.08	0.78	0.38	Tidak DIF
<i>Item5</i>	-1.38	-1.24	-0.14	2.31	0.13	Tidak DIF
<i>Item6</i>	-0.87	-0.80	-0.07	0.57	0.45	Tidak DIF
<i>Item7</i>	1.09	1.05	0.04	0.18	0.67	Tidak DIF
<i>Item9</i>	0.71	0.75	-0.04	0.19	0.66	Tidak DIF
<i>Item12</i>	-0.47	-0.52	0.05	0.33	0.56	Tidak DIF
<i>Item13</i>	-0.89	-0.98	0.09	1.08	0.30	Tidak DIF
<i>Item18</i>	-0.90	-0.89	0.00	0.00	0.99	Tidak DIF
<i>Item19</i>	0.56	0.47	0.09	1.13	0.29	Tidak DIF
<i>Item20</i>	0.40	0.28	0.13	2.24	0.13	Tidak DIF
<i>Item21</i>	-0.86	-0.86	0.00	0.00	0.96	Tidak DIF
<i>Item23</i>	0.54	0.62	-0.08	1.03	0.31	Tidak DIF
<i>Item24</i>	1.28	1.43	-0.15	3.04	0.08	Tidak DIF
<i>Item25</i>	0.51	0.65	-0.14	2.51	0.11	Tidak DIF
<i>Item28</i>	0.33	0.16	0.17	3.60	0.06	Tidak DIF
<i>Item30</i>	-0.33	-0.37	0.04	0.20	0.65	Tidak DIF
<i>Item31</i>	-0.34	-0.39	0.05	0.36	0.55	Tidak DIF
<i>Item33</i>	-0.14	-0.09	-0.04	0.29	0.59	Tidak DIF
<i>Item34</i>	-0.71	-0.50	-0.20	5.77	0.02	DIF
<i>Item36</i>	-0.87	-0.99	0.11	1.45	0.23	Tidak DIF
<i>Item38</i>	0.57	0.57	0.00	0.00	0.98	Tidak DIF
<i>Item39</i>	0.77	0.85	-0.08	1.15	0.28	Tidak DIF

Dari tabel diketahui bahwa *item 34* yang memiliki nilai *p-value* < 0.05 (signifikan) terdeteksi DIF, sehingga hipotesis yang menyatakan tidak terdapat perbedaan kecenderungan persetujuan antara kelompok laki-laki dan perempuan ditolak dan hipotesis alternatif diterima. Dengan kata lain, *item* tersebut memiliki kecenderungan persetujuan yang berbeda untuk kelompok perempuan dan laki-laki. Artinya *item 34* teridentifikasi mengandung bias respon terkait dengan jenis kelamin. Gambaran lebih jelas mengenai *item-item* alat ukur integritas yang terdeteksi DIF berdasarkan jenis kelamin dapat dilihat pada lampiran 10 mengenai plot perbedaan kecenderungan persetujuan antara kelompok laki-laki dan perempuan. *Item 34* yang teridentifikasi mengandung DIF tersebut, yang berbunyi “*Saya merasa lebih produktif mengerjakan tugas-tugas ketika mendekati*

deadline”, memiliki tingkat kecenderungan persetujuan yang lebih tinggi untuk kelompok laki-laki. Dengan kata lain *item* tersebut cenderung disetujui oleh kelompok laki-laki dibandingkan perempuan.

Pengujian DIF ini memberikan gambaran bahwa *item-item* pada alat ukur integritas pada awalnya masih mengandung DIF sehingga responden dari kelompok jenis kelamin berbeda memiliki kecenderungan memberi respon yang berbeda terhadap *item* dikarenakan atribusi yang melekat pada dirinya. *Item* yang mengandung DIF tersebut dieliminasi dari alat ukur untuk mengurangi bias pada hasil pengukuran.

4.2.6. Deskripsi Karakteristik Setiap *Band-Scale* Alat Ukur Integritas

Berdasarkan hasil pengujian *item fit* dan analisis DIF terhadap seluruh *item* pada alat ukur integritas, peneliti memutuskan menghilangkan *item* 34 sehingga total *item* alat ukur integritas menjadi 24 *item*. Dari 24 *item* tersebut, peneliti membuat deskripsi karakteristik setiap *band-scale* alat ukur integritas. Tidak seperti di dalam teori tes klasik dimana skor tes diinterpretasi dengan membandingkan posisinya dalam kelompok normatif, dalam IRT skor tes diinterpretasi berdasarkan *item*. Pembuatan deskripsi karakteristik *band-scale* alat ukur didasarkan pada *domain reference* atau *criterion reference*. Penyekoran tes menggunakan *criterion reference* dilakukan setelah melalui proses kalibrasi indeks *threshold* untuk setiap *item*. Kemudian diestimasi pula skala *true score* berbasiskan *item* yang selanjutnya dibuat rentang *true-score* untuk menjadi skala penafsiran skor. Untuk menafsirkan setiap skor tes, diidentifikasi *item-item* yang mewakili rentang *true-score* tertentu dan dibuat deskripsi tentang karakteristik yang direpresentasikan oleh himpunan *item* pada setiap rentang *true score* tersebut. Dalam pembuatan deskripsi karakteristik *band-scale* alat ukur integritas ini, peneliti mengalikan setiap rentang *true score* dengan SD sebesar 10 dan menjumlahkannya dengan *mean* 50. Berdasarkan tahapan atau proses tersebut, peneliti membuat deskripsi karakteristik *band-scale* seperti yang tercantum dalam tabel berikut.

