

Validitas dan Reliabilitas *Perceived Future Employability Scale* untuk Dewasa Muda

Anastasia Christy Matius, William Gunawan

Program Studi Psikologi, Fakultas Psikologi, Universitas Kristen Krida Wacana, Jakarta

Abstrak. Di era revolusi industri 4.0 saat ini, salah satu tantangan yang harus dihadapi berkaitan dengan lapangan pekerjaan. Terdapat penurunan kesempatan kerja, tetapi persaingan juga semakin meningkat. Angka pengangguran setiap tahun juga terus meningkat khususnya dari kalangan lulusan perguruan tinggi. Kesadaran akan kesiapan mahasiswa dalam memasuki dunia kerja setelah menyelesaikan studi merupakan suatu hal yang penting. Penelitian ini bertujuan untuk meneliti validitas dan reliabilitas alat ukur *Perceived Future Employability Scale* (PFES), guna mencari tahu apakah alat ukur tersebut dalam Bahasa Indonesia reliabel dan valid. Alat ukur ini diharapkan dapat membantu mahasiswa mendapatkan gambaran terkait kesiapan dirinya masuk ke dunia kerja setelah menyelesaikan studi, sehingga dapat lebih mempersiapkan diri serta mengetahui apa hal yang masih perlu ditingkatkan. Proses yang dilakukan meliputi pelaksanaan Diskusi Kelompok Terarah (*Focused Group Discussion/FGD*), terjemahan *forward-backward*, *expert review*, uji keterbacaan, *expert judgement*, dan pengambilan data. Desain penelitiannya berbentuk kuesioner *online*. Ukuran sampel penelitian ini adalah 467 orang, terdiri dari mahasiswa berusia 18-30 tahun. Penelitian ini melakukan pembuktian validitas konstruk (konvergen dan diskriminan) dan juga reliabilitas konsistensi internal. Hasilnya, PFES mempunyai korelasi positif yang kuat dengan alat ukur konvergen yang mengukur konstruk serupa, dan mempunyai korelasi negatif yang rendah dengan alat ukur diskriminan yang mengukur konstruk berbeda. Hasil yang diperoleh juga menunjukkan angka koefisien *Alpha Cronbach* yang tinggi, artinya skor yang dihasilkan PFES konsisten (reliabel). Dengan demikian, alat ukur PFES versi Bahasa Indonesia merupakan alat ukur yang reliabel dan valid.

Kata Kunci: alat ukur, kelayakan kerja di masa depan, mahasiswa, reliabilitas, validitas

Validity and Reliability of *Perceived Future Employability Scale* for Young Adults

Abstract. In this era of the 4th industrial revolution, one of the challenges we face is employment. There has been a decrease in job opportunities, while competition has also increased. The unemployment rate also continues to increase annually, especially among university graduates. Students need to be aware of how ready they are to enter the labor market after completing their studies. This study aims to test the validity and reliability of the *Perceived Future Employability Scale* (PFES), to identify whether the Indonesian version of the instrument is valid and reliable. PFES is expected to assist students to get an overview of how ready they are to enter the labor market, so they can prepare themselves better and know what still needs to improve. The process carried out includes a *Focused Group Discussion* (FGD), *forward-backward* translation, *expert review*, readability test, *expert judgment*, and data collection. The research design is in the form of an online questionnaire. The number of samples in this study amounted to 467 university students aged 18-30 years old. This study examined the construct validity and internal consistency reliability (Cronbach's alpha). PFES has a strong positive correlation with the instrument that measures similar construct, and has a low negative correlation with the instrument that measures different construct. The results also indicated a high level of reliability. Thus, the Indonesian version of the PFES measuring instrument is a valid and reliable measuring instrument.

Keywords: perceived future employability, reliability, tool of measurement, university student, validity

Korespondensi: William Gunawan. Email: william.gunawan@ukrida.ac.id

Saat ini Indonesia sedang memasuki era revolusi industri 4.0. Hal tersebut menghadirkan beberapa tantangan yang harus dihadapi, salah satunya adalah terkait lapangan pekerjaan. Agus Sartono, Deputi Koordinasi Bidang Pendidikan dan Agama Kemenko PMK, pada tahun 2019 memaparkan bahwa akan terjadi penurunan kesempatan kerja sebanyak 62.5% di masa sekarang ini. Hal tersebut tentu cukup mengkhawatirkan jika dibandingkan dengan jumlah pencari kerja (*jobseeker*) yang angkanya mencapai kurang lebih 31 juta orang setiap tahunnya (Mutiara, 2019).

Banyaknya jumlah pencari kerja di Indonesia juga belum sebanding dengan *Human Capital Index* (HCI) atau Indeks SDM Indonesia yang masih berada di kelas menengah, yaitu menempati posisi 87 dari 157 negara dengan angka .53 dan masih tertinggal jika dibandingkan negara-negara tetangga seperti Singapura (.88), Vietnam (.67), Malaysia (.62), Thailand (.60), dan Filipina (.55) (Pratama, 2019). Padahal, dalam era bonus demografi yang dimiliki Indonesia saat ini salah satu hal yang perlu ditingkatkan adalah Sumber Daya Manusia (SDM) yang berdaya saing dan terampil untuk memenuhi kebutuhan pasar kerja, karena banyaknya jumlah angkatan kerja yang harus dikelola tidaklah mudah.

Jumlah pengangguran juga masih menjadi salah satu masalah yang dihadapi oleh Indonesia. Beberapa tahun belakangan ini angka pengangguran lulusan diploma dan universitas malah meningkat. Sepanjang tahun 2017 hingga

2019, terdapat peningkatan jumlah pengangguran sebanyak 8.5% untuk lulusan diploma dan sebanyak 25% untuk lulusan universitas (Ulya, 2019). Jumlah lulusan perguruan tinggi yang terus bertambah tiap tahun seharusnya membuat perusahaan lebih mudah mendapatkan tenaga kerja, tetapi studi yang dilakukan oleh Willis Towers Watson pada tahun 2014 menunjukkan bahwa ternyata delapan dari sepuluh perusahaan di Indonesia masih sulit mendapatkan tenaga kerja siap pakai dari lulusan perguruan tinggi. Salah satu alasan banyak lulusan universitas sulit mendapat pekerjaan adalah ketimpangan antara kualifikasi tenaga kerja siap pakai yang dibutuhkan perusahaan dengan profil lulusan perguruan tinggi (Gewati, 2016).

Diskrepansi atau kesenjangan antara kebutuhan pasar tenaga kerja dengan ketersediaan tenaga kerja yang tidak sesuai penting untuk dikurangi guna memperkuat daya saing SDM di Indonesia. Fokus pada peningkatan kualitas SDM agar lebih mempunyai daya saing dapat dilakukan melalui pembekalan lulusan pendidikan dan pekerja dengan keterampilan teknis yang tepat, sesuai dengan yang dibutuhkan untuk memenuhi permintaan pasar tenaga kerja di masa sekarang ataupun masa depan (Handayani, 2015).

Selain membekali lulusan pendidikan dengan keterampilan yang tepat, kontribusi oleh komunitas aktivis psikologi juga dapat dilakukan dengan cara melakukan adaptasi, pembaharuan, ataupun pembuatan alat tes

psikologi yang valid dan reliabel. Suwartono (2017) menyatakan bahwa harus terdapat alat tes psikologi yang bisa mendukung dan meningkatkan daya saing serta kualitas masyarakat di Indonesia. Alat tes psikologi tersebut hendaknya dapat membantu mengidentifikasi individu terkait potensi mereka untuk memenuhi kriteria yang ada. Namun perlu diperhatikan bahwa dalam implementasinya, alat-alat tes psikologi di Indonesia perlu dilakukan adaptasi, bukan hanya membeli alat tes dari luar dan diterjemahkan. Purwono (2010) juga mendefinisikan adaptasi alat ukur sebagai “aktivitas yang dilakukan agar suatu alat ukur dapat digunakan dalam bahasa dan lingkungan budaya yang berbeda dengan bahasa dan lingkungan budaya di mana alat ukur tersebut dikembangkan”, sehingga alat ukur tidak hanya sekedar diterjemahkan tetapi juga melibatkan penyesuaian budaya.

Terdapat beberapa kelebihan mengadaptasi sebuah instrumen dibandingkan membuat instrumen yang baru. Mengadaptasi sebuah instrumen akan lebih efisien dari segi biaya dan juga waktu, lalu hasil adaptasi juga bisa menjadi suatu perbandingan antar budaya (Arafat et al, 2016). Dilakukannya adaptasi instrumen juga dapat semakin meningkatkan kemampuan instrumen tersebut untuk melakukan generalisasi dan juga memungkinkan untuk dapat mencari tahu perbedaan-perbedaan dalam populasi yang semakin beragam (Borsa et al, 2012).

