

Validasi skala penilaian instrumen perencanaan karier menggunakan Andrich Threshold

Stefanus Soejanto Sandjaja

Fakultas Psikologi, Universitas Kristen Krida Wacana, Jl. Tanjung Duren Raya No. 4 Jakarta Barat

Yuda Syahputra

Fakultas Ilmu Pendidikan dan Pengetahuan Sosial, Universitas Indraprasta PGRI, Jl. Nagka Raya No. 58 C
Jakarta Selatan

Lira Erwinda

Fakultas Ilmu Pendidikan, Universitas Negeri Padang, Jl. Prof. Dr. Hamka Air Tawar Padang, Sumatera Barat
E-mail: sandjaja@ukrida.ac.id

Abstract

Success in a career is related with subjective wellbeing, so that career development needs particular intervention from the counselor. Process of determining and preparing career made counsellor a valid and reliable inventory. However, inventory that is considered valid and reliable by the counselor still raises problems in administration, such as students confuse to decide their points in rating scale. The purpose of this study is to find a clear and unambiguous rating scale to make it easier for students to set a response point in their inventory. The research sample uses area random sampling consisting of six groups of test subjects, namely: group 1, n = 75; group 2, n = 61; group 3, n = 47; group 4, n = 146; group 5, n = 85; and group 6, n = 63. Data in this study were 5-point Likert scale political data collected using career planning inventory. The research data were analyzed using the Rasch model by testing the rating scale analysis through Threshold analysis between ratings. The results showed the rating scale in the Threshold analysis, the rating scale changed to a four-point Likert scale with a choice of very inappropriate, not appropriate, appropriate, and very appropriate.

Keywords: Career Planning; Likert Scale; Rating Scale; Threshold

Abstrak

Sukses dalam karier identik dengan kesejahteraan hidup individu, yang perlu penanganan lebih lanjut dari konselor. Dalam menentukan dan mempersiapkan karier dimasa depan, konselor mestinya menggunakan inventori yang valid dan reliabel. Namun, inventori yang dianggap valid dan reliabel oleh konselor masih memunculkan permasalahan dalam pengadministrasian, yaitu siswa bingung dalam menetapkan rating scale pada inventori. Tujuan penelitian ini adalah menemukan skala penilaian yang jelas dan tidak ambigu untuk memudahkan siswa menetapkan rating scale pada inventori. Sampel penelitian menggunakan area random sampling yang terdiri dari enam kelompok subjek tes, yaitu: kelompok 1, n = 75; kelompok 2, n = 61; kelompok 3, n = 47; kelompok 4, n = 146; kelompok 5, n = 85; dan kelompok 6, n = 63. Data dalam penelitian ini berupa data politomi 5-point Likert scale yang dikumpulkan menggunakan career planning inventory. Data penelitian dianalisis menggunakan model Rasch dengan menguji rating scale analysis melalui analisis Threshold antar rating. Hasil penelitian menunjukkan rating scale pada analisis Threshold, rating scale berubah menjadi empat point Likert scale dengan pilihan sangat tidak sesuai, kurang sesuai, sesuai, dan sangat sesuai.

Kata kunci: Perencanaan Karier; Rating Scale; Skala Likert; Threshold

Pendahuluan

Data Badan Pusat Statistik (BPS, 2017) menunjukkan bahwa pengangguran di Indonesia masih berada pada kategori tinggi, dengan tingkat pendidikan yang beraneka ragam seperti: pendidikan sekolah menengah kejuruan sebesar 11.41% dan sekolah menengah atas sebesar 8.29%. Kondisi ini disebabkan karena individu kesulitan menghadapi dunia kerja, yang bermula karena di sekolah tidak adanya persiapan karier siswa dalam mengambil keputusan (Afdal, Alizamar, Ildil, Syahputra, & Nurhastuti, 2019), untuk itu perlunya Kepedulian tentang masa depan untuk membantu individu melihat ke depan dan mempersiapkan apa yang akan terjadi selanjutnya dalam mencapai kesuksesan karir (Haenggli & Hirschi, 2020). Jung & Yoo (2020) menyatakan bahwa rencana karir individu harus dikerjakan secara konsisten selama sekolah menengah, sekolah menengah, dan perguruan tinggi berdasarkan bakat dan minat. Berdasarkan kondisi tersebut pentingnya perencanaan karier siswa dalam menentukan pilihan jurusan di sekolah sampai menentukan jurusan di perguruan tinggi, sehingga tidak ada lagi di masa depan masalah kesalahan dalam memilih jurusan (Afdal et al., 2019). Namun, kenyataannya individu hampir tidak pernah memiliki informasi karier yang jelas dan tegas. Seringkali informasi tersebut ambigu atau tidak tersedia pada saat pengambilan keputusan karier (Xu & Tracey, 2015).