Tabel 4.7. Deskripsi Karakteristik Setiap *Band-Scale* Alat Ukur Integritas

<i>Rentang True Score</i>	<i>Scaled Score</i>	Karakteristik
-1.5 s.d. -1	35-38	<p>Dari segi kejujuran, individu mempertimbangkan kejujuran meskipun memiliki kecenderungan melakukan <i>white-lie</i> jika dirasa perlu.</p> <p>Dari segi keteguhan, individu memiliki kemungkinan melakukan kecurangan jika tidak ada yang mengetahui dan jika merasa hal itu tidak akan merugikan pihak manapun.</p> <p>Dari segi <i>self-control</i>, individu cenderung menjaga jarak dan menjauhi hal-hal yang beresiko memancing emosi negatifnya.</p> <p>Dari segi <i>self-esteem</i>, individu masih cenderung meragukan apakah nilai-nilai moral yang ia percaya tepat.</p>
-1 s.d. 0	39-45	<p>Dari sisi kejujuran, individu benar-benar mementingkan kejujuran tanpa mempertimbangkan kepentingan diri sendiri serta dampaknya terhadap komitmen dengan orang lain</p> <p>Dari sisi keteguhan, individu terlalu berpegang teguh pada prinsip tanpa mempertimbangkan hal-hal atau kepentingan pihak lain yang mungkin mendapat imbas dari tindakan yang ia lakukan atas dasar prinsip tersebut.</p> <p>Dari sisi <i>self-control</i>, individu belum mengatasi emosi negatif secara tepat tetapi mampu menghadapi hal-hal yang beresiko memunculkan emosi negatif.</p> <p>Dari sisi <i>self-esteem</i>, individu mampu menerima kritika meskipun menyakitkan.</p>
0 s.d. 1	46-54	<p>Dalam hal kejujuran, individu tidak menyetujui kebohongan dalam bentuk apapun, termasuk <i>white lie</i>. Kejujuran disampaikan secara tegas dan keras. Jika ada pihak yang tersakiti, individu menganggapnya sebagai kewajaran.</p> <p>Dalam hal keteguhan, individu memegang erat prinsip tetapi mampu mencari solusi yang tidak menyalahi prinsip dalam menjalankan kewajiban.</p> <p>Dalam hal <i>self-control</i>, Individu terlihat memiliki <i>self-control</i> yang baik karena mampu menutupi emosi negatif yang ia rasakan meskipun dalam dirinya masih merasa sangat cemas.</p> <p>Dalam hal <i>self-esteem</i>, individu mampu mengakui kesalahan dan kekurangannya tetapi menuntut orang lain juga mampu menghargai keberhasilan yang ia lakukan.</p>

<i>Rentang True Score</i>	<i>Scaled Score</i>	Karakteristik
1 s.d. 1.5	55-59	<p>Individu dalam kategori ini sangat menekankan kejujuran tetapi mampu mempertimbangkan cara menyampaikan kejujuran tanpa menyakiti pihak-pihak tertentu.</p> <p>Individu dalam kategori ini, memiliki sisi keteguhan yang didukung kestabilan self. Prinsip-prinsip sudah terinternalisasi dalam diri individu serta tahan menghadapi pertentangan dan ketidakpercayaan.</p> <p>Individu dalam kategori ini memiliki <i>self-control</i> yang kuat. Individu mampu dengan cepat mengatasi rasa cemas yang ia hadapi.</p> <p>Individu dalam kategori ini, memiliki sisi <i>self-esteem</i> yang stabil. Individu mampu menjalankan prinsip-prinsip moral yang dianut dan memiliki keyakinan bahwa dirinya bermoral tanpa memerlukan pujian atau penghargaan dari pihak luar.</p>

Deskripsi karakteristik setiap *band-scale* alat ukur integritas yang dipaparkan pada tabel, penggunaannya masih terbatas pada tingkat integritas menengah. *Band-scale* belum mampu mendeskripsikan karakteristik-karakteristik integritas yang berada pada tingkat ekstrim, yaitu individu yang sangat memiliki integritas dan individu yang sama sekali tidak memiliki integritas.

BAB 5

KESIMPULAN, DISKUSI DAN SARAN

Bagian ini memaparkan kesimpulan dalam rangka menjawab pertanyaan penelitian. Bab ini juga membahas mengenai beberapa hal yang perlu didiskusikan mengenai penelitian yang telah dilakukan. Pada bagian akhir peneliti memberikan saran yang berkaitan dengan pengembangan alat ukur integritas.

5.1. Kesimpulan

Penelitian ini merancang alat ukur integritas berdasarkan teori yang dikemukakan oleh Rogers (1961) karena dianggap lebih mampu menggambarkan integritas sebagai variabel psikologis. Namun masih perlu dilakukan pengujian untuk membuktikan bahwa teori integritas tersebut memiliki konstruk yang valid.

Proses ujicoba alat ukur integritas memberikan hasil perhitungan validitas alat ukur serta pengujian unidimensionalitas berdasarkan metode *confirmatory factor analysis* (CFA) yang menegaskan bahwa konstruk alat ukur integritas yang dikemukakan oleh Rogers (1961) terbukti valid.

Penelitian ini juga melakukan pengujian menggunakan metode *first order* dan *second order* CFA untuk menguji asumsi unidimensionalitas. Pengujian *first order* CFA terhadap 40 *item* alat ukur integritas memberikan hasil yang *fit* pada RMSEA, tetapi *chi-square* tidak *fit*. Hal ini bisa dianggap *fit* mengingat *chi-square* memiliki sensitifitas terhadap jumlah sampel dan normalitas data. Namun umumnya indeks *chi-square* lebih diyakini sehingga peneliti memutuskan kembali melakukan pengujian *first order* CFA dengan mengeliminasi *item-item* yang memberi kontribusi kecil. Eliminasi menghasilkan total 26 *item* integritas yang akan menjalani kembali pengujian *first order* CFA. 26 *item* tersebut juga melalui pengujian empat faktor *second order* CFA. Hasil pengujian *first order* dan empat faktor *second order* CFA menunjukkan bahwa model yang digunakan *fit* dengan data sehingga dapat dipastikan bahwa *item-item* dalam alat ukur yang dirancang dalam penelitian ini *fit* untuk mengukur integritas. Hasil pengujian *first order* dan empat faktor *second order* CFA membuktikan bahwa alat ukur integritas yang dirancang dalam penelitian ini memenuhi asumsi unidimensionalitas.

Kesimpulannya, hasil pengujian alat ukur integritas menggunakan pendekatan IRT menunjukkan bahwa model *fit* dengan data.

Terkait model *polytomous* IRT, yaitu *Rating Scale Model* (RSM), yang peneliti terapkan pada alat ukur integritas dalam penelitian ini, hasil pengujian dari program QUEST menunjukkan bahwa dari 26 *item* alat ukur integritas, terdapat satu *item* (*item* 22) yang tidak *fit* dengan model. Namun secara keseluruhan, rata-rata nilai *infit mean square* yang berada pada rentang 0,77 dan 1,30 mengindikasikan bahwa model RSM yang diterapkan *fit* dengan data.

Pengujian *differential item functioning* (DIF) yang dilakukan dengan membuat klasifikasi partisipan berdasarkan jenis kelamin, ternyata mendeteksi DIF pada *item* 34. *Item* yang berbunyi ‘*Saya merasa lebih produktif mengerjakan tugas-tugas ketika mendekati deadline*’ tersebut memiliki tingkat kecenderungan persetujuan yang lebih tinggi untuk kelompok laki-laki. Ini berarti *item* tersebut cenderung disetujui oleh kelompok laki-laki dibandingkan perempuan. *Item* yang terdeteksi DIF disinyalir dapat mempengaruhi hasil pengukuran terhadap integritas. Partisipan dengan level *trait* (*ability*) yang sama dalam kelompok yang berbeda akan memiliki peluang (*probability*) yang berbeda untuk memilih kategori tertentu (“Sangat Sesuai”, “Sesuai”, “Tidak Sesuai” atau “Sangat Tidak Sesuai”) pada *item* yang terdeteksi DIF, maka *item* tersebut dieliminasi. Berdasarkan seluruh analisis dan pengujian yang dilakukan terhadap 26 *item* alat ukur integritas yang dirancang dalam penelitian ini, satu *item* (*item* 22) *didrop* karena tidak *fit* dan satu *item* lain (*item* 34) *didrop* karena terdeteksi DIF. Maka terdapat total 24 *item* yang digunakan sebagai alat ukur integritas.