Rothwell dan Arnold (2007) mengungkapkan bahwa kelayakan kerja (*employability*) merupakan salah satu hal penting yang perlu diperhatikan oleh lulusan sarjana untuk bisa mendapatkan pekerjaan dan mempertahankan pekerjaan tersebut. Salah satu hal penting yang menjadi modal utama dalam bersaing di dunia kerja berkaitan dengan seberapa *employable* individu tersebut. Kesadaran akan pentingnya kelayakan kerja harapannya dapat mendorong mahasiswa agar mengembangkan kompetensi yang dimilikinya serta mengetahui hal apa yang masih perlu ditingkatkan. Maka dari itu, berkaitan dengan kualitas SDM di Indonesia pada era bonus demografis ini yang masih perlu ditingkatkan, keberadaan suatu instrumen yang dapat digunakan untuk mengetahui dan mengukur kelayakan kerja, khususnya pada mahasiswa yang akan segera memasuki dunia kerja setelah menyelesaikan pendidikan, dapat dikatakan sebagai suatu hal yang diperlukan.

Akkermans dan Kubasch (2017) menyatakan bahwa pengembangan konseptual konstruk kelayakan kerja masih dalam tahap yang relatif awal, dan konstruk kelayakan kerja yang sudah ada biasanya berfokus pada pekerja dan orang-orang yang sudah ada di dalam dunia kerja. Berbagai definisi yang sudah ada menunjukkan bahwa konsep dari kelayakan kerja yang dirasakan (*perceived employability*) sudah cukup dikenal luas, tetapi definisi yang sudah ada belum dibangun atas dasar teori psikologis dan pendidikan yang kuat (William Gunawan et al, 2019).

Sudah terdapat beberapa alat ukur yang mengukur kelayakan kerja yang dirasakan pada karyawan atau individu yang sedang bekerja. Berntson dan Marklund (2007) menyusun sebuah skala kelayakan kerja yang dirasakan yang terdiri atas lima aitem dan mengukur *perceived skills, experience, network, personal traits*, serta *knowledge of the labour market* pada individu. Rothwell dan Arnold (2007) mengembangkan dan melakukan validasi skala *Self-Perceived Employability* (SPE) yang mencakup *internal employability* (empat butir), *external employability* (enam butir), dan satu butir yang mencakup kedua komponen tersebut. De Cuyper dan De Witte (2010) mengonstruksi skala kelayakan kerja yang terdiri dari 16 butir untuk mengukur 4 tipe kelayakan kerja yang dirasakan, yaitu *internal lateral/quantitative, internal upward/qualitative, external lateral/quantitative*, dan *external upward/qualitative*. Skala tersebut dikonstruksi untuk mengukur kelayakan kerja yang dirasakan pada orang dewasa yang bekerja atau ingin berpindah ke pekerjaan yang lain.

Selain alat ukur tersebut, Rothwell et al. (2008) serta Dacre Pool et al. (2014) juga sudah membuat skala yang bertujuan untuk mengukur kelayakan kerja bagi orang dewasa muda di konteks pendidikan. Rothwell et al. (2008) mengonstruksi skala SPE untuk mahasiswa universitas yang dimodifikasi dari skala SPE untuk orang dewasa tetapi dengan penambahan satu dimensi baru yaitu reputasi dari institusi pendidikan. Dacre Pool et al.

(2014) juga mengembangkan *Employability Development Profile* (EDP), yang secara khusus dibuat untuk pengembangan terkait mahasiswa sarjana pada institusi perguruan tinggi. Skala tersebut merupakan sebuah pengukuran berbentuk *self-report* yang berfokus pada berbagai aspek dari kelayakan kerja yang berdasarkan model *Career-EDGE*. Menurut Gunawan et al. (2019), berdasarkan beberapa kritik dari skala-skala yang sudah ada tersebut, skala-skala tersebut dikembangkan tanpa merujuk pada teori, serta hanya berfokus pada kelayakan kerja yang dirasakan saat ini saja daripada kelayakan kerja yang dirasakan di masa mendatang. Kelayakan kerja yang dirasakan di masa mendatang disini terkait khususnya ketika seorang dewasa muda telah menyelesaikan studi atau pelatihan, yang sebagaimana lebih relevan dialami kebanyakan orang dewasa muda di institusi-institusi pendidikan.

Karena belum adanya skala yang cocok untuk mengukur *Perceived Future Employability* (PFE) pada orang dewasa muda, maka hal tersebut mendasari dikembangkannya *Perceived Future Employability Scale* (PFES) untuk dewasa muda yang bertujuan untuk mengukur PFE pada orang dewasa muda (Gunawan et al., 2019). Alat ukur terkait kelayakan kerja yang sudah ada bagi orang dewasa muda seperti salah satunya yaitu SPE, kurang cocok digunakan pada orang dewasa muda seperti mahasiswa yang baru hendak akan memasuki dunia kerja, dikarenakan alat

ukur tersebut pada awalnya dibuat untuk mengukur kelayakan kerja yang dirasakan pada individu yang sudah bekerja (William Gunawan et al., 2019). Gunawan et al. (2019) menggunakan teori *future self* (Markus & Nurius, 1986) serta teori penetapan tujuan (*goal setting theory*) dan teori regulasi diri (*self regulation theory*) Bandura (2004) untuk kemudian mengusulkan konstruk PFE dan merancang alat ukur PFES untuk mengukur kelayakan kerja pada mahasiswa.

Kelayakan kerja yang dirasakan merupakan hal yang penting dalam *goal setting* dan *goal management* serta berpengaruh pada apa yang individu lakukan, pikirkan, dan rasakan terkait pekerjaan atau karir mereka (Fugate et al., 2004). Menurut Peetz dan Wilson (2014), persepsi individu akan dirinya sendiri tidak hanya terbatas pada konteks situasional saat ini, tetapi juga meluas dan meliputi persepsi dirinya di masa lalu dan juga masa depan. *Future selves* menurut Cross dan Markus (William Gunawan et al., 2019) mengacu pada representasi individu tentang seperti apa dirinya nanti di masa depan. Persepsi seseorang akan diri mereka di masa depan merefleksikan dan juga memengaruhi pikiran, perilaku, dan afek mereka di masa sekarang (Wilson et al., 2012).

Representasi karier seseorang di masa depan disebut dengan *perceived future employability* (Ellen et al., 2012). PFE bagi orang dewasa muda merujuk kepada persepsi mereka tentang seberapa *employable* dirinya di

masa depan setelah menyelesaikan edukasi, pelatihan, atau mengalami perkembangan diri (misalnya ketika sudah menyelesaikan pendidikan atau pelatihan dan siap masuk ke pasar tenaga kerja). PFE untuk orang dewasa muda secara lebih spesifik didefinisikan oleh Gunawan et al. (2021) sebagai penilaian seseorang terkait keterampilannya, pengalamannya, jejaring yang dimilikinya, sifat pribadi, pengetahuannya akan pasar tenaga kerja, dan reputasi institusi pendidikannya setelah individu tersebut menyelesaikan pendidikan dan/atau pelatihan, serta siap memasuki pasar tenaga kerja.

Menurut Gunawan et al. (2019), terdapat enam dimensi dari PFE, yaitu keterampilan di masa depan (*future skills*), pengalaman yang diharapkan (*expected experience*), jaringan/relasi di masa depan (*future networks*), *expected personal traits*, pengetahuan terkait pasar tenaga kerja di masa depan (*future labour market knowledge*), dan reputasi institusi pendidikan di masa depan (*anticipated reputation of educational institution*).

Future skills mengacu pada keterampilan-keterampilan (*skills*) dan kompetensi yang sudah akan dimiliki seseorang di masa depan nanti, yang dapat membantunya mendapatkan/mempertahankan pekerjaan, seperti misalnya kemampuan berpikir (*problem solving*, dll.) ataupun keterampilan-keterampilan teknis. *Expected experiences* mengacu pada pengalaman yang sudah akan dimiliki oleh individu di masa depan nanti, yang dapat

membantunya dalam mendapatkan/mempertahankan pekerjaannya, seperti misalnya pengalaman bekerja di bidang tertentu. *Future networks* mengacu pada relasi/jaringan yang sudah akan dimiliki individu di masa depan, yang dapat membantunya sukses di tempat kerja ataupun mendapatkan pekerjaan.