Kondisi ini mengakibatkan para peneliti membuat alat ukur untuk perencanaan karier atau kemampuan beradaptasi karier siswa (Hirschi, Freund, & Herrmann, 2014; Pignault & Houssemand, 2013; Porfeli & Savickas, 2012; Savickas & Porfeli, 2011; Stambulova, 2010; Wegemer & Eccles, 2019). Selain membuat alat ukur sendiri, para peneliti di berbagai belahan dunia mengadopsi instrumen holland yang telah disesuaikan dengan berbagai Bahasa di dunia dan disesuaikan dengan budaya (Bullock, Andrews, Braud, & Reardon, 2009; Gasser, Larson, & Borgen, 2007). Penggunaan inventori minat jabatan RIASEC dikembangkan terutama dipopulasi AS, dan diterapkan untuk individu yang berbeda secara budaya di seluruh dunia (Fan & Leong, 2016[Cina]; Iliescu, Ispas, Ilie, & Ion, 2013[Romania]; Morgan & de Bruin, 2017[Afrika]; Nagy, Trautwein, & Lüdtke, 2010[Jerman]).

Penggunaan alat ukur beradaptasi karier untuk memprediksi hasil karier yang penting, termasuk kepuasan karier (Zacher, 2014), gaji (Guan, Zhou, Ye, Jiang, & Zhou, 2015) dan prestasi kerja (Yu & Zheng, 2013). Diantaranya inventori yang sangat populer digunakan adalah *Career Adapt-Abilities Scale* (CAAS) adalah inventori untuk mengukur konstrukpsikosial dalam kemampuan karier di berbagai Negara (Savickas & Porfeli, 2012). Teori konstruksi karier Savickas (2013) mengemukakan bahwa kemampuan beradaptasi karier, yang merujuk pada sumber daya psikologis yang penting (kekhawatiran karier, kontrol karier, keingintahuan karier, kepercayaan karier) yang memfasilitasi pemecahan masalah individu dalam pengembangan karier mereka (Savickas, 1997). Berdasarkan hasil dari 13 negara menunjukkan inventori CAAS valid dan reliabel dalam pengukuran lintas budaya (Savickas & Porfeli, 2012).



Selanjutnya, para peneliti tidak hanya membahas tentang alat ukur perencanaan karier siswa. Namun, juga membahas tentang perencanaan karier para atlet setelah berhenti (Sorkkila, Ryba, Aunola, Selänne, & Salmela-Aro, 2017). Stambulova (2010) mengemukakan *Five-Step Career Planning Strategy (5-SCP)* yang merupakan kerangka kerja konseling untuk membantu atlet menuju masa berkarier setelah tidak lagi menjadi atlet atau setelah menjadi atlet. Empat langkah pertama dalam strategi ini berurusan dengan memetakan pengalaman masa lalu klien, situasi saat ini, masa depan yang dirasakan dan mengintegrasikan pengalaman (masa lalu, masa kini, dan masa depan) yang diproyeksikan ke dalam strategi karier dan kehidupan (Fernandez, Stephan, & Fouquereau, 2006).

Alat ukur yang valid dan reliabel pasti memiliki skala rating yang baik pula dalam menentukan *rating scale*. *Rating scale* merupakan bagian dari kriteria evaluatif dan kategori penilaian yang menunjukkan sejauh mana perilaku, keterampilan atau strategi tertentu yang ditampilkan oleh seorang individu (Vagle, 2014). Isi *rating scale* mewakili pilihan yang dibuat selama desainnya tentang apa yang dinilai dan bagaimana penilaiannya (Stiggins, Arter, Chappuis, & Chappuis, 2006). Jika *rating scale* tidak masuk akal (kriteria penilaian terlalu banyak atau terlalu sedikit atau tidak relevan), atau tidak dapat diimplementasikan dengan andal, maka interpretasi dan penggunaan skor yang berasal dari instrumen akan terjadi kekeliruan dalam membuat kesimpulan (Tierney & Simon, 2004).

