

UJI RELIABILITAS DAN VALIDITAS EKSTERNAL THE RAVEN'S STANDARD PROGRESSIVE MATRICES

Christiany Suwartono*¹, Cahyo Pratomo Amiseso², Restu Tri Handoyo³

¹Fakultas Psikologi, Universitas Katolik Indonesia Atma Jaya Jakarta.
Jl. Jenderal Sudirman 51, Jakarta 12930.

²Divisi Konseling dan Edukasi, Lembaga Psikologi Terapan Universitas Indonesia, Jakarta. Jl. Jenderal Sudirman 51, Jakarta 12930.

³Division of Psychiatry, University College London, London.

*christiany.suwartono@atmajaya.ac.id¹

Abstract

This study aims to provide the latest information about the reliability and validity of the Raven Standard Progressive Matrices (SPM) in the context of Indonesia. In this study, we use data (N = 583) that has been provided, with an age range between 13-61 years old. The researcher used Cronbach's Alpha method to test the reliability of SPM and the criterion validity method with the Culture Fair Intelligence Test (CFIT) as a measurement benchmark to test SPM's external validity. The results showed that SPM has reliability 0,84. SPM also proved valid with criterion validity method ($r_{(136)} = 0,64$, $r^2 = 0,41$, $p < 0,01$). Future research could explore the diagnostic utility of the SPM with specific groups, such as groups of intellectually gifted individuals and groups of individuals with disabilities intelligence.

Keywords: culture fair intelligence test, external validity, raven standard progressive matrices, reliability

Abstrak

Penelitian ini bertujuan untuk memberikan informasi terkini mengenai reliabilitas dan validitas *The Raven Standard Progressive Matrices* (SPM) dalam konteks Indonesia. Penelitian ini menggunakan data (N=583) yang telah tersedia, dengan rentang usia antara 13 – 61 tahun. Metode *Cronbach's Alpha* digunakan untuk menguji reliabilitas SPM dan metode validitas kriteria dengan *the Culture Fair Intelligence Test* (CFIT) sebagai alat ukur pembandingan untuk menguji validitas eksternal dari SPM. Hasil penelitian ini menunjukkan bahwa SPM memiliki reliabilitas sebesar 0,84. SPM juga terbukti valid dengan metode validitas kriteria ($r_{(136)} = 0,64$, $r^2 = 0,41$, $p < 0,01$). Penelitian selanjutnya dapat mengeksplorasi kegunaan diagnostik dari SPM dengan kelompok khusus, seperti kelompok individu cerdas istimewa dan kelompok individu dengan disabilitas inteligensi.

Kata kunci: *culture fair intelligence test*, *raven standard progressive matrices*, reliabilitas, validitas eksternal

Pendahuluan

The Raven's Standard Progressive Matrices (SPM) adalah alat tes inteligensi yang terdiri atas 60 soal. SPM terbagi dalam 5 set (A, B, C, D, dan E) dengan 12 soal pada masing-masing set. Soal dalam SPM didesain dalam bentuk matriks dengan bagian yang hilang. Partisipan diminta untuk melengkapi bagian yang hilang dengan memilih salah satu pilihan jawaban yang disediakan.

Pada awal pengembangannya, SPM ditujukan untuk mengukur dua komponen dari konsep *g factor* yang digagas oleh Spearman pada tahun 1923, yaitu *eductive ability* dan *reproductive ability*. Konsep *g factor* dapat dipahami sebagai kemampuan kognitif umum (*general cognitive ability*) (Raven, 2000). *Eductive ability* adalah kemampuan mencari makna dari sebuah ketidakteraturan dan mengorganisasi serta mengenali hubungan antar informasi untuk memecahkan personalan. *Reproductive ability* adalah kemampuan untuk menangkap, menarik kembali dan memproduksi kembali sebuah informasi eksplisit yang dikomunikasikan dari satu orang ke orang lain (Raven, 2000). Spearman (dalam Raven, 2008) menegaskan bahwa kedua kemampuan (*ability*) ini bukanlah faktor yang terpisah, namun secara analitis keduanya dapat dibedakan sebagai dua komponen.