Expected personal traits mengacu pada atribut-atribut pribadi (misalnya traits/sifat dan karakteristik) yang sudah akan dimiliki individu di masa depan, yang berguna untuk mengatasi situasi penuh tantangan di dunia kerja, khususnya atribut seperti kapasitas individu untuk belajar dan berkembang. *Future labour market knowledge* mengacu pada pengetahuan individu akan pasar tenaga kerja di masa depan, seperti kesadarannya akan peluang-peluang kerja yang tersedia ataupun tren yang ada di pasar tenaga kerja. *Anticipated reputation of educational institution* mengacu pada bagaimana reputasi institusi pendidikan seseorang berhubungan dengan *employment* individu tersebut di masa depan, seperti misalnya apakah nantinya institusi pendidikan individu tersebut dapat membantunya untuk lebih mudah mendapatkan pekerjaan.

PFES merupakan alat ukur baru yang sudah dirancang mengikuti tahap-tahap konstruksi alat ukur yang tepat, memiliki angka validitas dan reliabilitas internal ($\alpha = .95$) yang tergolong baik (Gunawan et al., 2019). Oleh karena itu penelitian ini bertujuan untuk mencari tahu apakah alat ukur PFES dalam versi bahasa Indonesia reliabel dan valid.

Metode

Penelitian ini merupakan penelitian kuantitatif dan metode pengumpulan data yang dipakai adalah menggunakan kuesioner *online*, yang disebar melalui platform *limesurvey*. Kuesioner tersebut mencantumkan informasi terkait tujuan pelaksanaan penelitian, penggunaan data partisipan, kontak peneliti, serta menanyakan ketersediaan partisipan dalam penelitian (*informed consent*). Rancangan penelitian ini bertujuan untuk meneliti properti psikometri, berupa reliabilitas dan validitas dari alat ukur PFES untuk dewasa muda versi bahasa Indonesia.

Instrumen penelitian

PFES untuk dewasa muda merupakan alat ukur yang disusun oleh Gunawan et al. (2019) yang terdiri dari enam *sub-scale* dan masing-masing *sub-scale* terdiri dari empat butir. Keenam *sub-scale* itu adalah keterampilan di masa depan (*future skills*), pengalaman yang diharapkan (*expected experience*), jaringan/relasi di masa depan (*future networks*), *expected personal traits*, pengetahuan terkait pasar tenaga kerja di masa depan (*future labour market knowledge*), dan reputasi institusi pendidikan di masa depan (*anticipated reputation educational institution*).

Respon partisipan diukur berdasarkan skor tipe Likert dengan rentang satu sampai enam. Angka satu artinya sangat tidak setuju (STS) dan angka enam yang artinya sangat setuju (SS). Skor pada PFES diperoleh dengan

cara menjumlahkan skor total butir per *sub-scale* dan juga skor total aitem secara keseluruhan. Skor total per *sub-scale* bertujuan untuk mendapatkan gambaran PFE pada tiap *sub-scale*, sedangkan total skor keseluruhan bertujuan untuk mendapatkan kesimpulan terkait gambaran PFE orang tersebut secara keseluruhan. Total skor yang semakin tinggi mengindikasikan semakin tinggi PFE individu tersebut. Sebaliknya, semakin rendah total skor yang diperoleh mengindikasikan semakin rendah PFE individu tersebut.

Selain PFES, instrumen penelitian lainnya adalah *Career Ambition Scale* (CAS) dan *Career Distress Scale* (CDS). CAS digunakan untuk mencari bukti validitas konvergen dan *career distress scale* digunakan untuk mencari bukti validitas diskriminan. Menurut Rothwell et al. (2009) CAS adalah instrumen yang mengukur persepsi kesuksesan karir di masa depan dan terdiri dari enam butir. Sedangkan CDS merupakan alat ukur yang mengukur *distress* seputar karier (Creed et al., 2016) dan terdiri dari sembilan butir. Respon partisipan pada kedua alat ukur tersebut diukur berdasarkan skor tipe Likert dengan rentang satu sampai enam. Angka satu artinya sangat tidak setuju dan angka enam artinya sangat setuju.

Sampel

Kriteria partisipan dalam penelitian ini adalah mahasiswa perguruan tinggi di Indonesia dengan rentang usia 18–30 tahun. Teknik sampling yang digunakan pada

penelitian ini adalah *incidental sampling* atau biasa disebut juga *convenience sampling*. *Incidental sampling* atau *convenience sampling* adalah proses memilih sampel penelitian yang paling tersedia untuk digunakan (Cohen & Swerdlik, 2009), artinya setiap orang yang memenuhi kriteria akan diambil sebagai sampel. Teknik sampling tersebut digunakan peneliti atas pertimbangan ketersediaan serta kemudahan untuk mendapatkan data. Hal tersebut juga disebabkan oleh waktu dan juga biaya penelitian yang terbatas.

Prosedur penelitian

Prosedur penelitian dimulai dengan pertama-tama melakukan *Focused Group Discussion* (FGD) dan dilakukan atas dasar pertimbangan bahwa alat ukur yang diadaptasi merupakan alat ukur yang berasal dari luar dan masih baru serta belum pernah diterjemahkan dan diuji di Indonesia. Peneliti melakukan FGD sebanyak satu kali dengan lima orang narasumber. Setelah melakukan FGD, yang menjadi temuan atau kesimpulan peneliti adalah tidak terdapat perbedaan signifikan antara pemahaman narasumber terkait definisi PFE dengan definisi aslinya secara teoritis. Hal tersebut disimpulkan setelah peneliti melakukan perbandingan antara definisi PFE menurut narasumber dengan definisi yang asli, dan ditemukan kata-kata kunci yang sama/serupa.

Tahapan selanjutnya adalah proses penerjemahan alat ukur. Peneliti melakukan

penerjemahan dari bahasa Inggris ke bahasa Indonesia (*forward translation*) menggunakan jasa lembaga penerjemah dan hasil terjemahan tersebut kemudian diterjemahkan kembali ke dalam bahasa Inggris (*backward translation*) menggunakan jasa lembaga penerjemah yang berbeda. Seluruh hasil terjemahan (*forward dan backward*) kemudian dibandingkan kembali dengan alat ukur asli dan di-*review* oleh dosen pembimbing peneliti yang mempunyai kriteria yaitu ahli di bidang karier, paham tentang alat ukur, pernah melakukan studi di luar negeri, dan juga merupakan penyusun alat ukur PFES berbahasa Inggris (alat ukur asli).

Peneliti kemudian melakukan uji keterbacaan untuk melihat *face validity* dari alat ukur. *Face validity* atau validitas tampak adalah penilaian terkait seberapa baik sebuah tes atau instrumen alat ukur mengukur apa yang hendak diukur dilihat dari 'tampilannya' atau 'tampaknya', seperti misalnya konten dari butir-butir aitem (Cohen & Swerdlik, 2009). Peneliti melakukan uji keterbacaan pada delapan orang subjek yang sesuai dengan kriteria partisipan untuk mencari tahu apakah aitem sudah cukup jelas dan dapat dipahami dari segi susunan kalimat dan kosakata yang digunakan.

Bersamaan dengan uji keterbacaan, peneliti juga melakukan *expert judgement* kepada tiga orang ahli dengan kriteria mempunyai latar belakang psikologi serta ahli dalam bidang karier atau psikologi pendidikan, untuk menilai validitas isi (*content validity*) dari alat ukur. *Content validity* atau validitas isi

berkaitan dengan sejauh mana isi tes/alat ukur dapat mewakili sampel tingkah laku yang hendak diukur (Cohen & Swerdlik, 2009). Salah satu metode yang dapat digunakan untuk menjadi bukti terkait validitas isi adalah dengan melakukan perhitungan CVI atau *content validity index* (Wulandari et al., 2019). Pada penelitian ini peneliti melakukan perhitungan validitas isi dengan menggunakan metode CVI. Terdapat tiga parameter yang pada perhitungan CVI, yaitu parameter relevansi (*relevancy*), kepentingan (*importancy*), dan kejelasan (*clarity*). Parameter relevansi merujuk kepada sejauh mana relevansi aitem dengan konstruk yang hendak diukur. Kepentingan merujuk kepada seberapa penting aitem tersebut jika dikaitkan dengan konstruk dan konteks penelitian. Kejelasan berfokus pada kejelasan aitem dalam artian apakah aitem tersebut dapat dipahami dan sudah jelas.