Berdasarkan pemaparan di atas, dapat dinyatakan bahwa para peneliti hanya berfokus pada validasi inventori dan strategi untuk meningkatkan perencanaan karier, dan melupakan apakah individu yang mengisi inventori penelitian benar-benar tidak ambigu dalam mengisi jawaban pertanyaan inventori. Tujuan pada penelitian ini adalah untuk mevalidasi skala penilaian inventori perencanaan karier yang jelas atau tidak ambigu untuk memudahkan siswa menentukan pilihan jawaban pada inventori. Untuk itu peneliti mengembangkan Inventori Perencanaan Karier Siswa yang dikembangkan dengan 4 tahapan perencanaan karier, yakni Pemahaman diri, pemahaman lingkungan, perumusan pilihan dan perencanaan tindakan (Afdal, 2015; Afdal et al., 2019). Inventori ini dikembangkan melalui 5 *rating scale*, yakni sangat tidak sesuai, tidak sesuai, kurang sesuai, sesuai, dan sangat sesuai, agar inventori ini lebih valid dan reliabel maka perlu penelitian lanjutan berkenaan dengan validitas dan reliabilitas pengembangan inventori perencanaan karier siswa.

Metode

Sampel penelitian berjumlah 477 siswa dari SMA, SMK, MAN dengan rentang usia 16-18 tahun yang tersebar pada empat kabupaten (Agam, Dharmasraya, Pasaman, dan Pesisir Selatan) dan dua kota (Bukittinggi dan Payakumbuh) di Sumatera Barat, Indonesia. Alasan memilih sampel penelitian dari berbagai sekolah adalah bertitik tolak dari tujuan penelitian, yaitu untuk mevalidasi skala penilaian atau pilihan jawaban

inventori perencanaan karier yang mudah dipahami siswa di berbagai sekolah untuk itu perlunya sampel yang *representative* dari berbagai daerah di Sumatera Barat. Teknik pengambilan sampel dengan teknik *area random sampling*, *penyebaran instrument* dibantu mahasiswa pascasarjana dan guru bimbingan dan konseling yang tergabung di Majelis Guru Bimbingan Konseling (MGBK) disetiap daerah. Seluruh sampel penelitian terdiri dari enam kelompok subjek Tes di lokasi yang berbeda, yaitu: subjek Tes 1, n = 75 dari Kabupaten Agam; subjek Tes 2, n = 61 dari Kota Bukittinggi; subjek Tes 3, n = 47 dari Kabupaten Dharmasraya; subjek Tes 4, n = 146 dari Kabupaten Pasaman; subjek Tes 5, n = 85 dari Kota Payakumbuh; dan subjek Tes 6, n = 63 dari Kabupaten Pesisir Selatan.

Data kelompok subjek tes dikumpulkan melalui *career planning inventory* yang dikembangkan oleh Afdal (2015) dengan mengukur empat aspek, yaitu: 1) pemahaman diri, salah satu item berbunyi “Saya mengetahui bakat yang dimiliki”, 2) pemahaman lingkungan, salah satu item berbunyi “Saya memahami bahwasanya terdapat persyaratan tertentu untuk memasuki suatu pekerjaan/karier tertentu”, 3) perumusan pilihan, salah satu item berbunyi “Rencana pilihan karier yang akan digeluti sesuai dengan bakat yang dimiliki”, dan 4) perencanaan tindakan, salah satu item berbunyi “Saya memahami bahwa perencanaan yang baik akan mempermudah saya untuk memasuki suatu pekerjaan/karier tertentu”. Kuisioner menggunakan 5-point Likert scale dengan jumlah item 63. Data penelitian dianalisis menggunakan model Rasch dengan menguji *rating scale analysis* melalui analisis Threshold antar rating (Alagumalai, Curtis, & Hungi, 2005; Bond & Fox, 2015; Marsinun, Erwinda, & Syahputra, 2020; Sumintono & Widhiarso, 2015). Berdasarkan hasil analisis threshold akan dipertimbangkan skala rating yang digunakan dalam inventori, sehingga seluruh siswa tidak bingung dalam menetapkan pilihan jawaban pada inventori. Sebelum menguji validasi rating scale, berikut disampaikan hasil uji validasi instrument menggunakan rasch model.

Tabel 1

Hasil Uji Validasi Instrumen

Estimation	Values
Item Reliabilities	.87
Person Reliabilities	.95
Cronbach alpha (KR-20)	.96
Separation index of Person	4.17
Mean Item	0.00
Mean Person	0.60
Mean INTFIT MNSQ Item	1.01
Mean INFIT MNSQ Person	1.00
Mean OUTFIT MNSQ Item	1.00
Mean OUTFIT MNSQ Person	1.00
Raw Variance Explained by measures	33.0 %
Raw Variance Unexplained by measures	67.0 %



Hasil

Untuk mencapai tujuan pada penelitian ini ada dua tahapan penilaian yang akan dilakukan yaitu: (1) menilai kualitas *career planning inventory* di Sumatera Barat, (2) menilai validitas skala peringkat yang diberikan tes *career planning inventory*, yang dijelaskan sebagai berikut:

Kualitas Career Planning Inventory

Inventory dikatakan baik bila mengukur apa yang seharusnya diukur, pada Tabel 2 disampaikan kualitas measure dari *Career planning inventory*.