Peneliti kemudian merancang norma alat ukur integritas berdasarkan *domain referenced*, yaitu melakukan penyekoran tes setelah melalui proses kalibrasi dan estimasi skala *true score* (θ). *Item-item* yang mewakili setiap rentang *true score* diidentifikasi dan dibuat deskripsi karakteristiknya, sehingga dapat disimpulkan bahwa alat ukur yang dikembangkan dalam penelitian ini merupakan alat ukur integritas yang reliabel, valid, memiliki model dan *item-item* yang *fit* serta kontekstual dengan pekerja di Indonesia.

5.2. Diskusi

Pengembangan alat ukur integritas dalam penelitian ini melibatkan dua pengujian *first order* CFA dan satu pengujian empat faktor *second order* CFA. Pengujian *first order* CFA pertama melibatkan 40 *item* integritas namun hasil pengujian hanya memenuhi kriteria *fit* pada RMSEA. Hal ini bisa dianggap *fit*. Namun indeks *chi-square* umumnya lebih diyakini sehingga peneliti memutuskan kembali melakukan pengujian *first order* CFA dengan mengeliminasi *item-item* integritas yang memberi kontribusi kecil. Proses eliminasi menghasilkan total 26 *item* integritas yang akan menjalani kembali pengujian *first order* CFA. 26 *item* integritas tersebut juga mendapat pengujian empat faktor *second order* CFA. Seluruh hasil pengujian kembali menunjukkan bahwa model *closed fit* dengan data serta disusun oleh indikator yang valid untuk mengukur konstruk integritas. Hasil juga menunjukkan bahwa 26 *item* tersebut mengukur satu variabel laten, yaitu integritas. Sehingga alat ukur integritas yang dirancang dalam penelitian ini terbukti mampu memenuhi asumsi unidimensionalitas.

Dari 26 *item* yang digunakan sebagai alat ukur integritas, satu *item* didrop karena tidak *fit* dan satu *item* lain didrop karena terdeteksi DIF. Maka diperoleh total 24 *item* yang dipertahankan dan digunakan sebagai alat ukur integritas. Selanjutnya peneliti mengembangkan deskripsi karakteristik setiap *band-scale* untuk alat ukur integritas berdasarkan *domain referenced*, namun deskripsi karakteristik tersebut perlu dievaluasi lebih lanjut. Dibutuhkan keterlibatan *content specialist* serta psikolog yang berpengalaman untuk merancang deskripsi kemampuan atau karakteristik yang direpresentasikan oleh himpunan *item* pada setiap *band-scale* agar kemampuan tiap individu yang dites dapat dideskripsikan dengan merujuk kepada *band-scale* tersebut.

Terkait jumlah *item*, tidak ada standar yang menyebutkan batasan jumlah *item* yang dapat dikatakan banyak atau sedikit sehingga sulit mengkategorikan 24 *item* alat ukur integritas yang dirancang dalam penelitian ini panjang atau pendek. Selain itu, dalam pendekatan modern (IRT) dijelaskan bahwa alat ukur yang memiliki sedikit *item* bisa lebih reliabel dibandingkan alat ukur yang memiliki *item* lebih banyak (Embretson dan Reise, 2000). Tetapi ada kemungkinan bahwa *item* dalam alat ukur integritas ini, baik ketika berjumlah 40 *item* maupun 26 *item*,

terlalu sedikit sehingga kurang mampu menggali integritas yang dimiliki oleh partisipan. Padahal jumlah 40 *item* tersebut ditetapkan berdasarkan pendapat ahli dan proses elisitasi, karena ketika diberikan alat ukur di atas 40 *item* partisipan memperlihatkan kejenuhan dan tidak berkonsentrasi. Sedangkan jumlah 26 *item* diperoleh berdasarkan hasil pengujian model *fit*.

Langkah selanjutnya yang dilakukan dalam penelitian ini bertujuan untuk mendeteksi bias respon dalam *item-item* alat ukur integritas, peneliti menggunakan analisis DIF dengan metode *Item Parameter Comparison*. Pengujian terhadap jenis kelamin menemukan satu *item* terdeteksi DIF. *Item* yang berbunyi “*Saya merasa lebih produktif mengerjakan tugas-tugas ketika mendekati deadline*” tersebut cenderung disetujui oleh kelompok laki-laki dibandingkan perempuan. Berkaitan dengan hal ini, Harter (dalam Peterson dan Seligman, 2004) menjelaskan bahwa laki-laki memiliki kecenderungan lebih terbuka dalam menyuarakan pendapat dibandingkan perempuan. Namun tidak berarti setiap perempuan pasti lebih tertutup menyuarakan pendapatnya dibandingkan laki-laki.

Terkait penggunaan IRT dalam perancangan alat ukur, Hambleton, Swaminathan dan Rogers (1991) menyatakan kemungkinan suatu data *fit* ketika menerapkan model IRT tertentu tetapi tidak *fit* ketika menerapkan model IRT yang lain. Hal ini diistilahkan sebagai *falsifiable model*. Berkaitan dengan penjelasan tersebut, penerapan *polytomous* IRT menggunakan *rating scale model* (RSM) dalam perancangan alat ukur integritas pada penelitian ini dinilai tepat karena alat ukur memiliki model yang *fit* dengan data.

Poin penting selanjutnya terkait penggunaan IRT dalam perancangan alat ukur adalah rentang *true score* ($tetha/\theta$) yang bervariasi. Alat ukur integritas dalam penelitian ini memiliki *item-item* yang mampu mengukur variasi *tetha* dari rentang -1.45 sampai dengan 1.56. Namun variasi rentang *item* lebih kecil daripada variasi kemampuan partisipan yang berada pada rentang -3.84 sampai dengan 3.79. Hal ini terjadi karena *item-item* alat ukur integritas dalam penelitian ini disinyalir kurang mampu menggali situasi yang paling memunculkan integritas partisipan. Situasi-situasi tersebut seharusnya dapat tergal melalui proses elisitasi dan *focus group discussion* (FGD) yang dilakukan peneliti. Tetapi hal ini kemungkinan terjadi karena peneliti dapat dikatakan masih kurang familiar

dengan konteks pekerja dan perusahaan di Indonesia, meskipun peneliti telah mendiskusikan topik ini dengan beberapa praktisi dan akademisi terkait.