SPM ditujukan untuk mencakup usia anak-anak sampai dewasa. SPM dapat diberikan pada responden usia anak saat mereka sudah memahami cara menemukan bagian yang hilang untuk melengkapi sebuah pola (Raven, 1988). Waktu administrasi SPM relatif singkat yakni, 20 menit bila diberi batas waktu, walaupun sebenarnya tidak ada batas waktu dalam administrasi alat tes ini (Raven, 2000). SPM adalah alat tes yang bersifat non-verbal, sehingga dapat diadministrasikan dalam berbagai kelompok budaya (Pind, Gunnarsdóttir, & Jóhannesson, 2003). Berdasarkan beberapa karakteristik tersebut, SPM menjadi salah satu alat tes paling populer dan telah digunakan dalam berbagai konteks budaya (Moran, 1986; Pind, dkk, 2003).

The Raven's Progressive Matrices dikembangkan pada pertengahan tahun 1930-an dan dipublikasikan pertama kali pada tahun 1938 (Raven, 1988). Sejak awal publikasinya, *Raven Progressive Matrices* mengalami beberapa kali revisi yang memunculkan dua seri baru, yaitu *Coloured Progressive Matrices* (CPM) dan *Advanced Progressive Matrices* (APM) untuk menutup kelemahan seri awal (SPM) dalam hal daya beda pada kelompok responden dengan kemampuan rendah dan responden dengan kemampuan tinggi (Raven, 2000). Raven (2000) menjelaskan bahwa kelemahan ini bersumber dari cakupan rentang usia SPM yang relatif luas. Oleh karena itu, dikembangkanlah CPM untuk mengakomodir kelompok responden dengan kemampuan rendah, yang identik dengan kelompok anak dan APM untuk kelompok dengan kemampuan tinggi yang identik dengan kelompok dewasa dan berpendidikan tinggi.

Standarisasi SPM pertama kali dilakukan pada tahun 1938 dengan menggunakan data yang berasal dari 1407 sampel anak-anak di Ipswich, Inggris, sementara norma untuk dewasa dikembangkan selama masa perang dunia kedua (Raven, 1988). Standarisasi kedua dilakukan pada tahun 1979 karena beberapa penelitian (Heron & Chown, 1967; Kyle, 1977) menunjukkan adanya kenaikan skor bila dibandingkan dengan norma Ipswich (Raven, 1983). Menyikapi hasil standarisasi kedua tahun 1979, yang memang pada akhirnya menunjukkan kenaikan bila dibandingkan dengan norma Ipswich, Raven berpendapat bahwa kenaikan skor antar waktu tidak hanya terjadi pada SPM tetapi juga pada alat tes lain, yang mengukur kemampuan eduktif. Hal ini menguatkan pendapat Flynn (1999), yang menyebutkan bahwa skor IQ mengalami

kenaikan dari waktu ke waktu. Oleh karena itu menjadi penting dalam penggunaan sebuah alat ukur inteligensi untuk memberikan interpretasi hasil tes berdasarkan norma mutakhir.

Validitas dan reliabilitas adalah dua karakteristik penting dalam sebuah tes psikologis. Salah satu cara yang dapat digunakan untuk menguji validitas sebuah alat ukur adalah dengan membandingkannya dengan tes lain yang mengukur konstruk yang sama. Dalam hal ini, validitas tes inteligensi dapat diuji dengan membandingkannya dengan tes prestasi akademik atau dengan tes inteligensi lainnya (Jensen, 1980; Raven, Raven, & Court, 1998). Korelasi *The Raven's progressive matrices* dengan tes prestasi akademik, secara umum berada dalam rentang antara 0,20 – 0,60 (Raven, Raven, & Court, 1998) dengan korelasi yang lebih tinggi ditemukan pada matematika dan ilmu alam, dibandingkan dengan bahasa atau skor prestasi akademik secara umum. Reliabilitas dapat diuji salah satunya dengan melihat konsistensi internal antara skor soal dengan skor total alat ukur.