Tabel 2
Hasil Uji Menggunakan *Summary Statistic*

Measure	Pelaksanaan					
	Kelompok Subjek Test 1	Kelompok Subjek Test 2	Kelompok Subjek Test 3	Kelompok Subjek Test 4	Kelompok Subjek Test 5	Kelompok Subjek Test 6
Mean Person	.60	1.61	1.35	1.15	1.32	1.62
Mean Item	.00	.00	.00	.00	.00	.00
Infit MNSQ Item	1.01	.99	1.00	1.00	1.00	1.01
Outfit MNSQ Item	1.00	1.01	1.01	1.01	1.05	1.00
Separation Item	2.57	2.21	1.43	3.85	1.52	1.47
Separation Person	4.17	4.46	4.02	3.94	3.73	4.25
Raw variance explained by measures	33.0 %	34.7 %	29.7 %	29.3 %	22.2 %	31.4 %
Reliability Person	.95	.95	.94	.94	.93	.95
Reliability Item	.87	.83	.67	.94	.70	.68
Cronbach Alpha (Kr-20)	.96	.96	.95	.95	.93	.96

Berdasarkan Tabel 2 di atas dapat dijelaskan bahwa kemampuan siswa memberikan jawaban terhadap inventori sangat baik, terlihat dari perbandingan nilai *mean person* dan item. Nilai *mean item* menjadi titik normal yang dimulai dari nilai 0.00. Sedangkan, nilai *mean person* menjadi tolak ukur seberapa tinggi kemampuan responden dalam memberikan jawaban terhadap inventori, hal ini terlihat dari nilai yang selalu bervariasi setiap dilakukan di tempat atau daerah yang berbeda. Sebagai contoh kelompok subjek tes 2 s/d tes 6 nilai *mean person* cenderung di atas (1.00), hal ini berarti kecenderungan person menjawab item tergolong mudah, berbeda dengan nilai *mean person* pada kelompok subjek tes 1 tidak jauh berbeda dengan nilai *mean item*, hal ini berarti ada kesamaan kualitas item dengan kualitas individu.

Selanjutnya, untuk melihat reliabilitas pada *Career Planning Inventory* dapat dilihat skor masing-masing tes *reliability person* pada kelompok subjek tes 1 s/d tes 6 cenderung tinggi atau berada di atas (.93), hal ini dapat dinyatakan bahwa keragaman person pada setiap kelompok tes dan kualitas jawaban yang diberikan person sangat bagus. Sama halnya dengan nilai *reliability item* pada kelompok subjek tes 4 menunjukkan kualitas item dalam kondisi sangat bagus (.94). Sedangkan, pada kelompok subjek tes 1 dan 2, menunjukkan kualitas item dalam pengukuran perencanaan karier adalah bagus. Namun,

hal yang berbeda ditunjukkan pada kelompok subjek tes 3, 5, dan 6, menunjukkan kualitas item dalam pengukuran (0.67 - 0.80) adalah cukup. Sementara nilai masing-masing tes pada alpha Cronbach (KR-20) adalah kelompok subjek tes 1 (0.96), tes 2 (.96), tes 3 (.95), tes 4 (.95), tes 5 (.93), dan tes 6 (.96), hal ini menunjukkan bahwa interaksi antara person dan item bagus sekali pada seluruh kelompok subjek tes 1, 2, 3, 4, 5, dan 6.

Disamping itu nilai sensitifitas item pada masing-masing kelompok subjek tes cenderung mendekati +1.00 logit (INFIT MNSQ) dan nilai sensitifitas item secara keseluruhan pada masing-masing tes mendekati +1.00 logit (OUTFIT MNSQ). Hal ini menunjukkan kesesuaian item dalam pengukuran perencanaan karier siswa pada kelompok subjek tes 1 s/d 6 telah sesuai, hal ini berarti menunjukkan apakah item fit sesuai dengan model Rasch. Kesesuaian dalam pengukuran dilihat dari *infit* dan *outfit mean square* dengan nilai kuadrat tengah 1.0 (Alagumalai et al., 2005; Bond & Fox, 2015; Sumintono & Widhiarso, 2015). Lebih lanjut, nilai *raw variance explained by measures* pada masing-masing tes berada di atas 20 %, hal ini menunjukkan bahwa persyaratan minimum unidimensional 20% telah terpenuhi (Linacre, 2011; Sumintono & Widhiarso, 2015). Dengan demikian kondisi ini menyatakan bahwa persyaratan unidimensionalitas inventori terpenuhi. Sehingga dapat dinyatakan 63 item yang digunakan dalam *Career planning inventory* adalah valid dan baik untuk kondisi pengukuran di daerah kelompok subjek tes. Hasil ini akan relatif sama pada pengukuran di Sumatera Barat, Indonesia.