5.3. Saran

Tuntutan terhadap alat ukur integritas dapat dipastikan semakin luas. Pengembangan sangat diperlukan untuk mendapatkan alat ukur integritas yang berkualitas, khususnya menggunakan pendekatan *item response theory* (IRT). Alat ukur integritas yang dirancang dalam penelitian ini dapat digunakan karena terbukti valid dan *fit* untuk mengukur integritas secara kontekstual terkait kondisi pekerja dan perusahaan di Indonesia. Hasil-hasil pengujian yang telah dipaparkan juga membuktikan bahwa alat ukur integritas yang dikembangkan menggunakan pendekatan modern (IRT) lebih baik dan lebih informatif daripada alat ukur yang dikembangkan menggunakan pendekatan klasik. Tetapi alat ukur integritas yang dirancang dalam penelitian ini memiliki kelemahan-kelemahan, sehingga penelitian selanjutnya diharapkan melakukan beberapa perbaikan.

Penelitian ini merancang 24 *item* integritas dengan melibatkan 1210 pekerja sebagai partisipan. Dari awal peneliti berupaya melibatkan partisipan yang memiliki latar belakang bervariasi. Namun ternyata rentang *item* masih lebih kecil daripada rentang kemampuan partisipan. Rentang yang kecil (skor homogen) kemungkinan terjadi karena kelompok partisipan terlalu homogen sehingga varians yang diperoleh kecil. Hal ini menyebabkan analisis apapun yang dilakukan dengan alat ukur ini akan memberi korelasi yang rendah, padahal dasar dari IRT dan Lisrel menggunakan data korelasi. Peneliti selanjutnya sebaiknya melibatkan kelompok sampel yang dipastikan lebih heterogen agar diperoleh varians yang lebih tinggi.

Kelemahan lain dari alat ukur yang dirancang dalam penelitian ini adalah jumlah *item* yang disinyalir masih terlalu sedikit. Penelitian selanjutnya yang berminat mengembangkan integritas juga disarankan menggunakan *item* dengan jumlah yang lebih banyak serta lebih bervariasi agar lebih mampu mengukur integritas yang dimiliki partisipan serta diperoleh deskripsi karakteristik *band-scale* yang mampu mendekati kriteria-kriteria ekstrim dari integritas.

Alat ukur integritas yang dikonstruksi dalam penelitian ini bersifat kontekstual dan ditujukan khusus bagi kepentingan promosi pekerja pada level manager. Mengingat integritas merupakan karakter positif yang memiliki sifat universal, sangat baik apabila penelitian berikutnya mengembangkan alat ukur yang mampu mengukur integritas secara umum.

Integritas secara sederhana terlihat ketika individu harus membuat keputusan di antara pilihan situasi yang kontras. Gambaran situasi tersebut diperoleh melalui proses elisitasi dan *focus group discussion* (FGD). Proses ini sangat penting dan memiliki pengaruh sangat signifikan terhadap alat ukur yang dirancang. Peneliti yang tertarik mengembangkan alat ukur integritas harus melaksanakan proses elisitasi yang mampu menggali situasi kontras terkait konteks integritas yang diharapkan. Jika memungkinkan FGD sebaiknya melibatkan kelompok-kelompok yang berbeda agar memperoleh masukan yang lebih bervariasi.

Selain itu, alat ukur integritas yang dikonstruksi dalam penelitian ini menggunakan pilihan respon yang meminta partisipan menentukan secara pasti kesesuaian atau ketidaksesuaian dirinya dengan pernyataan pada *item* alat ukur. Peneliti lain yang tertarik mengembangkan alat ukur integritas dapat mempertimbangkan penggunaan pilihan respon yang bersifat netral. Tentu sangat menarik mencermati pengaruh pilihan opsi tertentu terhadap alat ukur integritas yang dikembangkan.

Terakhir, peneliti juga menyarankan agar ada penelitian yang melakukan proses *equating* terhadap *item* politomi agar dapat diperoleh bank soal untuk *item* integritas. Dengan adanya bank soal *item-item* integritas, tentu dapat dikembangkan alat ukur integritas yang berbentuk *computerized adaptive testing* (CAT). Alat ukur integritas yang terstandarisasi akan sangat bermanfaat secara praktis dan ekonomis

DAFTAR PUSTAKA

- Anastasi, A., & Urbina S. (1997). *Psychological testing*. United States of America: Prentice-Hall International.
- Barrett, P. (2001). *Pre-employment integrity testing: Current methods, problems and solutions*. Paper presented at: British Computer Society: Information Security Specialist Group March 29th-30th, 2001, Milton Hill, Oxford.
- Blasi, A. (2004). Moral functioning: Moral understanding and personality. In D.K. Lapsey & D. Narvaez, *Moral development, self and identity* (pp. 335-348). New Jersey: Lawrence Erlbaum Associates.
- Carter, S. (1996). *Integrity*. New York: Basic Books, A Division Of Harper Collins Publishers.
- Craig, S.B., & Smith, J.A. (2000). Integrity and personality: A person-oriented investigation. In D. Norris (Chair) *Patterns, Patterns Everywhere! Application of Person-oriented Methodology to Problems in Industrial-Organizational Psychology*. Symposium presented at the annual conference of the Society for Industrial-Organizational Psychology in New Orleans, April, 2000.
- Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. New York: Rinehart and Winston Inc.
- Deci, E.L., & Ryan, R.M. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, 4, 227-268.
- DeVellis, R.F. (2003). *Scale development: Theory and application (2nd ed)*. USA: Sage Publication Inc.
- Egberink, I.J.L., & Veldkamp, B.P. (2007). The development of a computerized adaptive test for integrity. In D.J. Weiss (Ed.), *Proceedings of the 2007 GMAC Conference on Computerized Adaptive Testing*. Retrieved February, 14th 2012 from www.psych.umn.edu/psylabs/CATCentral/
- Ellingson, J.E., Sackett, P.R., & Hough, L.H. (1999). Social desirability corrections in personality measurement: Issues of applicant comparison and construct validity. *Journal Of Applied Psychology*, 84, 155-166.