Penelitian mengenai reliabilitas dan validitas *Raven Progressive Matrices* setidaknya dilakukan di beberapa negara berikut: Islandia (Pind dkk., 2003), Afrika Selatan (Rushton, Skuy, & Bons, 2004), dan Italia (Balboni, Naglieri, & Cubelli, 2010) dalam dua dekade terakhir. Penelitian-penelitian tersebut menunjukkan bahwa *Raven Progressive Matrices* memiliki reliabilitas dengan rentang antara 0,69 – 0,93 dan validitas dengan rentang koefisien antara 0,28 – 0,64. Publikasi tentang uji validitas dan reliabilitas *Raven Progressive Matrices* di Indonesia belum banyak dilakukan (Kumolohadi & Suseno, 2012; Mangestuti & Aziz, 2007; Sukarti, 1980; Suwartono, Hendriks, Hidajat, Halim, & Kessels, 2016), dan ketika dilakukan disarankan untuk menggunakan sampel besar dan mencakup rentang usia yang luas (Rindermann & Nijenhuis, 2012; Sukarti, 1980) agar hasil dapat mendekati representasi kelompok usia yang ditargetkan oleh *Raven Progressive Matrices* dalam konteks populasi Indonesia.

Penelitian saat ini bertujuan untuk memberikan informasi terkini mengenai reliabilitas dan validitas SPM dalam konteks Indonesia. Penelitian ini dapat dijadikan sebagai acuan untuk menentukan apakah SPM memiliki properti psikometri, dalam hal ini reliabilitas dan validitas, yang memadai. Hal ini tentu sejalan dengan upaya untuk memberi pertimbangan mengenai relevansi penggunaan SPM di Indonesia saat ini. Untuk pembuktian validitas, penelitian ini menggunakan *Cattell's Culture Fair Intelligence Test* (LPSP3 UI, 2009) sebagai kriteria eksternal untuk bukti validitas dari SPM. CFIT diasumsikan tidak terpengaruh terhadap pengalaman budaya tertentu yang bisa mempengaruhi jawaban dari peserta tes terhadap soal-soal. Soal-soal CFIT pun didesain untuk mengukur *fluid intelligence*. *Fluid intelligence* merupakan keluaran utama dari faktor biologis mempengaruhi perkembangan intelektual dan diasumsikan tidak terpengaruh dengan faktor-faktor budaya. Meskipun perbedaan individual dapat ada dalam sebuah budaya, tidak ada perbedaan dalam hal *fluid intelligence* antar budaya (Nenty & Dinero, 1981).

Metode Penelitian

Dalam penelitian ini, data mentah skor SPM diperoleh dari data yang dimiliki Lembaga Psikologi Terapan UI dan pengambilan data penelitian di Unika Atma Jaya (N=538). Data mentah skor CFIT diperoleh dari data yang dimiliki Unika Atma Jaya

dan *Bright Consulting* (N=138). Sebagaimana dijelaskan pada table 1, rentang usia partisipan berkisar antara 13 – 61 tahun.

Tabel 1.
Data Demografis Partisipan

Data demografi	SPM	CFIT
Besar sampel	538	138
Jenis kelamin		
Laki-laki	53,50	35,50
Perempuan	46,50	64,50
Usia		
Jangkauan	13,43 – 61	17,22 – 61
M	17,27	23,75
SD	5,43	7,59
Pendidikan terakhir		
SD	62,83	-
SMP	11,52	0,72
SMA	14,13	55,07
SMK	3,72	13,77
D3	0,60	2,17
S1	6,32	24,64
S2	0,90	3,62
Pekerjaan		
Karyawan	3,00	11,60
Konsultan	2,20	8,70
Mahasiswa	19,10	73,90
Pelajar	74,30	0,70
Lainnya	1,30	5,07