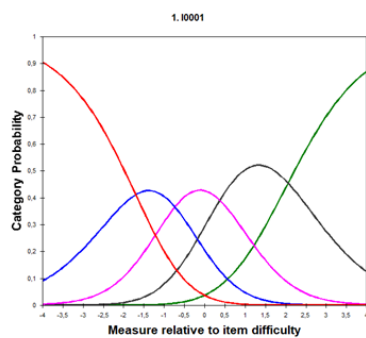
Skala Peringkat yang diberikan Career planning inventory dapat dipahami oleh Responden

Inventori yang baik digunakan adalah inventori yang tidak membingungkan responden dalam memilih jawaban yang disediakan oleh inventori itu sendiri. Skala peringkat yang diberikan harus dipahami dengan baik oleh responden. Parameter Threshold antar rating yang digunakan untuk menunjukkan standar adalah 1.4 – 5.0 logit (Sumintono & Widhiarso, 2015). Inventori menggunakan *5-point Likert scale*. Analisis Threshold antar rating yang disampaikan pada Tabel 3 berikut.

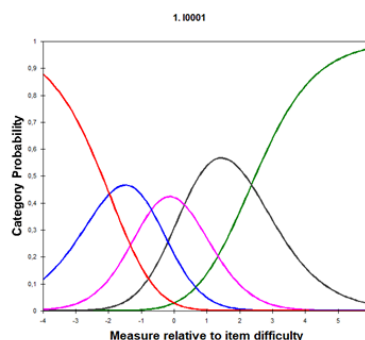
Tabel 3. Hasil Uji Menggunakan *Category Structure*

Kategori	Label	Kelompok Subjek Test 1	Kelompok Subjek Test 2	Kelompok Subjek Test 3	Kelompok Subjek Test 4	Kelompok Subjek Test 5	Kelompok Subjek Test 6
1	Sangat tidak sesuai	None	None	None	None	None	None
2	Tidak sesuai	-1.70	-1.98	-1.94	-2.05	-1.74	-1.72
3	Kurang sesuai	-.73	-.69	-1.70	-1.03	-.72	-.78
4	Sesuai	.37	.31	.56	.50	.35	.18
5	Sangat sesuai	2.06	2.36	3.08	2.58	2.11	2.32

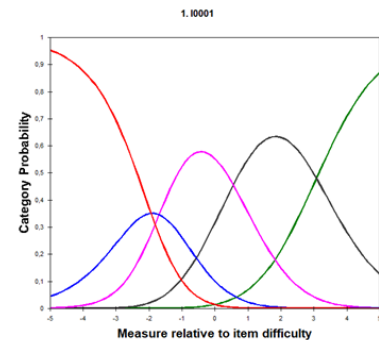
Berdasarkan Tabel 3 di atas, menunjukkan jawaban responden terhadap *rating scale* yang diberikan pada *Career planning inventory* pada masing-masing kelompok subjek tes. Hasil di atas menunjukkan kestabilan pilihan jawaban yang dapat di respon oleh responden, hal ini dinyatakan dari nilai Threshold setiap pilihan 1 sampai 5 tidak jauh berbeda pada masing-masing kelompok subjek tes 1, 2, 5, dan 6. Pada *rating scale* 1 atau sangat tidak sesuai menjadi titik awal sehingga dikatakan none. Hal ini menunjukkan nilai Threshold pada pilihan 2 = tidak sesuai dan 3 = kurang sesuai, tidak memenuhi standar antar rating (1.4 – 5.0) atau kurang dari 1.4 setiap antar rating (Engelhard Jr & Wind, 2017). Sementara pada kelompok subjek tes 3 dan 4 menunjukkan nilai Threshold pada pilihan 2 = tidak sesuai, tidak memenuhi standar antar rating. Kesulitan dalam memberikan atau menetapkan jawaban diperjelas dengan menggunakan gambar sebagai berikut.