Setelah melakukan uji keterbacaan dan *expert judgement*, peneliti berdiskusi kembali bersama dosen pembimbing untuk merevisi susunan kalimat dan/atau pemilihan kosakata beberapa aitem PFES dengan mempertimbangkan saran dari hasil uji keterbacaan dan *expert judgement*. Usai melakukan proses tersebut, langkah selanjutnya adalah melakukan pengambilan data dari tanggal 15 Desember 2020 hingga 2 Januari 2021. Penyebaran kuesioner *online* dilakukan dengan cara memasang iklan di media sosial *Instagram*. Target iklan disesuaikan dengan kriteria partisipan penelitian, yaitu rentang usia

diantara 18-30 tahun dan berlokasi di Indonesia. Kuesioner juga disebarluaskan secara pribadi oleh peneliti lewat aplikasi LINE dan Whatsapp. Setelah data terkumpul, peneliti mulai mengeliminasi data yang tidak memenuhi kriteria, kemudian mulai mempersiapkan semua data yang telah didapatkan untuk diolah dan masuk ke dalam tahap analisis data.

Analisis data

Teknik analisis data pada penelitian ini adalah analisis psikometrik menggunakan aplikasi *Statistical Package for the Social Sciences* (SPSS) versi 25. Terdapat beberapa analisis yang dilakukan oleh peneliti. Pertama, analisis *item-total correlation* dilakukan untuk melihat keselarasan antara skor tiap aitem dengan skor total skala. *Item total correlation* adalah metode analisis aitem yang mengorelasikan skor aitem dengan skor aitem total (Crocker & Algina, 2008). *Item total correlation* didapatkan dengan cara mencari nilai *corrected item-total correlation*. Analisis *item discrimination* dengan menggunakan *independent sample t-test* juga dilakukan untuk melihat daya beda aitem, yaitu untuk melihat apakah aitem dapat membedakan kelompok subjek yang mempunyai hasil skor tinggi (*upper group*) dan kelompok subjek yang mempunyai hasil skor rendah (*lower group*) (Cohen & Swerdlik, 2009).

Selanjutnya peneliti juga mencari bukti validitas konstruk (*construct validity*). Validitas konstruk menunjukkan sejauh mana

hasil tes mampu mengungkap suatu *trait* atau konstruk teoritik yang hendak diukurnya (Cohen & Swerdlik, 2009). Validitas konstruk terdiri dari validitas konvergen dan validitas diskriminan. Bukti validitas konvergen didapatkan dengan cara mengorelasikan tes dengan alat tes lainnya yang memiliki konstruk yang sama, sehingga hasil korelasinya diharapkan tinggi dikarenakan mengukur konstruk yang sama. Bukti validitas diskriminan didapatkan dengan cara mengorelasikan tes dengan alat tes lain yang memiliki konstruk berbeda. Alat ukur konvergen pada penelitian ini adalah CAS dan alat ukur diskriminannya adalah CDS.

Terakhir, teknik perhitungan reliabilitas yang digunakan dalam penelitian ini adalah reliabilitas konsistensi internal. Perhitungan reliabilitas konsistensi internal bertujuan untuk melihat konsistensi antar aitem dalam suatu instrumen (Urbina, 2014). Peneliti menggunakan teknik hitung koefisien alfa (α) untuk mendapatkan hasil perhitungan reliabilitas konsistensi internal.

Hasil

Penelitian ini bertujuan untuk mencari tahu dan memberikan informasi mengenai reliabilitas dan validitas alat ukur PFES untuk dewasa muda versi bahasa Indonesia.

Deskripsi partisipan

Responden yang mengisi kuesioner online pada awalnya berjumlah sebanyak 576.

Setelah mengecek kelengkapan data dan kesesuaian dengan kriteria penelitian, sebanyak 109 data di eliminasi dan didapatkan hasil akhir responden sebanyak 467 ($N = 467$). Responden yang mengisi kuesioner online merupakan mahasiswa/mahasiswa perguruan tinggi yang ada di Indonesia dengan rentang usia 18-30 tahun. Mayoritas responden berjenis kelamin perempuan dengan persentase sebesar 70.9% ($n = 331$), sedangkan responden laki-laki sebesar 29.1% ($n = 136$), seperti dapat dilihat pada Tabel 1.

Persebaran usia didominasi oleh responden yang berusia 18-20 tahun (58%) Sebanyak 53.4% responden berasal dari Perguruan Tinggi Negeri (PTN) dan sisanya sebanyak 46.7% berasal dari Perguruan Tinggi Swasta (PTS). Terdapat total sebanyak 210 perguruan tinggi dari

keseluruhan data yang didapatkan, terdiri dari 92 PTN dan 118 PTS.

Wilayah perguruan tinggi didominasi wilayah Jawa (76.7%). Berdasarkan jenjang pendidikannya, partisipan didominasi oleh mahasiswa program sarjana S1 sebanyak 92.5% ($n = 432$), tetapi terdapat juga mahasiswa/mahasiswi D3, D4, S2, dan mahasiswa/mahasiswi yang sedang menjalani pendidikan profesi. Dilihat dari tingkatan semesternya, mayoritas partisipan merupakan mahasiswa/mahasiswi semester 1 (26.6%) dan semester 3 (22.3%). Sedangkan jika dilihat dari fakultasnya, mayoritas partisipan berasal dari fakultas ilmu sosial dan politik (20.3%), ekonomi dan bisnis (17.6%), teknik (10.1%), serta keguruan dan ilmu pendidikan (9.9%).

Tabel 1*Data Demografi*

Data Demografi	<i>n</i>	%
Jenis kelamin	136	29.1
Laki-laki	331	70.9
Perempuan		
Usia	271	58.0
18-20	175	37.5
21-25	21	4.4
26-30		
Perguruan tinggi	249	53.3
PTN (Perguruan Tinggi Negeri)	218	46.7
PTS (Perguruan Tinggi Swasta)		
Wilayah perguruan tinggi	358	76.7
Jawa	21	4.5
Kalimantan	17	3.6
Sulawesi	56	12.0
Sumatera	14	3.0
DLL (Bali, NTB, NTT, Maluku, Papua)		

Catatan. N = 467.

Analisis butir

Sebelum meneliti validitas dari alat ukur, peneliti terlebih dahulu melakukan analisis butir yaitu dengan cara melakukan perhitungan *item discrimination* dan juga *item-total correlation*. Pada penelitian ini, uji *item discrimination* dilakukan dengan melakukan perhitungan *independent sample t-test*. Peneliti membagi data menjadi kelompok atas (*upper*) dan bawah (*lower*) dengan pembagian data 27% teratas dan 27% terbawah kemudian membandingkan kedua kelompok tersebut untuk mengetahui apakah perbedaan antara *upper group* dan *lower group* signifikan atau tidak. Hasil perhitungan menunjukkan bahwa seluruh butir mempunyai signifikansi dibawah .05 ($p < .05$) yang artinya seluruh butir PFES valid untuk

membedakan antara *lower group* dan *upper group*.

Pengujian *item total correlation* dilakukan untuk melihat kesesuaian antara fungsi butir dengan fungsi alat ukur secara keseluruhan. Cara menghitung *item total correlation* adalah dengan mencari nilai *corrected item-total correlation*. Nilai yang lebih besar dari .30 mengindikasikan bahwa butir tersebut valid dan butir tidak perlu di eliminasi (Sugiyono, 2019). Hasil uji *item-total correlation* PFES disajikan pada Tabel 2. Berdasarkan hasil yang didapatkan, seluruh butir PFES mempunyai nilai *item-total correlation* di atas .30 dengan rentang nilai antara .501 sampai .738. Hal tersebut menunjukkan bahwa

seluruh butir PFES valid dalam artian tiap butirnya mengukur hal yang selaras dengan

yang hendak diukur oleh alat ukur secara keseluruhan.

Tabel 2

Hasil Perhitungan Analisis Butir PFES

Sub-skala	Butir	Kode	Item-total correlation
<i>Future networks</i>	Butir 1	NW1	.584*
	Butir 2	NW2	.675*
	Butir 3	NW3	.653*
	Butir 4	NW4	.702*
<i>Expected experiences</i>	Butir 5	EX1	.611*
	Butir 6	EX2	.716*
	Butir 7	EX3	.697*
	Butir 8	EX4	.718*
<i>Expected personal traits</i>	Butir 9	TR1	.661*
	Butir 10	TR2	.719*
	Butir 11	TR3	.646*
	Butir 12	TR4	.632*
<i>Anticipated reputation of educational institution</i>	Butir 13	ED1	.501*
	Butir 14	ED2	.507*
	Butir 15	ED3	.564*
	Butir 16	ED4	.621*
<i>Future labour market knowledge</i>	Butir 17	LB1	.700*
	Butir 18	LB2	.704*
	Butir 19	LB3	.693*
	Butir 20	LB4	.637*
<i>Future skills</i>	Butir 21	SK1	.709*
	Butir 22	SK2	.738*
	Butir 23	SK3	.709*
	Butir 24	SK4	.718*

Catatan. Uji item discrimination menggunakan perhitungan *independent sample t-test*.