Penelitian ini menggunakan dua buah instrumen, yaitu: *Raven's Standard Progressive Matrices* (Raven, 2000; Raven dkk., 1998) dan *Cattell's Culture Fair Intelligence Test* format 3A (LPSP3 UI, 2009). CFIT format 3A digunakan sebagai alat ukur pembandingan untuk menguji validitas kriteria SPM. Administrasi CFIT relatif cepat dengan instruksi verbal yang detail (Colom & Abad, 2007). CFIT termasuk dalam kategori tes non-verbal yang terdiri dari empat subtes: serial, analogi, matriks dan pengkategorisasian. Dalam subtes serial, partisipan diminta untuk memilih satu dari lima pilihan yang paling benar untuk menjawab permasalahan matriks yang progresif. Dalam subtes analogi, partisipan diminta untuk mengidentifikasi dari lima gambar, gambar mana yang paling berbeda dari kelompoknya. Pada subtes matriks, partisipan diminta untuk melengkapi sebuah matriks dari gambar geometris yang disajikan. Dalam subtes pengkategorisasian, partisipan diminta untuk memilih satu dari lima pilihan yang memiliki hubungan topologi yang sama di antara bentuk-bentuk geometris yang disajikan (Nenty & Dinero, 1981). Dari hasil penelitian sebelumnya, diketahui bahwa CFIT memiliki koefisien Alpha sebesar 0,786 (LPSP3 UI, 2009).

Teknik statistik deskriptif digunakan untuk memberikan gambaran mengenai skor tes yang diperoleh dengan menyajikan tendensi sentral, standar deviasi, skor minimum dan maksimal dari tes yang digunakan. Uji *t-test dependent sample*, digunakan untuk menguji apakah ada perbedaan skor antara SPM dan CFIT. Berkenaan skor mentah kedua tes tidak sama, maka dilakukan perubahan skor total dari kedua tes menjadi z-score kemudian dilakukan uji t. Teknik yang digunakan untuk mengetahui reliabilitas dari SPM adalah teknik Alpha Cronbach Berdasarkan koefisien reliabilitas yang diperoleh dengan mencari *standard error of measurement* (Kaplan dan Saccuzzo, 2013). Validasi tipe kriteria, yaitu korelasi dengan alat tes lain yang mengukur konstruk yang sama dilakukan dengan mengkorelasikan SPM dengan CFIT dengan menggunakan teknik korelasi *Pearson product-moment*. Penentuan *shared variance* dilakukan dengan menghitung koefisien determinasi (r^2) guna mengetahui kebermaknaan dari hubungan yang telah didapatkan dan bagaimana tiap variabel dipengaruhi satu sama lainnya (Gravetter & Wallnau, 2005).

Data mentah SPM dan CFIT dikumpulkan dari basis data yang dimiliki oleh LPT UI dan pengambilan data penelitian di Unika Atma Jaya dan *Bright Consulting*. Data yang diperoleh dari LPT UI berjumlah 400 data, merupakan data dari tahun 2014-2016. Hasil pengambilan data penelitian di Unika Atma Jaya dan *Bright Consulting* pada tahun 2015 berjumlah 138 data. Data mentah kemudian diolah menjadi skor total. Data mentah yang dikumpulkan tersebut diadministrasikan dari rentang waktu tahun 2014 – 2016. Administrasi kedua alat dilakukan sesuai dengan manual untuk masing-masing alat oleh para administrator di kedua lembaga tersebut.

Hasil Dan Pembahasan

Data deskriptif

Tabel 2.

Sebaran data dari alat ukur penelitian.

	SPM	CFIT_3A	Serial	Analogi	Matriks	Kategorisasi
<i>Mean</i>	47.20	24.25	7.25	5.54	5.96	5.47
<i>Median</i>	48	25	7	5	6	6
<i>Modus</i>	47	25	7	4	6	6
<i>SD</i>	6.06	4.43	1.62	2.27	1.44	1.41
Jumlah soal	60	50	13	14	13	10
Skor <i>Minimum</i>	13	7	3	0	1	1
Skor <i>Maximum</i>	60	34	11	11	9	9

Pada Tabel 2, disajikan gambaran sebaran data dari dua tes inteligensi, yaitu *Raven's Standard Progressive Matrices* (SPM) dan *Cattell's Culture Fair Intelligence Test* (CFIT) yang digunakan dalam penelitian ini.