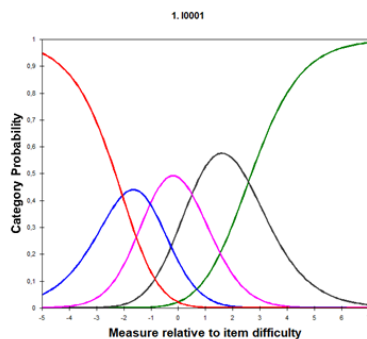
Gambar 1. Kelompok Subjek Tes 1



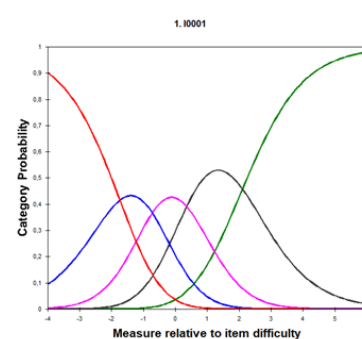
Gambar 2. Kelompok Subjek Tes 2



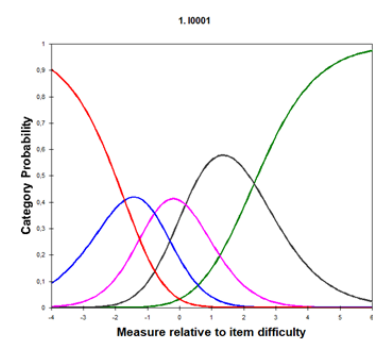
Gambar 3. Kelompok Subjek Tes 3



Gambar 4. Kelompok Subjek Tes 4



Gambar 5. Kelompok Subjek Tes 5



Gambar 6. Kelompok Subjek Tes 6

Pada gambar di atas, menunjukkan garis-garis menjadi *rating scale*, garis merah = pilihan 1, garis biru = pilihan 2, garis merah muda = 3, garis hitam = 4, dan garis hijau = 5, terlihat *rating scale* yang diberikan pada seluruh kelompok subjek tes 1, 2, 4, 5, dan 6 cenderung memilih jawaban 1 = sangat tidak sesuai, 4 =sesuai, dan 5 = sangat sesuai. Hal ini terlihat dari seluruh gambar yang menunjukkan garis merah, hitam, dan hijau muncul ke atas atau bergerak ke arah atas. Sedangkan, garis biru dan merah muda cenderung di bawah garis lainnya. Kondisi ini disebabkan karena siswa kebingungan atau kesulitan memilih pilihan jawaban 2 = tidak sesuai dengan pilihan jawaban 3 = kurang sesuai, kedua pilihan jawaban tersebut sulit dimengerti oleh siswa. Sementara, siswa cenderung memberikan jawaban 1 = sangat tidak sesuai, 4 = sesuai, dan 5 = sangat sesuai.

Sementara pada kelompok subjek tes 3, sangat terlihat jelas bahwa garis biru tertimpah dengan garis lainnya, kondisi ini dapat dinyatakan responden sangat kesulitan menentukan pilihan jawaban tidak sesuai. Dengan kata lain, pilihan jawaban 2 = tidak sesuai pada kelompok subjek tes 3 dapat dihilangkan. Apabila garis yang menandakan *rating scale* pada inventori tidak dihilangkan atau digabungkan, maka akan berdampak pada bias sosial dalam memberikan *rating scale* jawaban. Bias sosial yang dimaksud adalah keinginan untuk membuat perasaan interviewer menjadi senang karena mau menjadi responden. Responden akan cenderung menjawab dengan memilih netral (titik di tengah; Garland, 1991). Akan tetapi, jumlah titik respon 2, 3, dan 4 memiliki kelemahan dalam hal stabilitas dan kurang merepresentasikan jawaban responden (Hofmans, Joeri, Theuns, & Mairesse, 2007).

Pembahasan

Skala penilaian adalah serangkaian deskripsi yang memiliki skor dari karakteristik kinerja yang dilakukan peserta tes (Luoma, 2004; Pan, 2016). Dengan demikian, mereka "membentuk bagian dari definisi mereka tentang konstruk yang dinilai dalam tes" (Luoma, 2004). Dalam skala penilaian memiliki dua komponen, yaitu: skala peringkat mencakup domain yang akan dinilai (konstruk) dan penyelarasan antara kinerja peserta ujian atau tingkat deskripsi perilaku yang telah ditentukan sebelumnya. Oleh karena itu, sejauh menyangkut penilaian bahasa, setiap pengembangan skala peringkat juga harus menanggung pertimbangan dari dua komponen di atas (Pan, 2016).