* $p < .05$

Validitas tangpang

Peneliti menggunakan uji keterbacaan untuk melihat *face validity* dari alat ukur. Partisipan uji keterbacaan berjumlah delapan orang dan mereka diminta untuk memberikan penilaian pada tiap butir dengan tiga opsi pilihan yaitu 'sangat dimengerti', 'kurang dimengerti', dan 'tidak dimengerti'. Partisipan yang memberikan penilaian 'kurang dimengerti' dan 'tidak dimengerti' pada suatu butir diminta untuk memberikan saran terkait apa yang sebaiknya perlu diubah atau ditambahkan pada

kalimat tersebut sehingga butir dapat dipahami dengan lebih baik.

Butir yang mendapatkan penilaian 'kurang dimengerti' atau 'tidak dimengerti' sebanyak minimal 20% dari partisipan perlu dievaluasi kembali. Maka dari itu, butir yang tidak perlu di revisi adalah butir yang diberikan penilaian 'sangat di mengerti' oleh lebih dari 80% jumlah partisipan (Sousa & Rojjanasrirat, 2011). Berdasarkan hasil uji keterbacaan, didapatkan jumlah sebanyak 14 butir PFES yang perlu direvisi.

Validitas isi

Peneliti melakukan perhitungan validitas isi dengan menggunakan metode CVI. Pada *proses expert judgement*, tiga orang ahli diminta untuk memberikan penilaian terhadap seluruh butir PFES dari skala angka satu sampai empat pada tiga parameter, yaitu relevansi (*relevancy*), kepentingan (*importance*), dan kejelasan (*clarity*).

Terdapat dua tipe perhitungan dalam mengukur kesepakatan CVI pada ketiga parameter tersebut, yaitu I-CVI (*item-CVI*) dan S-CVI (*scale-CVI*) (Beckstead; Polit, Beck, Owen dalam Wulandari, et al, 2019). I-CVI mengukur validitas tiap butir secara individual dan S-CVI mengukur validitas alat ukur secara keseluruhan. Rumus I-CVI dapat dilihat pada Gambar 2.

Gambar 2

Rumus I-CVI

$$I - CVI = \frac{\text{Jumlah ahli yang memberi rating 3 atau 4}}{\text{Jumlah ahli keseluruhan}}$$

Angka I-CVI sebesar .78 atau lebih dari hasil perhitungan tiga atau lebih ahli dapat dianggap

sebagai bukti validitas isi yang baik. Hasil perhitungan I-CVI dapat dilihat pada Tabel 3.

Tabel 3

Hasil Perhitungan I-CVI PFES

Butir	Kode	Relevansi	Kepentingan	Kejelasan
Butir 1	NW1	.67	.67	.33
Butir 2	NW2	1.00	1.00	.67
Butir 3	NW3	1.00	1.00	.33
Butir 4	NW4	1.00	1.00	.67
Butir 5	EX1	1.00	1.00	.33
Butir 6	EX2	1.00	1.00	.33
Butir 7	EX3	1.00	1.00	.67
Butir 8	EX4	.67	.67	.33
Butir 9	TR1	1.00	1.00	.33
Butir 10	TR2	1.00	1.00	.67
Butir 11	TR3	1.00	1.00	.67
Butir 12	TR4	1.00	1.00	.67
Butir 13	ED1	1.00	1.00	.33
Butir 14	ED2	1.00	1.00	.67
Butir 15	ED3	1.00	1.00	.67
Butir 16	ED4	1.00	1.00	.33
Butir 17	LB1	1.00	1.00	.67
Butir 18	LB2	.67	.67	.33
Butir 19	LB3	1.00	1.00	.67
Butir 20	LB4	1.00	1.00	.67
Butir 21	SK1	1.00	1.00	.67
Butir 22	SK2	1.00	1.00	1.00
Butir 23	SK3	1.00	1.00	.33
Butir 24	SK4	1.00	1.00	.67

Setelah itu, dilakukan pengujian S-CVI pada setiap parameter. S-CVI didapatkan dengan cara menghitung rata-rata I-CVI pada tiap parameter, yaitu jumlah skor I-CVI dibagi jumlah butir keseluruhan (Gambar 2). Angka S-CVI sebesar .90 ke atas mengindikasikan

adanya validitas isi yang baik. Alat ukur PFES memiliki angka S-CVI sebesar .96 pada parameter relevansi dan kepentingan, tetapi mempunyai angka .54 pada parameter kejelasan. Hasil perhitungan S-CVI dapat dilihat pada Tabel 4.

Tabel 4

Perhitungan S-CVI PFES Berdasarkan Parameter

Parameter	Nilai S-CVI
Relevansi	.96*
Kepentingan	.96*
Kejelasan	.54

Catatan. *S-CVI > .90

Berdasarkan hasil perhitungan tersebut, pada parameter relevansi dan kepentingan ditemukan tiga butir (butir 1, 8, 18) yang mempunyai nilai I-CVI dibawah .78, sedangkan pada parameter kejelasan seluruh butir mempunyai nilai I-CVI dibawah .78 kecuali butir 22. Berdasarkan perhitungan S-CVI, secara keseluruhan parameter relevansi dan kepentingan menunjukkan validitas yang baik, tetapi nilai parameter kejelasan mempunyai nilai dibawah .90. Maka dari itu peneliti memutuskan untuk merevisi beberapa butir terkait kejelasannya sebelum digunakan untuk mengambil data. Peneliti mengubah susunan kalimat dan kata yang kurang jelas, salah satu contohnya adalah butir nomor 2 yang berbunyi "...saya akan telah membangun jaringan sosial yang dapat membantu saya melakukan pekerjaan saya dengan baik" diubah menjadi "...saya sudah membangun jejaring sosial yang dapat membantu saya melakukan pekerjaan saya dengan baik".

Validitas konvergen dan diskriminan

Perhitungan validitas konvergen dilakukan dengan cara mengorelasikan alat ukur dengan alat ukur lain yang memiliki konstruk yang sama sehingga hasil korelasinya diharapkan tinggi dikarenakan mengukur konstruk yang sama. Sedangkan perhitungan validitas diskriminan dilakukan dengan cara mengorelasikan alat ukur dengan alat ukur lain yang memiliki konstruk berbeda. Peneliti mengorelasikan alat ukur PFES dengan alat ukur *Career Ambition* (CA) untuk mengukur validitas konvergen, lalu mengorelasikan alat ukur PFES dengan alat ukur *Career Distress* (CD) untuk mengukur validitas diskriminan. Uji korelasi *Spearman* (statistik non-parametrik) digunakan karena hasil skor alat ukur CA maupun CD menunjukkan distribusi yang tidak normal.

Setelah melakukan perhitungan, maka diperoleh hasil nilai koefisien korelasi (r)

validitas konvergen sebesar .639 ($p < .05$) dan validitas diskriminan sebesar -.318 ($p < .05$) (Tabel 5). Menurut Urbina (2014), secara umum koefisien korelasi sebesar .40 atau lebih sudah dianggap cukup. Berdasarkan pengkategorian nilai koefisien korelasi, nilai .00 – .199 mempunyai korelasi sangat rendah, nilai .20 – .399 mempunyai korelasi rendah, nilai .40 – .599 mempunyai korelasi sedang, nilai .60 – .799 mempunyai korelasi kuat, dan nilai .80 – 1.00 mempunyai korelasi sangat kuat (Sugiyono, 2019).

Hasil perhitungan menunjukkan bahwa alat ukur PFES dan CA tersebut mempunyai

hubungan positif yang signifikan dan termasuk dalam kategori kuat ($r = .639$). Meskipun mempunyai korelasi yang kuat, hal tersebut tidak berarti bahwa alat ukur PFES dan CA mengukur hal yang sama persis. Kedua alat ukur tersebut masih tetap dapat dibedakan dikarenakan tingginya hasil koefisien korelasi yang didapatkan tidak sampai terlalu dekat dengan angka 1 (sangat kuat). Sedangkan alat ukur PFES dan CD mempunyai hubungan negatif yang signifikan namun termasuk dalam kategori rendah ($r = .315$) yang artinya korelasi tersebut tidak cukup kuat dan dapat dikatakan tidak mengukur konstruk yang sama.