Pada SPM, ada partisipan yang berhasil menjawab semua soal tes sesuai dengan kunci jawaban (mendapat skor maksimal 60). Pada CFIT, tidak ada yang bisa mencapai skor maksimal, yaitu skor 50. Melihat hasil tendensi sentral (*Mean*, *Median*, dan *Modus*) antara SPM dan CFIT, maka dapat dikatakan bahwa SPM lebih dirasakan mudah untuk dijawab oleh partisipan. Hasil ini didukung dari hasil uji *t dependent sample* yang menghasilkan $t_{(137)} = 5,44$ ($p < 0,01$), dengan hasil skor standar SPM lebih tinggi daripada CFIT. Hasil SPM yang lebih tinggi ini patut mendapat pertimbangan

karena di sisi lain, hasil ini bisa dikarenakan CFIT memiliki batas waktu yang cukup ketat (12,5 menit) dibandingkan dengan SPM (20 menit). Dengan demikian, dengan SPM, partisipan memiliki kemungkinan mendapat skor yang lebih tinggi dikarenakan mereka dapat lebih tenang dalam mengerjakan soal dikarenakan batas waktu pengerjaan tes yang lebih longgar.

Dalam uji reliabilitas, besar sampel yang digunakan sejumlah 538 partisipan. Hasil analisis reliabilitas dengan teknik Alpha Cronbach menghasilkan 0,84 ($M = 47,20$, $SD = 6,06$) untuk 60 soal SPM. Berdasarkan Kaplan dan Saccuzzo (2013), alat tes ini reliabel karena koefisien Alpha Cronbach yang dihasilkan di atas 0,70. Ada pun jangkauan koefisien *corrected item-total correlation* dari 60 soal ini adalah 0,05 – 0,52 (Median = 0,26). Dalam pengukuran, kekonsistenan hasil pengukuran dipengaruhi oleh kesalahan-kesalahan dalam pengukuran, karenanya peneliti mencari nilai *standard error of measurement* untuk mengestimasi skor tes yang mungkin diperoleh seseorang. Ada pun estimasi yang dihasilkan pada tingkat kepercayaan 68% adalah $\pm 2,42$ dari skor total yang diperoleh dengan menggunakan SPM. Hal ini berarti, jika seseorang mendapatkan skor total SPM = 40, maka dengan tingkat kepercayaan 68%, estimasi skor yang sebetulnya bisa diperoleh orang tersebut antara 37,58 sampai 42,42. Hasil ini sekaligus juga memberikan bukti reliabilitas dari alat tes SPM dengan metode Alpha Cronbach.

Pada validasi kriteria, ditemukan besaran korelasi sedang dan signifikan antara SPM dan CFIT. Korelasi *Pearson* antara SPM dan CFIT menghasilkan $r_{(136)} = 0,64$, $r^2 = 0,41$, $p < 0,01$. Hal ini berarti ada hubungan yang bukan kebetulan antara SPM dan CFIT pada tingkat kepercayaan 99%. Berdasarkan teori, hubungan ini dikarenakan persamaan konstruk yang diukur oleh SPM dan CFIT dengan koefisien korelasi sebesar 0,64. Hal ini juga berarti sebesar 41% ($r^2 \cdot 100\%$) keakuratan skor SPM ditentukan oleh skor CFIT, begitu pula sebaliknya. Hasil ini sekaligus juga memberikan bukti validitas eksternal dari alat tes SPM atas dasar korelasinya yang signifikan dengan CFIT. SPM masih dapat digunakan dalam mengukur inteligensi umum, yang didefinisikan dengan konsep *g factor*. Ada bagian dari SPM dan CFIT yang mengukur hal yang sama, yaitu penalaran induktif-deduktif. Pada kedua tes, secara umum partisipan diminta untuk berpikir konseptual dan melakukan abstraksi. Hasil penelitian ini mendukung penelitian yang sebelumnya yang mengkorelasikan SPM dengan alat tes inteligensi lain, yaitu *Intelligenz Struktur Test* (Kumolohadi & Suseno, 2012; Mangestuti & Aziz, 2007) dan *Wechsler Adult Intelligence Scale* – edisi keempat (Suwartono dkk., 2016). Hasil dari analisis korelasi antara total skor SPM, CFIT beserta domain-domainnya dan usia tersaji dalam Tabel 3 berikut ini.