- Embretson, S.E. & Reise, S.P. (2000). *Item response theory for psychologists*. New Jersey: Lawrence Erlbaum Associates, Publishers.
- Erhard, W., Jensen, M.C. and Zaffron, S. (2011) “Integridad: Un Modelo Positivo Que Incorpora Fenomenos Normativos de Moral, Etica y Legalidad – Abreviado (Integrity: A Positive Model that Incorporates the Normative Phenomena of Morality, Ethics, and Legality - Abridged) (March 18, 2011)”. *Harvard Business School NOM Unit Working Paper No. 10-061; Barbados Group Working Paper No. 10-01; Simon School Working Paper No. 10-07*. Available at SSRN: <http://ssrn.com/abstract=1756285>.
- Friedenberg, L. (2011, July). *Constructing a personality scale: A hands-on project for teaching psychological testing*. Presented at the biennial meeting of the International Conference for Teaching Psychology, Vancouver, British Columbia.
- Gough, H.G. (1990). Testing for leadership with the california psychological inventory. In K.E. Clark & M.B. Clark (Eds.), *Measures of leadership* (pp. 343-345). West Orange, NJ: Leadership Library of America.
- Guastello, S.J., & Rieke, M.L. (1991). A review and critique of honesty test research. *Behavioral Sciences And The Law*, 9, 501-523.
- Guilford, J.P., & Fruchter, B. (1978). *Fundamental statistics in psychology and education (6th ed)*. Tokyo: McGraw-Hill.
- Hambleton, R.K., Swaminathan H., & Rogers H.J. (1991). *MMSS fundamentals of item response theory (volume 2)*. California: Sage Publications.
- Harris, M.H., & Sackett, P.R. (1987). A factor analysis and item response theory analysis of an employee honesty test. *Journal Of Business And Psychology*, 2(2), pp. 122-135.
- Harris, W.G. (1987). A components analysis of a pre-employment integrity measure: A replicated study. *Psychological Reports*, 60, 1051-1055.
- Hunter, R. (In preparation). *Criterion-related validity of integrity tests in South Africa*.
- Impelman, K. (2006). *Predicting counter-productive workplace behavior: Item level analysis of an integrity test*. Ann Arbor: Proquest Information and Learning Company.

- Iwao, S. (1997). Consistency orientation and models of social behavior: Is it not time for West to meet East? *Japanese Psychological Research*, 39, 323-332.
- Joreskog, K.G., & D. Sorbom. (1996). *LISREL 8: User's reference guide*. Chicago: Scientific Software International, Inc.
- Kaplan, R.M., & Saccuzzo, D.P. (2005). *Psychological testing: Principles, applications & Issues*. CA: Thomson Wadsworth.
- Kohlberg, L. (1976). Moral stages and moralization: The cognitive-developmental approach. In T. Lickona (Ed.), *Moral development and behavior: Theory, research and social issues* (pp. 31-53). New York: Holt, Rinehart & Winston.
- Kumar, R. (1996). *Research methodology: Step-by-step guide for beginners*. California: Sage Publication Inc.
- Lapsey, D.K., & Hill, P.L. (2009). The development of the moral personality. In D. Narvaez & D.K. Lapsey, *Personality, identity and character* (pp. 185-213). Cambridge: Cambridge University Press.
- Lasson, E.D. (1992). *Preemployment honesty testing: Construct validity issues and a test of the person-situation question*. Doctoral Dissertation, Wayne State University, Detroit, MI.
- Marchus, B., & Schuler, H. (2004). Antecedents of counterproductive behavior at work: A general perspective. *Journal Of Applied Psychology*, Vol. 89, No.4, 647-660.
- Markus, H.R., Kitayama, S., & Heiman, R. (1996). Culture and basic psychological principles. In E.T. Higgins & A.W. Kruglanski (Eds.), *Social Psychology: Handbook of basic principles* (pp. 857-913). New York: Guilford Press.
- Martelli, T.A. (1988, Aug.). *Preemployment screening for honesty: The construct validity, criterion related validity and test retest reliability of a written integrity test*. Unpublished Doctoral Dissertation, Ohio University.
- Mitchell, M.L., & Janina M.J. (2012). *Research design explained 8th revised edition*. United Kingdom: Wadsworth Publishing Company.
- Morgeson, F.P., Campion, M.A., Dipboye, R.L., Hollenbeck, J.R., Murphy, K., & Schmitt, N. (2007). *Personnel Psychology*, 60, 683-729.

- Mount, M., Ilies, R., & Johnson, E. (2006). Relationship of personality traits and counterproductive work behaviors: The mediating effects of job satisfaction. *Personnel Psychology, 59*, 591-622.
- Mumford, M.D., Zaccaro, S.J., Harding, F.D., Jacobs, T.O., & Fleishman, E.A. (2000). Leadership skills for a changing world: Solving complex social problems. *Leadership Quarterly, 11*, 11-35.
- O'Bannon, M.R., Goldinger, L.A., & Appleby, G.S. (1989). *Honesty and integrity testing: A practical guide*. Atlanta, GA: Applied Information Resources.
- Ones, D.S. (1993). *Establishing construct validity for integrity tests*. Doctoral Dissertation, University of Iowa, Iowa City, IA.
- Ones, D.S., & Viswesvaran, C. (2001). Integrity tests and other criterion-focused occupational personality scales (COPS) used in personnel selection. *International Journal Of Selection And Assessment, 9*, 31-38.
- Ones, D.S., Viswesvaran, C., & Reiss, A.D. (1996). Role of social desirability in personality testing for personnel selection: The red herring. *Journal Of Applied Psychology, 81*, 660-679.
- Ones, D.S., Viswesvaran, C., & Schmidt, F. (1995). Integrity tests: Overlooked facts, resolved issues and remaining questions. *American Psychologist, 50*, 456-4567.
- Ones, D.S., Viswesvaran, C., & Schmidt, F. (2003). Personality and absenteeism: A meta-analysis of integrity tests. *European Journal Of Personality, 17*, S19-S38.
- Palanski, M.E., & Yammarino, F.J. (2007). Integrity and leadership: Clearing the conceptual confusion. *European Management Journal 25(3)*, 171-184.
- Papalia, D.E., Olds, W.S., & Feldman. (2004). *Human development (9th ed)*. New York: McGraw Hill.
- Permatasari, M. (2011). *Pengaruh gaya berpikir, integritas dan usia dalam perilaku kerja kontraproduktif*. Tesis. Tidak Diterbitkan. Depok: Fakultas Psikologi Universitas Indonesia.
- Peterson, C., & Seligman, M.E. (2004). Integrity. In *Character strengths and virtues: A handbook and classification* (pp. 249-271). Oxford: Oxford University Press.