Alat ukur yang valid dan reliabel pasti memiliki skala *rating* yang baik pula dalam menentukan *rating scale* agar mudah siswa memahami pilihan jawaban yang disediakan oleh alat ukur tersebut. *Rating scale* merupakan bagian dari kriteria evaluatif dan kategori penilaian yang menunjukkan sejauh mana perilaku, keterampilan atau strategi tertentu yang ditampilkan oleh seorang individu (Vagle, 2014). Isi *rating scale* mewakili pilihan yang dibuat selama desainnya tentang apa yang dinilai dan bagaimana penilaiannya (Stiggins et al., 2006). Jika *rating scale* tidak masuk akal (kriteria penilaian terlalu banyak atau terlalu sedikit atau tidak relevan), atau tidak dapat diimplementasikan dengan andal, maka interpretasi dan penggunaan skor yang berasal dari instrument akan terjadi kekeliruan dalam membuat kesimpulan (Tierney & Simon, 2004).

Setiap pengukuran memiliki skala penilaian yang baik untuk memperkecil kekeliruan dalam membuat kesimpulan, alat ukur yang sering digunakan adalah *Career Adapt-Abilities Scale* (CAAS) dengan menggunakan skala penilaian model *Likert-type* 5 poin: 1 (tidak kuat), 2 (agak kuat), 3 (kuat), 4 (sangat kuat), dan 5 (terkuat). Pada skala penilaian 5 poin yang digunakan pada Yordania valid atau mudah dipahami oleh responden untuk memberi respon jawaban (Mazahreh, Stoltz, & Wolff, 2017). Lalu, tanggapan terhadap skala kepuasan kerja dinilai sangat tinggi, dengan 62% dari peserta menunjukkan 4 atau lebih tinggi pada skala 5 poin (Mazahreh et al., 2017). Selanjutnya,



penelitian Tsai, Hsu, & Hsu (2017) juga mengembangkan skala kecemasan karier untuk mahasiswa pariwisata untuk masuk ke dunia kerja, dalam skala kecemasan karier menggunakan skala Likert 4 poin (1 = sangat tidak setuju dan 4 = sangat setuju); skor yang tinggi menunjukkan tingkat kecemasan yang tinggi.

Berdasarkan hasil penelitian menunjukkan kecenderungan kelompok subjek test 1, 2, 3, 4, 5, dan 6 pada tabel 2 menunjukkan *rating scale* 1 = sangat tidak sesuai, 4 = sesuai, dan 5 = sangat sesuai dapat dipahami responden. Namun, *rating scale* tidak sesuai dan kurang sesuai perlu dipertimbangkan untuk digunakan dalam *Career planning inventory*. Hal ini disebabkan responden kesulitan dalam memberikan atau menetapkan *rating scale* tidak sesuai dan kurang sesuai. Artinya kesulitan item tinggi menggambarkan responden mengalami kesulitan menyetujui item karena konten item kuat, yang berarti item menggambarkan konten dengan cara ekstrem yang membuat penilai sangat setuju atau tidak setuju dengan item tersebut (I. J. Park & Jung, 2015). Keunggulan yang diberikan model peringkat Rasch adalah memberikan indeks kesulitan item yang mewakili tingkat kesulitan person terkait dengan kesepakatan dengan suatu item (I.-J. Park, Han, & Ryu, 2019).

Dari hasil temuan pada penelitian ini, dapat dianalisis bahwa masyarakat yang tinggal di Sumatera Barat kesulitan dengan skala penilaian yang terlalu banyak. Kontradiksi dengan masyarakat Cina yang dapat membedakan *rating scale* sampai 7 point dengan menggunakan skala likert pada pengukuran *Career Personality Styles Inventory* (CPSI), yang secara khusus berkontribusi pada penerapan teori Holland kepribadian karier di antara populasi Cina (Fu & Zhang, 2019). Senadah dengan Hofmans, Joeri, Theuns, & Mairesse (2007) menyatakan jumlah titik respon 7 memiliki stabilitas, indeks validitas, reliabilitas, dan kekuatan diskriminasi yang baik serta jumlah titik respon ini lebih disukai oleh responden. Berdasarkan analisis Threshold antar rating perlunya penggabungan antara pilihan jawaban tidak sesuai dan kurang sesuai menjadi satu pilihan jawaban kurang sesuai. Sehingga nilai Threshold menjadi 1.4, karena parameter Threshold antar rating yang digunakan untuk menunjukkan standar adalah 1.4 – 5.0 logit (Sumintono & Widhiarso, 2015). Namun, perlunya pertimbangan stabilitas item, karena jumlah titik respon 2, 3, dan 4 memiliki kelemahan merepresentasikan jawaban responden (Hofmans et al., 2007).