Tabel 3.
Matriks korelasi antar instrumen penelitian dan usia.

	SPM	CFIT_3A	Serial	Analogi	Matriks	Kategorisasi
SPM	1					
CFIT_3A	0,64**	1				
Serial	0,36**	0,66**	1			
Analogi	0,44**	0,74**	0,26**	1		
Matriks	0,51**	0,64**	0,25**	0,27**	1	
Kategorisasi	0,42**	0,60**	0,30**	0,20*	0,31**	1
Usia	0,04	-0,36**	-0,20*	-0,31**	-0,36**	-0,03

*. korelasi signifikan pada Level of Confidence 0,05 (dua-arah)

** . korelasi signifikan pada Level of Confidence 0,01 (dua-arah)

Hasil analisis menunjukkan bahwa skor total dari SPM berkorelasi signifikan dengan semua domain yang diukur CFIT, yaitu: Serial, Analogi, Matriks, and Pengkategorisasian. Berdasarkan semua domain CFIT, domain Matriks memiliki hubungan tertinggi dengan SPM, yaitu $r_{(136)} = 0,51$, $r^2 = 0,26$, $p < 0,01$. Hal ini berarti ada hubungan yang bukan kebetulan antara SPM dan subtes Matriks dari CFIT pada tingkat kepercayaan 99%. Hubungan ini dikarenakan persamaan isi dan tipe soal yang diukur oleh SPM dan subtes Matriks pada CFIT dengan koefisien korelasi sebesar 0,51. Dalam subtes Matriks dari CFIT, partisipan diminta untuk melengkapi sebuah matriks dari gambar geometris yang disajikan. Hal ini juga berarti sebesar 26% ($r^2 \cdot 100\%$) keakuratan skor SPM ditentukan oleh skor subtes Matriks pada CFIT, begitu pula sebaliknya. Kemudian, *shared variance* terbesar kedua dari domain CFIT terhadap SPM adalah Analogi. Pada domain Analogi, partisipan diminta untuk mengidentifikasi gambar yang paling berbeda dari kelompoknya. Ada pun *shared variance* ($r^2 \cdot 100\%$) subtes Analogi dari CFIT yaitu sebesar 19,36%. Hal ini dapat dipahami karena isi soal pada SPM, yakni partisipan diminta untuk menemukan kesesuaian pola visual dan desain gambar dari persoalan analogi. Dengan demikian, hal ini mendukung domain Matriks dan Analogi memiliki *shared variance* yang lebih besar dengan SPM dibandingkan domain Serial dan Kategorisasi dari CFIT (fokus dari SPM adalah matriks dan kesesuaian pola visual).

Kekuatan dari penelitian ini adalah besar sampel yang cukup besar, yaitu 538 orang dan rentang usia sampel terbesar yakni usia tingkat masa sekolah. Hal ini bisa menjadi kekuatan dari penelitian ini dalam sumbangan pembuatan norma karena pengguna terbesar dari tes ini di area pendidikan (sekolah). SPM dapat digunakan oleh lembaga layanan psikologi untuk menyusun norma populasi khusus anak sekolah, khususnya tingkat sekolah menengah pertama. Pada praktiknya, SPM merupakan alat yang masih digunakan lembaga layanan psikologi untuk menangani kasus-kasus tes minat bakat bagi siswa-siswa sekolah menengah pertama, berkaitan dengan penjurusan jika kelak melanjutkan ke jenjang sekolah menengah atas. SPM rutin digunakan dalam beberapa tahun terakhir, namun sifatnya insidental; dalam artian SPM diadministrasikan berdasarkan insiden kebutuhan penggunaan alat ukur tersebut dalam konteks seleksi atau pemetaan potensi siswa sekolah menengah. Para peserta tes minat bakat tidak diukur dalam beberapa tahun kemudian, baik dengan alat tes yang sama maupun lain. Padahal, hal tersebut bisa juga digunakan untuk menguji reliabilitas dengan metode *test-retest* dan validitas konstruk dari SPM.