- Petrick, J.A., & Quinn, J.F. (2000). The integrity capacity construct and moral progress in business. *Journal Of Business Ethics*, 34 (3-4), 331-343.
- Power, F.C. (2004). The moral self in community. In D.K. Lapsey & D. Narvaez, *Moral development, self and identity* (pp. 47-65). New Jersey: Lawrence Erlbaum Associates.
- Puka B. (2004). Altruisme and character. In D.K. Lapsey & D. Narvaez, *Moral development, self and identity* (pp. 161-187). New Jersey: Lawrence Erlbaum Associates.
- Rogers, C.R. (1961). *On becoming a person: A therapist's view of psychotherapy*. Boston: Houghton Mifflin.
- Schlenker D., & Forsyth, D. (1977). On the ethics of psychological research. *Journal Of Experimental Social Psychology*, 13, 369-396.
- Schlenker, B.R., Miller, M.L., & Johnson, R.M. (2009). Moral identity, integrity and personal responsibility. In D. Narvaez & D.K. Lapsey, *Personality, identity and character* (pp. 316-340). Cambridge: Cambridge University Press.
- Suh, E.M. (2001). *Culture, identity consistency and subjective well-being*. Unpublish Manuscript, University of California-Riverside.
- US Congressional Office of Technology Assessment. (1990). *The use of integrity tests for pre-employment screening, OTA-SET-442*. Washington, DC: US Government Printing Office.
- Wanek, J.E., Sackett, P.R., & Ones, D.S. (2003). Towards an understanding of integrity test similarities and difference: An item-level analysis of seven tests. *Personnel Psychology*, 56, 873-894.

LAMPIRAN



Lampiran 1: Potongan Alat Ukur Integritas

Pernyataan	Pilihan			
	SS	S	TS	STS
Sebesar apapun resiko yang saya hadapi, saya tidak akan melakukan kecurangan				
Saya mengikuti perintah atasan meskipun bertentangan dengan prinsip pribadi				
Saya bersikap ramah kepada kolega yang tidak saya sukai				
Saya bersedia mengakui kesalahan di depan orang banyak				
Saya bersedia menolong rekan kerja meski ia tidak pernah mendukung saya				
Kemarahan saya mudah terpicu oleh orang-orang di sekitar saya				
Saya merasa lebih produktif mengerjakan tugas-tugas ketika mendekati deadline				
Saya tidak melakukan kecurangan meskipun akan kalah dari saingan saya				
Saya melaksanakan yang telah saya janjikan				
Saya tidak akan berhenti mengerjakan suatu tugas sebelum selesai				
Saya akan berterus terang meskipun merusak hubungan baik				
Saya mampu menerima kritikan meskipun menyakitkan				

Lampiran 2: Indeks *Fit* Hasil Uji *First Order CFA 40 Item* Integritas

Degrees of Freedom = 740
Minimum *Fit* Function Chi-Square = 933.79 (P = 0.00)
Normal Theory Weighted Least Squares Chi-Square = 825.07 (P = 0.016)
Estimated Non-centrality Parameter (NCP) = 85.07
90 Percent Confidence Interval for NCP = (19.04 ; 159.43)
Minimum *Fit* Function Value = 9.43
Population Discrepancy Function Value (F0) = 0.86
90 Percent Confidence Interval for F0 = (0.19 ; 1.61)
Root Mean Square Error of Approximation (RMSEA) = 0.034
90 Percent Confidence Interval for RMSEA = (0.016 ; 0.047)
P-Value for Test of Close *Fit* (RMSEA < 0.05) = 0.98

Expected Cross-Validation Index (ECVI) = 9.95
90 Percent Confidence Interval for ECVI = (9.28 ; 10.70)
ECVI for Saturated Model = 16.57
ECVI for Independence Model = 140.31

Chi-Square for Independence Model with 780 Degrees of Freedom = 13810.74
Independence AIC = 13890.74
Model AIC = 985.07
Saturated AIC = 1640.00
Independence CAIC = 14034.94
Model CAIC = 1273.48

Saturated CAIC = 4596.24

Normed *Fit* Index (NFI) = 0.93

Non-Normed *Fit* Index (NNFI) = 0.98

Parsimony Normed *Fit* Index (PNFI) = 0.88

Comparative *Fit* Index (CFI) = 0.99

Incremental *Fit* Index (IFI) = 0.99

Relative *Fit* Index (RFI) = 0.93

Critical N (CN) = 89.25

Root Mean Square Residual (RMR) = 0.060

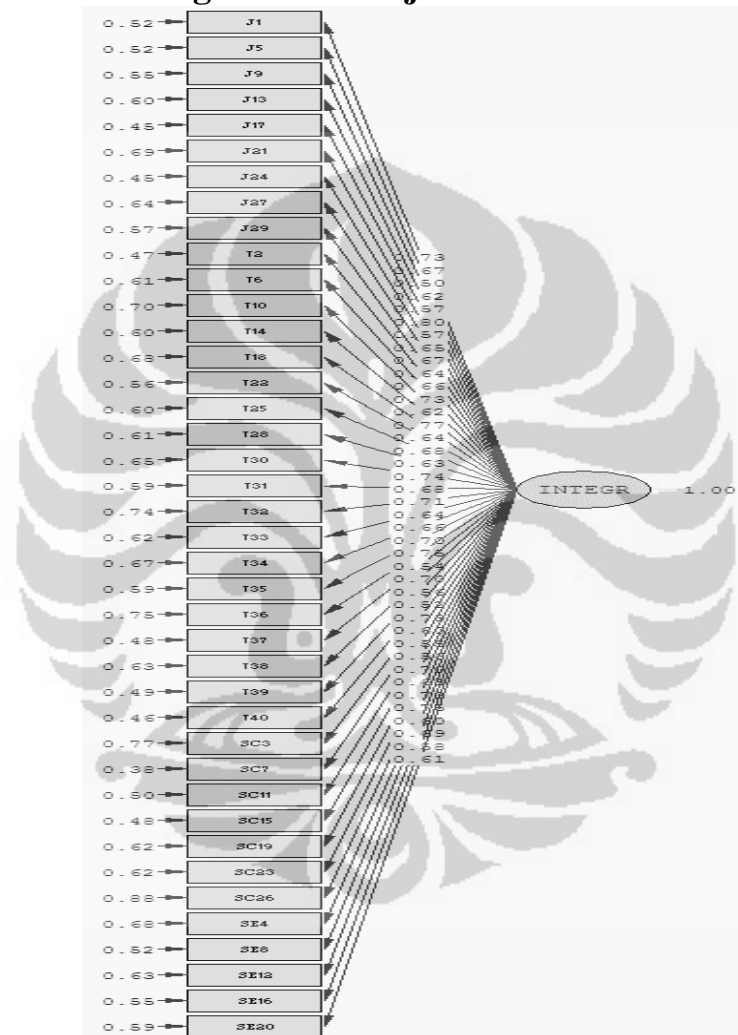
Standardized RMR = 0.059

Goodness of *Fit* Index (GFI) = 0.71

Adjusted Goodness of *Fit* Index (AGFI) = 0.67

Parsimony Goodness of *Fit* Index (PGFI) = 0.64

Lampiran 3: Path Diagram Hasil Uji *First Order CFA* 40 *Item* Integritas



Chi-Square=825.07, df=740, P-value=0.01582, RMSEA=0.034

Lampiran 4: Indeks *Fit* Hasil Uji *First Order CFA 26 Item Integritas*

Degrees of Freedom = 299

Minimum *Fit* Function Chi-Square = 334.98 (P = 0.074)