Tabel 5

Hasil Validitas Konvergen dan Diskriminan

Alat Ukur	r
PFES - CA	.639**
PFES - CD	-.315**

Catatan. ** $p < .001$

Reliabilitas konsistensi internal

Peneliti menggunakan perhitungan *Cronbach's Alpha* untuk mendapatkan perhitungan reliabilitas konsistensi internal. Hasil perhitungan reliabilitas *Alpha Cronbach* dapat dilihat pada Tabel 6. Berdasarkan tingkatan reliabilitasnya, nilai reliabilitas *Alpha Cronbach* sebesar .00 – .20 termasuk kurang reliabel, nilai .21 – .40 termasuk *agak* reliabel, nilai .41 – .60 termasuk cukup reliabel, nilai .61 – .80 termasuk reliabel, dan nilai .81 – 1.00 termasuk sangat reliabel (Arikunto, 1998).

Hasil perhitungan menunjukkan bahwa alat ukur PFES secara keseluruhan

mempunyai nilai koefisien *Alpha Cronbach* sebesar .905 yang artinya termasuk dalam kategori sangat reliabel. Keenam *sub-scale* PFES juga mempunyai reliabilitas yang termasuk dalam kategori sangat reliabel dengan rentang .854 sampai .897. Nilai reliabilitas yang tinggi menunjukkan skor yang dihasilkan oleh PFES dapat dipercaya dan konsisten. Alat ukur CA dan CD yang digunakan untuk mendapatkan bukti validitas konvergen dan validitas diskriminan juga termasuk reliabel. CAS mempunyai reliabilitas sebesar .689 sedangkan CDS mempunyai reliabilitas sebesar .891.

Tabel 6*Hasil Perhitungan Reliabilitas Konsistensi Internal*

Skala	Cronbach's α
<i>Perceived Future Employability Scale (PFES)</i>	.951
<i>Future networks</i>	.854
<i>Expected experiences</i>	.868
<i>Expected personal traits</i>	.857
<i>Anticipated reputation of educational institution</i>	.897
<i>Future labour market knowledge</i>	.880
<i>Future skills</i>	.888

Pembahasan

PFES bertujuan untuk mengukur persepsi orang dewasa muda terkait seberapa *employable* dirinya ketika menyelesaikan studi. PFES terdiri dari 24 butir dan enam *sub-scale* yaitu *future networks*, *expected experiences*, *expected personal traits*, *anticipated reputation of educational institution*, *future labour market knowledge*, dan *future skills*. Butir-butir PFES berbentuk skala Likert dari satu (sangat tidak setuju) sampai enam (sangat setuju) dengan perolehan skor minimal 24 dan maksimal 144.

Bukti validitas konvergen didapatkan dengan cara mengorelasikan total skor PFES dengan alat ukur CA. Hasil menunjukkan bahwa terdapat hubungan positif yang signifikan antara alat ukur PFES dengan CA dengan angka koefisien korelasi sebesar .639 ($p < .05$), dan termasuk dalam kategori kuat. Hasil tersebut sesuai dengan penelitian sebelumnya yang telah dilakukan oleh Gunawan et al. (2019) ketika melakukan validasi awal alat ukur PFES, yang juga menunjukkan bahwa ada hubungan positif yang signifikan antara PFES dengan CA ($r = .65$).

Bukti validitas diskriminan didapatkan dengan cara mengorelasikan alat ukur PFES dengan alat ukur CD. Hasil menunjukkan bahwa terdapat hubungan negatif yang signifikan antara alat ukur PFES dengan CD angka koefisien korelasi sebesar $-.315$ ($p < .05$), dan termasuk dalam kategori rendah. Hasil tersebut sesuai dengan penelitian sebelumnya yang telah dilakukan oleh Gunawan et al. (2019), yang menunjukkan bahwa ada hubungan negatif yang signifikan antara PFES dengan CD ($r = -.390$).

Metode perhitungan reliabilitas yang digunakan dalam penelitian ini adalah reliabilitas konsistensi internal dengan perhitungan *Alpha Cronbach*. Angka koefisien reliabilitas yang didapatkan adalah sebesar .951. Berdasarkan hasil yang didapatkan maka dapat disimpulkan bahwa alat ukur PFES reliabel dan mempunyai reliabilitas yang tinggi, artinya skor yang dihasilkan PFES dapat dipercaya dan konsisten. Temuan peneliti juga sesuai dengan penelitian sebelumnya yang telah dilakukan oleh Gunawan et al. (2019), yang menunjukkan

bahwa PFES mempunyai koefisien reliabilitas yang tinggi yaitu sebesar .950.

Uraian hasil pembuktian validitas dan reliabilitas yang telah dilakukan menunjukkan bahwa alat ukur PFES berbahasa Indonesia merupakan alat ukur yang valid dan reliabel. Alat ukur PFES dapat dikatakan valid dilihat dari hasil uji analisis butir (*item discrimination* dan *item-total correlation*) yang menunjukkan bahwa seluruh butir mempunyai daya beda yang baik dan juga valid, dalam artian tiap butirnya mengukur hal yang selaras dengan apa yang hendak diukur oleh alat ukur secara keseluruhan. Hasil pembuktian validitas konstruk juga menunjukkan bahwa PFES mengukur konstruk yang serupa dengan alat ukur konvergen dan mengukur konstruk yang berbeda dengan alat ukur diskriminan.

Alat ukur PFES juga dapat dikatakan reliabel dilihat dari hasil perhitungan reliabilitas konsistensi internal, yang menunjukkan bahwa angka koefisien *Alpha Cronbach* PFES tergolong cukup tinggi. Selain itu, alat ukur PFES berbahasa Indonesia juga mempunyai koefisien korelasi validitas konstruk (konvergen dan diskriminan) dan juga koefisien reliabilitas konsistensi internal yang tidak jauh berbeda dari PFES yang berbahasa Inggris (alat ukur asli).

Keterbatasan penelitian ini berkaitan dengan teknik *sampling*. Metode *sampling* yang digunakan dalam penelitian ini adalah *convenience sampling*, sehingga proporsi sampelnya kurang dapat mewakili populasi

secara keseluruhan. Menurut data demografi, berdasarkan jenis kelaminnya mayoritas responden merupakan perempuan (70.9%), berdasarkan usianya mayoritas responden berumur 18-20 tahun (58%), berdasarkan wilayah perguruan tingginya mayoritas responden merupakan mahasiswa/mahasiswi perguruan tinggi yang berada di Jawa (76.7%), dan berdasarkan jenjang perguruan tingginya mayoritas responden merupakan mahasiswa/mahasiswi sarjana S1 (92.5%).

Komposisi dari sebagian besar data demografi yang diteliti menunjukkan adanya proporsi yang tidak seimbang. Ketidakseimbangan proporsi ini menyebabkan populasi tidak dapat terwakili secara merata sehingga belum bisa dipakai menjadi suatu acuan (adanya penormaan), maka dari itu keterbatasannya adalah alat ukur ini belum dapat dipakai di seluruh Indonesia. Penelitian ini masih merupakan tahapan awal agar nantinya PFES dapat digunakan secara luas di Indonesia.

Dalam proses standarisasi, sebuah instrumen diadministrasikan ke sampel yang besar dan representatif sesuai dengan target populasi dari instrumen tersebut. Kelompok sampel standarisasi tersebut kemudian berfungsi untuk menetapkan norma (Anastasi, 1982). Perbandingan berbasis norma membutuhkan norma yang valid dengan menggunakan sampel yang mewakili target populasi (Gandek & Ware, 1998).

Sampel normatif harus mewakili target populasi se-representatif mungkin. Guna mendapatkan sampel yang representatif, beberapa peneliti mengandalkan *random sampling*. Jika data yang dikumpulkan tidak seimbang, sampel harus dikelompokkan berdasarkan kovariat yang paling memengaruhi hasil skor secara signifikan, misalnya seperti jenis kelamin, wilayah, latar belakang etnis, dan lain-lain. Stratifikasi sangat diperlukan dalam kasus yang memiliki ukuran sampel kecil, karena lebih rentan terhadap data yang tidak seimbang (Lenhard & Lenhard, 2021).