Keterbatasan dari penelitian ini terkait dengan keterwakilan jenjang usia. Penelitian ini memiliki rentang usia 13,43 – 61 tahun, namun kelompok partisipan terbesar adalah usia masa sekolah dan kuliah. Oleh karenanya, penelitian selanjutnya dapat menyasar ke kelompok usia karyawan dan lebih tua. Penelitian selanjutnya juga dapat menginvestigasi kegunaan diagnostik dari SPM dengan kelompok khusus, seperti individu cerdas istimewa dan disabilitas inteligensi. Hal ini bisa dilakukan dengan analisis *The Receiver-Operating Characteristic* (ROC). Analisis ROC ini biasanya digunakan di kalangan medis untuk memvisualisasikan dan menganalisis kinerja tes diagnostik, namun analisis ROC juga dapat dimanfaatkan dalam mengevaluasi validitas diagnostik suatu alat tes yang biasa digunakan dalam psikologi klinis (Pintea & Moldovan, 2009). Selain itu, analisis ROC juga bisa digunakan untuk memilih yang skor *cut-off* yang optimal (Zou, O'Malley, & Mauri, 2007).

Simpulan

Berdasarkan hasil analisis, dapat disimpulkan bahwa alat tes *Standard Progressive Matrices* ini reliabel menurut metode koefisien Alpha dan valid dengan metode validitas kriteria dengan mengkorelasikan dengan alat tes *Cattell's Culture Fair Intelligence Test*. Beberapa kekuatan dari penelitian ini di antaranya adalah melibatkan jumlah partisipan yang cukup besar dan adanya penambahan data yang relatif baru (tahun 2014 - 2016), meski masih terbatas pada jangkauan usia partisipan. Dengan demikian, untuk penelitian selanjutnya, dapat menyasar ke kelompok usia karyawan dan lebih tua serta bisa dilakukan investigasi kegunaan diagnostik dari SPM dengan kelompok khusus, seperti kelompok individu cerdas istimewa dan kelompok individu dengan disabilitas inteligensi dengan menggunakan analisis *The Receiver-Operating Characteristic* atau ROC.

Penghargaan

LPT UI telah memberikan akses pada data SPM. Kemudian Unika Atma Jaya dan *Bright Consulting* yang telah menyediakan tempat dan partisipan untuk berpartisipasi dalam penelitian ini. Peneliti juga mengucapkan terima kasih kepada para asisten penelitian.

Daftar Pustaka

- Balboni, G., Naglieri, J. A., & Cubelli, R. (2010). Concurrent and Predictive Validity of the Raven Progressive Matrices and the Naglieri Nonverbal Ability Test. *Journal of psychoeducational assessment, 28*(3), 222-235.
- Colom, R., & Abad, F. (2007). Culture Fair Intelligence Test. In N. J. Salkind (Ed.), *Encyclopedia of Measurement and Statistics*. Sage Publications, Inc. <http://dx.doi.org/10.4135/9781412952644.n116>
- Flynn, J. R. (1999). Searching for justice: the discovery of IQ gains over time. *American psychologist, 54*(1), 5 - 20.
- Gravetter, F. J., & Wallnau, L. B. (2005). *Essentials of statistics for the behavioral sciences*. Belmont, CA: Thomson/Wadsworth.
- Heron, A., & Chown, S. (1967). *Age and function*. London: Churchill.
- Jensen, A. R. (1980). *Bias in mental testing*. New York: The Free Press.
- Kaplan, R. M., & Saccuzzo, D. P. (2013). *Psychological testing: Principles, applications, and issues*. (8th ed.). Belmont, CA: Thomson Wadsworth.
- Kumolohadi, R. & Suseno, M. N. (2012). Intelligenz struktur test dan standard progressive matrices: Dari konsep inteligensi yang berbeda menghasilkan tingkat inteligensi yang sama. *Jurnal Inovasi dan Kewirausahaan, 1*(2), 79 - 85. Diunduh 24 Februari 2017 dari <http://bit.ly/2lfwBS5>
- Kyle, J.G. (1977). Raven's Progressive Matrices: Thirty years later. *British Psychological Society Bulletin, 30*, 406-407.