Normal Theory Weighted Least Squares Chi-Square = 299.39 (P = 0.48)

Estimated Non-centrality Parameter (NCP) = 0.39

90 Percent Confidence Interval for NCP = (0.0 ; 44.96)

Minimum *Fit* Function Value = 3.38

Population Discrepancy Function Value (F0) = 0.0039

90 Percent Confidence Interval for F0 = (0.0 ; 0.45)

Root Mean Square Error of Approximation (RMSEA) = 0.0036

90 Percent Confidence Interval for RMSEA = (0.0 ; 0.039)

P-Value for Test of Close *Fit* (RMSEA < 0.05) = 1.00

Expected Cross-Validation Index (ECVI) = 4.07

90 Percent Confidence Interval for ECVI = (4.07 ; 4.52)

ECVI for Saturated Model = 7.09

ECVI for Independence Model = 64.72

Chi-Square for Independence Model with 325 Degrees of Freedom = 6354.97

Independence AIC = 6406.97

Model AIC = 403.39

Saturated AIC = 702.00

Independence CAIC = 6500.70

Model CAIC = 590.85
Saturated CAIC = 1967.41

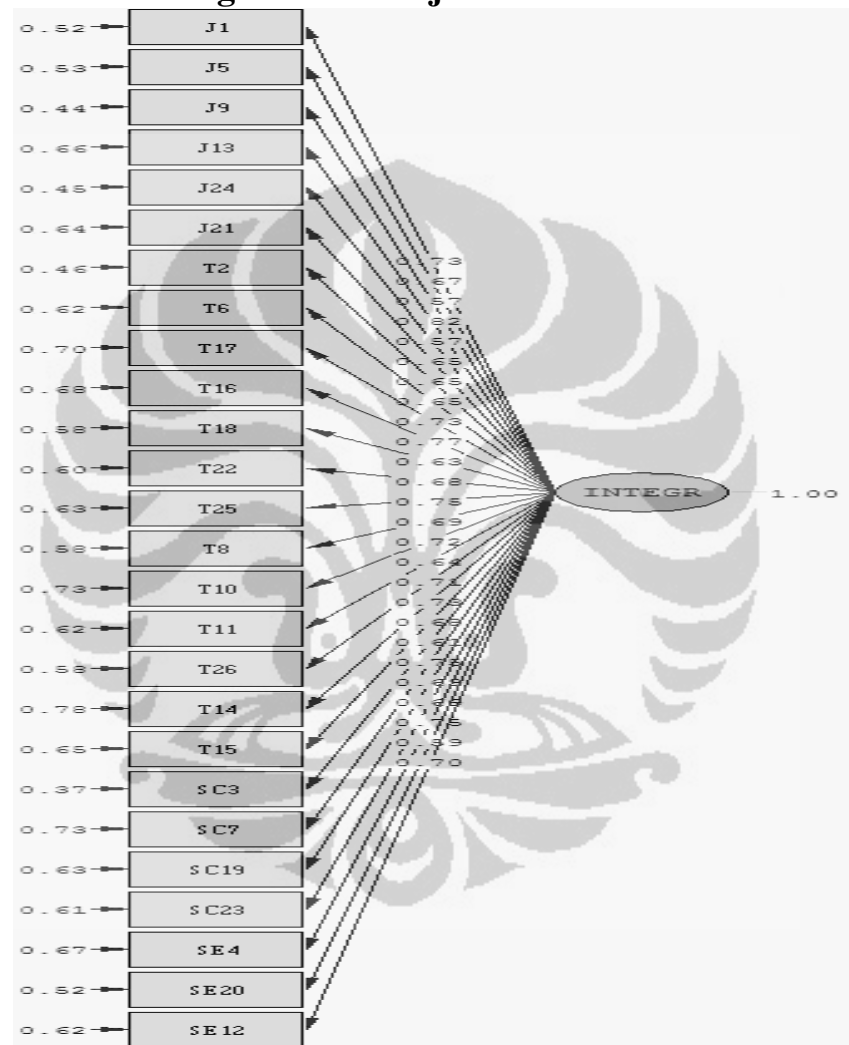
Normed *Fit* Index (NFI) = 0.95
Non-Normed *Fit* Index (NNFI) = 0.99
Parsimony Normed *Fit* Index (PNFI) = 0.87
Comparative *Fit* Index (CFI) = 0.99
Incremental *Fit* Index (IFI) = 0.99
Relative *Fit* Index (RFI) = 0.94

Critical N (CN) = 107.04

Root Mean Square Residual (RMR) = 0.056
Standardized RMR = 0.052

Goodness of *Fit* Index (GFI) = 0.81
Adjusted Goodness of *Fit* Index (AGFI) = 0.78
Parsimony Goodness of *Fit* Index (PGFI) = 0.69

Lampiran 5: Path Diagram Hasil Uji *First Order CFA* 26 Item Integritas



Chi-Square=299.39, df=299, P-value=0.48285, RMSEA=0.004

Lampiran 6: Indeks *Fit* Hasil Uji *Second Order CFA 26 Item Integritas*

Degrees of Freedom = 296

Minimum *Fit* Function Chi-Square = 333.62 (P = 0.065)

Normal Theory Weighted Least Squares Chi-Square = 297.36 (P = 0.47)