Berdasarkan studi literatur yang dilakukan oleh peneliti, peneliti menemukan beberapa poin yang perlu diperhatikan ketika hendak melakukan penormaan, yaitu ukuran sampel yang mewakili populasi (*sample size*), *feasibility*, serta komposisi sampel yang sesuai populasi. Ketiga hal tersebut saling berkaitan. Pertama, terkait *sample size*, biasanya sampel normatif dibagi menjadi sub-sampel yang lebih kecil (misalnya kelompok usia, dan lain-lain.). Pada prosedur *norming* secara konvensional, ukuran sub-sampel sebanyak 250 ($n = 250$) hingga 500 ($n = 500$) biasanya direkomendasikan untuk menghasilkan norma dengan presisi yang cukup tinggi (tergantung pada tujuan instrumen tersebut serta *scale properties*-nya) (Lenhard & Lenhard, 2021).

Kedua, *feasibility* untuk mendapatkan sampel juga perlu dipertimbangkan. Sebagai contoh, misalnya pada sub-sampel kelompok

usia, sebaiknya tidak mempunyai rentang yang terlalu luas/terlalu jauh. Di sisi lain, jika rentangnya terlalu sempit maka juga dapat menghasilkan total keseluruhan ukuran sampel yang sangat besar (karena akan ada lebih banyak kelompok usia). Oleh karena itu, antara ukuran sampel total dan rentang kelompok usia harus diseimbangkan dengan hati-hati dan mempertimbangkan *feasibility*-nya (Lenhard & Lenhard, 2021).

Ketiga, komposisi sampel norma juga merupakan suatu pertimbangan yang penting (Tett et al, 2009). Semakin besar variasinya, maka harus semakin memadai pula berbagai sub-sampel yang dipilih agar instrumen tersebut memiliki penerapan (*applicability*) yang tinggi. Contohnya, semakin besar distribusi geografis dari sampel standarisasi, maka penerapan (*applicability*) instrumen tersebut juga akan semakin tinggi (Hegde & Pomaville, 2013). Terkait ukuran sampel atau *sample size* dan komposisinya, minimal sampel harus mewakili karakteristik usia dan jenis kelamin penduduk, serta harus mewakili berbagai daerah di suatu negara (Gandek & Ware, 1998).

Terakhir, implikasi dari penelitian ini adalah alat ukur PFES telah diadaptasi ke dalam bahasa Indonesia dan bisa digunakan sebagai alat ukur yang dapat memberikan gambaran terkait PFE. Bagi mahasiswa perguruan tinggi, mengetahui seperti apa PFE yang mereka miliki dapat membantu memberikan gambaran terkait seberapa siap mereka untuk masuk ke

dunia kerja nantinya setelah menyelesaikan studi. Selain itu, karena alat ukur PFES terdiri dari enam *sub-scale*, maka skor total tiap *sub-scale* juga dapat membantu mahasiswa mengidentifikasi atau mengetahui hal apa yang masih kurang atau perlu ditingkatkan pada dirinya dan apa yang sudah cukup baik, dengan harapan bahwa setelah ia mengetahui informasi tersebut ia akan dapat lebih membekali dirinya pada aspek yang masih kurang, agar bisa menjadi lebih siap masuk ke dunia kerja nantinya.

Simpulan

Penelitian ini bertujuan untuk mencari tahu dan memberikan informasi mengenai reliabilitas dan validitas alat ukur PFES untuk dewasa muda versi bahasa Indonesia.

Berdasarkan hasil penelitian, peneliti menyimpulkan bahwa alat ukur PFES berbahasa Indonesia reliabel dan valid. Hal tersebut peneliti simpulkan berdasarkan angka reliabilitas konsistensi internal yang tinggi ($\alpha = .951$), artinya menunjukkan bahwa skor yang dihasilkan dari PFES dapat dipercaya dan konsisten (reliabel). Selain itu butirnya juga mempunyai daya beda yang baik (*item discrimination*) dan nilai *item-total correlation* seluruh butir juga mempunyai nilai diatas .30, yang artinya setiap butir valid dalam mengukur konstruk yang selaras dengan yang diukur oleh alat ukur secara keseluruhan.

Peneliti juga menyimpulkan bahwa PFES merupakan alat ukur yang valid berdasarkan

hasil perhitungan validitas konstruk. Berdasarkan bukti validitas konstruk yang didapatkan, skor total PFES berkorelasi secara positif (dalam kategori kuat) dengan alat ukur konvergen yang mengukur konstruk yang serupa, yaitu CA, dan berkorelasi negatif (dalam kategori rendah) dengan alat ukur diskriminan yang mengukur konstruk yang berbeda, yaitu CD.

Namun, meski PFES merupakan alat ukur yang valid dan reliabel, alat ukur ini belum dapat dijadikan alat ukur standar yang dapat digunakan di seluruh Indonesia secara merata dikarenakan keterbatasan pada teknik *sampling*-nya, sehingga sampel pada penelitian ini belum cukup mewakili populasi untuk dapat membuat penormaan.

Saran

Keterbatasan penelitian ini berkaitan dengan teknik *sampling* yang dipakai. Metode *sampling* yang digunakan dalam penelitian ini adalah *convenience sampling* sehingga jika dilihat dari data demografi, proporsi sampel yang didapatkan kurang representatif atau kurang dapat mewakili populasi secara keseluruhan. Responden pada penelitian ini mayoritas berjenis kelamin perempuan, kemudian dari asal perguruan tinggi didominasi oleh perguruan tinggi yang berada di wilayah Jawa, dan mayoritas merupakan mahasiswa sarjana S1. Maka dari itu keterbatasannya adalah alat ukur ini belum dapat dipakai di seluruh Indonesia dan belum dapat

distandarisasi hasil skornya untuk membuat norma.

Penelitian selanjutnya dapat menggunakan teknik *sampling* lain yang lebih mumpuni dan melakukan *upscale*, dalam artian mencari sampel yang lebih banyak agar lebih mewakili populasi Indonesia secara keseluruhan, serta lingkup wilayah yang lebih luas dengan menjangkau tidak hanya perguruan tinggi yang ada di wilayah Jawa tetapi juga perguruan tinggi yang berada di luar Jawa. Penelitian selanjutnya juga dapat lebih berfokus pada mahasiswa S1 agar rentang usianya tidak terlalu luas.

Terkait prosedur pelaksanaan penelitian, peneliti juga menyarankan bahwa penelitian selanjutnya melakukan pengecekan kembali terhadap keterbacaan butir-butir PFES, dikarenakan pada penelitian ini baru dilakukan uji keterbacaan sebanyak satu kali dan belum sempat diujikan kembali setelah direvisi. Uji keterbacaan disarankan untuk dilakukan kembali guna mengetahui apakah butir-butir yang memerlukan revisi pada penelitian ini sudah cukup jelas setelah direvisi, atau memerlukan revisi tambahan.

Secara teoritis, variabel PFE juga dapat diteliti lebih dalam dengan variabel lainnya yang masih berhubungan dan relevan dalam konteks pendidikan dan karier, seperti misalnya bagaimana kaitannya dengan perbedaan asal fakultas dan jurusan, pengalaman bekerja atau magang yang dimiliki, *networking* yang dimiliki, dll. Penelitian selanjutnya juga bisa

mencari tahu faktor demografis apa yang paling memengaruhi PFES agar dapat dijadikan pertimbangan atau acuan ketika menentukan sub-sampel pada sampel standarisasi untuk melakukan penormaan/standarisasi hasil tes.

Metode analisis lainnya juga dapat digunakan, seperti misalnya uji analisis faktor untuk menganalisis struktur hubungan diantara variabel-variabel dalam instrumen PFES. Ataupun menggunakan model alternatif lain yang lebih mutakhir (*modern*) seperti *Item Response Theory* (IRT), atau biasa disebut teori respons butir, yang mempunyai kelebihan yaitu tidak terikat dengan sampel. Salah satu bentuk IRT yang banyak digunakan adalah *Rasch model*. Saran untuk penelitian selanjutnya dapat menggunakan model politomi pada teori respon butir, atau menggunakan Rasch model untuk data politomi (*Rasch Rating Scale Model*).

Secara praktis, PFES berbahasa Indonesia dapat digunakan pada konteks (*setting*) pendidikan, yaitu pada kalangan mahasiswa khususnya untuk mahasiswa S1 di perguruan tinggi yang berada di wilayah Jawa. PFES dapat digunakan sebagai alat ukur komplementer bersamaan dengan instrumen alat ukur karier lainnya karena bisa bermanfaat untuk memberikan gambaran terkait kesiapan mahasiswa untuk masuk ke dunia kerja nantinya, berdasarkan enam *sub-scale* yang ada. PFES juga dapat bermanfaat bagi *career center* atau pusat karier yang ada pada perguruan tinggi di Indonesia, seperti misalnya dapat digunakan sebagai salah satu sarana dalam membuat

program intervensi karier untuk meningkatkan *employability* mahasiswa.