- Lembaga Pengembangan Sarana Pengukuran dan Pendidikan Psikologi Fakultas Psikologi Universitas Indonesia (LPSP3 UI). (2009). *Buku Petunjuk Penggunaan: CFIT Culture Fair Intelligence Test Skala 3 Bentuk A dan Bentuk B*. Depok: LSP3 FPUI.
- Mangestuti, R. & Aziz, R. (2007). Validasi tes inteligensi SPM dan IST pada mahasiswa fakultas psikologi UIN Malang. *Ringkasan Laporan Penelitian*. Diunduh 24 Februari 2017 dari <http://bit.ly/2lfMqYC>
- Moran, A. (1986). The reliability and validity of Raven's Standard Progressive Matrices for Irish apprentices. *Applied psychology*, 35(4), 533-538.
- Nenty, H. J., & Dinero, T. E. (1981). A cross-cultural analysis of the fairness of the Cattell Culture Fair Intelligence Test using the Rasch model. *Applied Psychological Measurement*, 5(3), 355-368.
- Pind, J., Gunnarsdóttir, E. K., & Jóhannesson, H. S. (2003). Raven's Standard Progressive Matrices: new school age norms and a study of the test's validity. *Personality and Individual Differences*, 34(3), 375-386.
- Pintea, S., & Moldovan, R. (2009). The Receiver-Operating Characteristic (ROC) analysis: Fundamentals and applications in clinical psychology. *Journal of Cognitive and Behavioral Psychotherapies*, 9(1), 49-66.
- Raven, J. (1983). The Progressive Matrices and Mill Hill Vocabulary Scale in Western Societies. In S. H. Irvine & J. W. Berry (Eds.), *Human Assessment and Cultural Factors*. New York: Plenum Press.
- Raven, J. (1988). *Manual for Raven's progressive matrices and vocabulary scales. Section 3: The Standard Progressive Matrices*. Oxford: Oxford Psychologist Press.
- Raven, J. (2000). The Raven's progressive matrices: change and stability over culture and time. *Cognitive Psychology*, 41(1), 1-48. doi:10.1006/cogp.1999.0735
- Raven, J., Raven, J. C., & Court, J. H. (1998). *Manual for Raven's progressive matrices and vocabulary scales*. Oxford: Oxford Psychologists Press.
- Rindermann, H., & Nijenhuis, J. (2012). Intelligence in Bali - A case study on estimating mean IQ for a population using a various correction based on theory and empirical findings. *Intelligence*, 40, 395 - 400.
- Rushton, J. P., Skuy, M., & Bons, T. (2004). Construct Validity of Raven's Advanced Progressive Matrices for African and Non-African Engineering Students in South Africa. *International journal of selection and assessment*, 12(3), 220-229.
- Sukarti (1980). *The Coloured Progressive Matrices pada anak anak Sekolah Dasar di desa*. Yogyakarta: Departemen Psikometri Fakultas Psikologi UGM.
- Suwartono, C., Hendriks, M. P. H., Hidajat, L. L., Halim, M. S., & Kessels, R. P. C. (2016). External Validity of the Indonesian Wechsler Adult Intelligence Scale – Fourth Edition (WAIS-IV-ID). *Naskah dalam proses publikasi*.
- Zou, K. H., O'Malley, J., & Mauri, L. (2007). Receiver-operating characteristic analysis for evaluating diagnostic tests and predictive models. *Circulation*, 115, 654-657.