Estimated Non-centrality Parameter (NCP) = 1.36

90 Percent Confidence Interval for NCP = (0.0 ; 45.85)

Minimum *Fit* Function Value = 3.37

Population Discrepancy Function Value (F0) = 0.014

90 Percent Confidence Interval for F0 = (0.0 ; 0.46)

Root Mean Square Error of Approximation (RMSEA) = 0.0068

90 Percent Confidence Interval for RMSEA = (0.0 ; 0.040)

P-Value for Test of Close *Fit* (RMSEA < 0.05) = 0.99

Expected Cross-Validation Index (ECVI) = 4.11

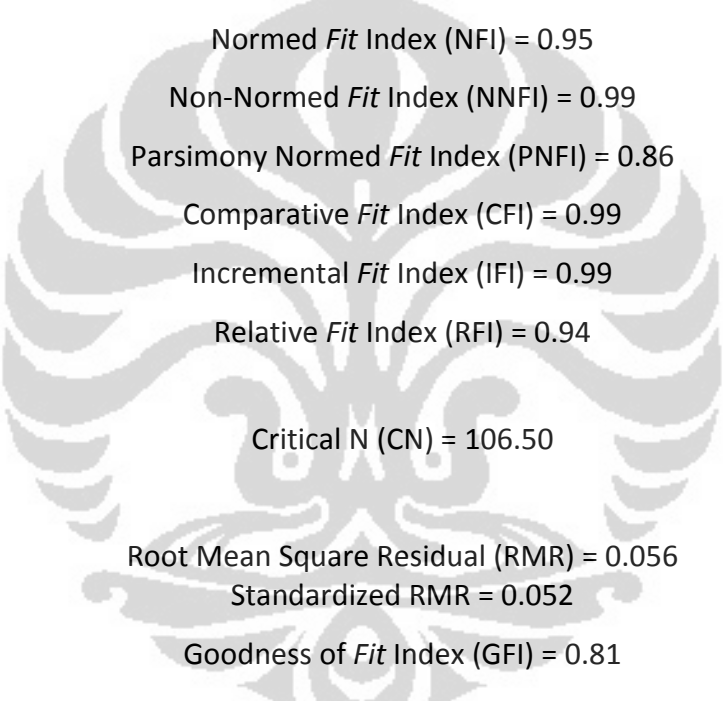
90 Percent Confidence Interval for ECVI = (4.10 ; 4.56)

ECVI for Saturated Model = 7.09

ECVI for Independence Model = 64.72

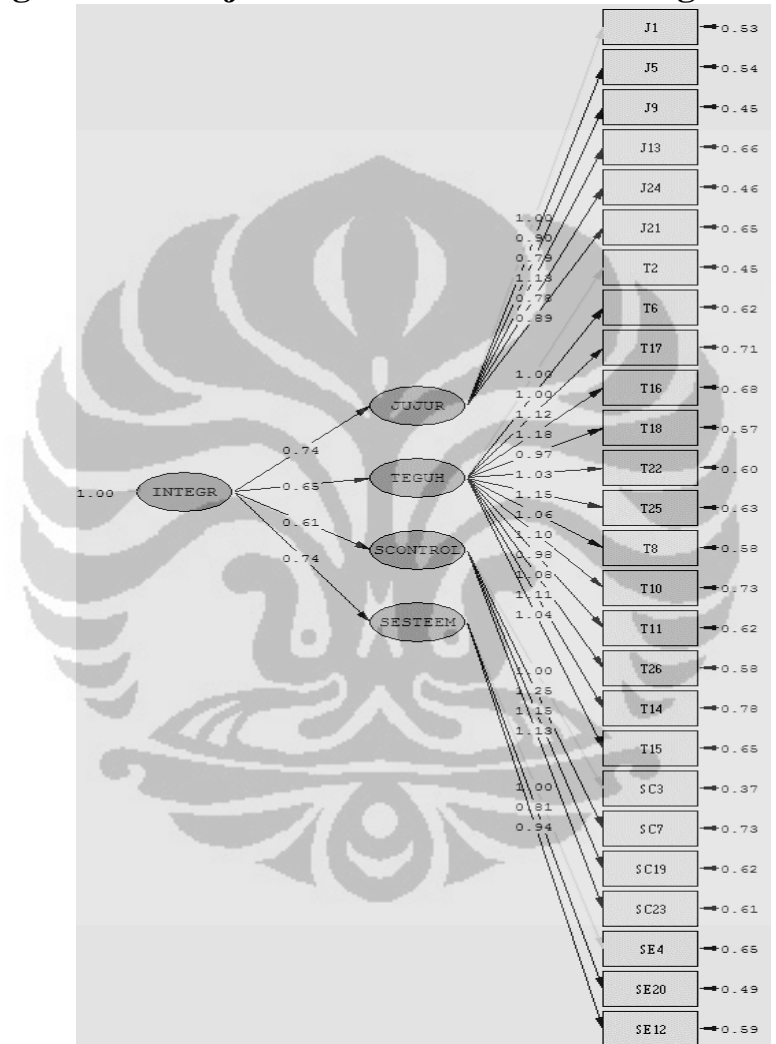
Chi-Square for Independence Model with 325 Degrees of Freedom = 6354.97

Independence AIC = 6406.97



Model AIC = 407.36
Saturated AIC = 702.00
Independence CAIC = 6500.70
Model CAIC = 605.64
Saturated CAIC = 1967.41
Normed *Fit* Index (NFI) = 0.95
Non-Normed *Fit* Index (NNFI) = 0.99
Parsimony Normed *Fit* Index (PNFI) = 0.86
Comparative *Fit* Index (CFI) = 0.99
Incremental *Fit* Index (IFI) = 0.99
Relative *Fit* Index (RFI) = 0.94
Critical N (CN) = 106.50
Root Mean Square Residual (RMR) = 0.056
Standardized RMR = 0.052
Goodness of *Fit* Index (GFI) = 0.81
Adjusted Goodness of *Fit* Index (AGFI) = 0.78
Parsimony Goodness of *Fit* Index (PGFI) = 0.69

Lampiran 7: Path Diagram Hasil Uji *Second Order* CFA 26 Integritas *Item* Pada 100 Partisipan



Chi-Square=297.36, df=296, P-value=0.46691, RMSEA=0.007

Lampiran 8: Plot *Item Fit* (25 *item*)

DATA SERIBU IRT INTEGRITAS 25 *Item*

Item Fit 30/ 5/12 13:50
all on all (N =1000 L = 25 Probability Level= .50)

	INFIIT							
	MNSQ	.63	.71	.83	1.00	1.20	1.40	1.60
1 <i>item</i> 1					*			
2 <i>item</i> 2					*			
3 <i>item</i> 3				*				
4 <i>item</i> 4					*			
5 <i>item</i> 5						*		
6 <i>item</i> 6					*			
7 <i>item</i> 7					*			
8 <i>item</i> 9					*			
9 <i>item</i> 12						*		
10 <i>item</i> 13						*		
11 <i>item</i> 18					*			
12 <i>item</i> 19					*			
13 <i>item</i> 20						*		
14 <i>item</i> 21						*		
15 <i>item</i> 23						*		
16 <i>item</i> 24							*	
17 <i>item</i> 25					*			
18 <i>item</i> 28					*			
19 <i>item</i> 30					*			
20 <i>item</i> 31						*		
21 <i>item</i> 33						*		
22 <i>item</i> 34						*		
23 <i>item</i> 36					*			
24 <i>item</i> 38						*		
25 <i>item</i> 39							*	

Lampiran 9: Plot Analisis DIF Berdasarkan Jenis Kelamin

Plot of Standardised Differences