Pendanaan

Penelitian ini dilakukan dengan bantuan dana dari LPPM (Lembaga Penelitian Dan Pengabdian kepada Masyarakat) Universitas Kristen Krida Wacana (UKRIDA).

Referensi

- Akkermans, J., & Kubasch, S. (2017). Trending topics in careers: A review and future research agenda. *Career Development International, 22*(6), 586–627. <https://doi.org/10.1108/CDI-08-2017-0143>
- Arafat, S., Chowdhury, H., Qusar, M., & Hafez, M. (2016). Cross cultural adaptation and psychometric validation of research instruments: A methodological review. *Journal of Behavioral Health, 5*(3), 129. <https://doi.org/10.5455/jbh.20160615121755>
- Arikunto, S. (1998). Prosedur penelitian: Suatu pendekatan praktek. *Rineka Cipta*.
- Bandura, A. (2004). Health promotion by social cognitive means. *Health Education and Behavior, 31*(2), 143–164. <https://doi.org/10.1177/1090198104263660>
- Berntson, E., & Marklund, S. (2007). The relationship between perceived employability and subsequent health. *Work & Stress, 21*(3), 279–292. <https://doi.org/10.1080/02678370701659215>
- Borsa, J. C., Damásio, B. F., & Bandeira, D. R. (2012). Cross-cultural adaptation and validation of psychological instruments: Some considerations. *Paidéia (Ribeirão Preto), 22*(53), 423–432. <https://doi.org/10.1590/S0103-863X2012000300014>
- Cohen, R. J., & Swerdlik, M. E. (2009). Psychological testing and assessment: An introduction to test and measurement (7th ed.). McGraw-Hill.
- Creed, P. A., Hood, M., Praskova, A., & Makransky, G. (2016). The Career Distress Scale. *Journal of Career Assessment, 24*(4), 732–746. <https://doi.org/10.1177/1069072715616126>
- Dacre Pool, L., Qualter, P., & J. Sewell, P. (2014). Exploring the factor structure of the CareerEDGE employability development profile. *Education + Training, 56*(4), 303–313. <https://doi.org/10.1108/ET-01-2013-0009>
- De Cuyper, N., & De Witte, H. (2010). Temporary employment and perceived employability: Mediation by impression management. *Journal of Career Development, 37*(3), 635–652. <https://doi.org/10.1177/0894845309357051>
- Ellen, P. S., Wiener, J. L., & Fitzgerald, M. P. (2012). Encouraging people to save for their future: Augmenting current efforts with positive visions of the future. *Journal of Public Policy & Marketing, 31*(1), 58–72. <https://doi.org/10.1509/jppm.09.089>
- Fugate, M., Kinicki, A. J., & Ashforth, B. E. (2004). Employability: A psycho-social construct, its dimensions, and applications. *Journal of Vocational Behavior, 65*(1), 14–38. <https://doi.org/10.1016/j.jvb.2003.10.005>
- Gandek, B., & Ware, J. E. (1998). Methods for validating and norming translations of health status questionnaires. *Journal of Clinical Epidemiology, 51*(11), 953–959. [https://doi.org/10.1016/S0895-4356\(98\)00086-9](https://doi.org/10.1016/S0895-4356(98)00086-9)
- Gewati, M. (2016). *Kenapa lulusan perguruan tinggi makin susah mendapat pekerjaan?* Edukasi.Kompas.Com. <https://edukasi.kompas.com/read/2016/04/23/17424071/Kenapa.Lulusan.Perguruan.Tinggi.Makin.Susah.Mendapat.Pekerjaan.?page=all>
- Gunawan, W., Creed, P. A., & Glendon, A. I. (2021). Young adults' perceived future employability: Antecedents and consequences. *International Journal for Educational and Vocational Guidance,*

- 21(1), 101–122. <https://doi.org/10.1007/s10775-020-09430-7>
- Gunawan, William, Creed, P. A., & Glendon, A. I. (2019). Development and initial validation of a perceived future employability scale for young adults. *Journal of Career Assessment, 27*(4), 610–627. <https://doi.org/10.1177/1069072718788645>
- Handayani, T. (2015). Relevansi lulusan perguruan tinggi di Indonesia dengan kebutuhan tenaga kerja di era global. *Jurnal Kependudukan Indonesia, 10*(1), 53. <https://doi.org/10.14203/jki.v10i1.57>
- Hegde, M. N., & Pomaville, F. (2013). *Assessment of communication disorders in children: Resources and protocols (2nd ed.)*. Plural Publishing Inc.
- Lenhard, W., & Lenhard, A. (2021). Improvement of norm score quality via regression-based continuous norming. *Educational and Psychological Measurement, 81*(2), 229–261. <https://doi.org/10.1177/0013164420928457>
- Markus, H., & Nurius, P. (1986). Possible selves. *American Psychologist, 41*(9), 954–969. <https://doi.org/10.1037/0003-066X.41.9.954>
- Mutiara, P. (2019). *Tantangan SDM unggul untuk Indonesia maju di era revolusi industri 4.0*. Kemenkopmk.Go.Id <https://www.kemenkopmk.go.id/tantangan-sdm-unggul-untuk-indonesia-maju-di-era-revolusi-industri-40>
- Peetz, J., & Wilson, A. E. (2014). Marking time. *Personality and Social Psychology Bulletin, 40*(1), 44–56. <https://doi.org/10.1177/0146167213501559>
- Pratama, A. M. (2019). *Indeks modal manusia Indonesia kalah jauh dari Singapura dan Vietnam*. Bandung.Kompas.Com. <https://bandung.kompas.com/read/2019/08/14/142900226/indeks-modal-manusia-indonesia-kalah-jauh-dari-singapura-dan-vietnam>
- Purwono, U. (2010). *50 tahun Himpunan Psikologi Indonesia (HIMPSI): Redefinisi psikologi Indonesia dalam keberagaman*. Himpsi Jaya.
- Rothwell, A., & Arnold, J. (2007). Self perceived employability: Development and validation of a scale. *Personnel Review, 36*(1), 23–41. <https://doi.org/10.1108/00483480710716704>
- Rothwell, A., Herbert, I., & Rothwell, F. (2008). Self-perceived employability: Construction and initial validation of a scale for university students. *Journal of Vocational Behavior, 73*(1), 1–12. <https://doi.org/10.1016/j.jvb.2007.12.001>
- Rothwell, A., Jewell, S., & Hardie, M. (2009). Self-perceived employability: Investigating the responses of post-graduate students. *Journal of Vocational Behavior, 75*(2), 152–161. <https://doi.org/10.1016/j.jvb.2009.05.002>
- Sousa, V. D., & Rojjanasrirat, W. (2011). Translation, adaptation and validation of instruments or scales for use in cross-cultural health care research: A clear and user-friendly guideline. *Journal of Evaluation in Clinical Practice, 17*(2), 268–274. <https://doi.org/10.1111/j.1365-2753.2010.01434.x>
- Sugiyono. (2019). *Metode penelitian kuantitatif, kualitatif, dan R&D*. Alfabeta.
- Suwartono, C. (2017). Alat tes psikologi konteks Indonesia: Tantangan psikologi di era MEA. *Jurnal Psikologi Ulayat, 3*(1), 1. <https://doi.org/10.24854/jpu12016-51>
- Tett, R. P., Fitzke, J. R., Wadlington, P. L., Davies, S. A., Anderson, M. G., & Foster, J. (2009). The use of personality test norms in work settings: Effects of sample size and relevance. *Journal of Occupational and Organizational Psychology, 82*(3), 639–659. <https://doi.org/10.1348/096317908X336159>
- Ulya, F. N. (2019). *BPS: Pengangguran meningkat, lulusan SMK*

mendominasi. Money.Kompas.Com.
<https://money.kompas.com/read/2019/11/05/155358926/bps-pengangguran-meningkat-lulusan-smk-mendominasi>

Urbina, S. (2014). *Essentials of psychological testing 2nd edition*. John Wiley & Sons, Inc.

Wilson, A. E., Buehler, R., Lawford, H., Schmidt, C., & Yong, A. G. (2012). Basking in projected glory: The role of subjective temporal distance in future self-appraisal. *European Journal of Social Psychology*,

42(3), 342–353. <https://doi.org/10.1002/ejsp.1863>

Wulandari, P. Y., Suminar, D. R., & Hendriani, W. (2019). Adaptasi dan validasi skala strategi sibling conflict. *Jurnal Psikologi*, 18(2), 151. <https://doi.org/10.14710/jp.18.2.151-162>



Received 12 April 2021
Revised 31 December 2021
Accepted 6 January 2022