Berdasarkan hasil tersebut dapat dinyatakan bahwa *rating scale* berubah menjadi 4-point *Likert scale* dengan pilihan sangat tidak sesuai, kurang sesuai, sesuai, dan sangat sesuai. Sama halnya dengan Garland (1991) menjelaskan bahwa bias sosial menjadi pertimbangan saat menentukan jumlah *rating scale*. Jumlah *rating scale* genap (jumlah titik 4) lebih disarankan (Becker, 2018) daripada jumlah *rating scale* ganjil (jumlah titik 5) karena bias sosial dapat dikurangi. Syahputra, Sandjaja, Afdal, & Ardi (2019) mengungkapkan bahwa empat *rating scale* pada instrumen lebih dimengerti daripada Lima, karena lima *rating scale* membuat siswa cenderung memilih titik respon tengah yang merupakan titik aman.

Kesimpulan

Hasil temuan menunjukkan *rating scale* pada instrumen berubah menjadi 4-point *Likert scale* dengan pilihan sangat tidak sesuai, kurang sesuai, sesuai, dan sangat sesuai. Hal ini disebabkan dari penggabungan antara *rating scale* tidak sesuai dan kurang sesuai. Sehingga nilai Threshold menjadi 1.4, karena parameter Threshold antar rating yang digunakan untuk menunjukkan standar adalah 1.4 – 5.0 logit (Sumintono & Widhiarso, 2015). Hal ini dapat dinyatakan bahwa *Career planning inventory* dapat digunakan di Sumatera Barat dengan skala 4-point likert.

Hasil temuan mengungkapkan bahwa masyarakat Sumatera Barat terindikasi dalam memberikan respon jawaban bingung terhadap *rating scale* yang banyak (>5) cenderung menggunakan skala likert dengan 4-point dan maksimal 5-point. Kondisi ini disebabkan karena siswa kebingungan dengan pilihan tidak sesuai dengan pilihan kurang sesuai, kedua pilihan jawaban tersebut sulit dimengerti oleh siswa. Sementara, siswa cenderung memberikan pilihan sangat tidak sesuai, sesuai, dan sangat sesuai. Berdasarkan hasil penelitian diharapkan para peneliti lebih memperhatikan *rating scale* dalam mengembangkan instrumen, karena kesalahan menentukan pilihan jawaban akan berdampak pada penarikan kesimpulan dalam penelitian.

Referensi

- Afdal, A. (2015). *Career planning inventory*. Padang: Universitas Negeri Padang.
- Afdal, A., Alizamar, A., Ildil, I., Syahputra, Y., & Nurhastuti, N. (2019). Career planning differences between male. *Specialusis Ugdymas / Special Education*, 1(39), 99–108. <https://doi.org/10.21277/se.v1i39.409>
- Alagumalai, S., Curtis, D. D., & Hungi, N. (2005). *a Pplied Rasch Measurement : a Book of Exemplars Education in the Asia-Pacific Region : Issues , Concerns and Prospects*. Netherlands: Springer.
- Becker, A. (2018). Not to scale? An argument-based inquiry into the validity of an L2 writing rating scale. *Assessing Writing*, 37(2018), 1–12.
- Bond, T. G., & Fox, C. M. (2015). *Applying the Rasch Model, Fundamentals Measurement in the Human Science (3rd edition)*. New York: Routledge.
- BPS. (2017). *Keadaan Ketenagakerjaan Indonesia Agustus 2017*. Indonesia: Badan Pusat Statistik.
- Bullock, E. E., Andrews, L., Braud, J., & Reardon, R. C. (2009). Holland's theory in an international context: Applicability of RIASEC structure and assessments. *Career Planning & Adult Development Journal*, 25(4), 29–58.
- Engelhard Jr, G., & Wind, S. (2017). *Invariant Measurement with Raters and Rating Scales: Rasch Models for Rater-Mediated Assessments*. New York: Routledge Taylor dan Francis Group.
- Fan, W., & Leong, F. L. (2016). Introduction to the special issue: Career development and intervention in Chinese contexts. *Career Development Quarterly*, 64(3), 192–202.
- Fernandez, A., Stephan, Y., & Fouquereau, E. (2006). Assessing reasons for sports career termination: Development of the Athletes' Retirement Decision Inventory (ARDI). *Psychology of Sport and Exercise*, 7(2006), 407–421.



- Fu, M., & Zhang, L. (2019). Developing and Validating the Career Personality Styles Inventory. *Measurement and Evaluation in Counseling and Development*, 52(1), 38–51. <https://doi.org/10.1080/07481756.2018.1435193>
- Garland, R. (1991). The Mid-Point on a Rating Scale: Is it Desirable? *Marketing Bulletin*, 2(1), 66–70.
- Gasser, C. E., Larson, L. M., & Borgen, F. H. (2007). Concurrent validity of the 2005 Strong Interest Inventory: An examination of gender and major field of study. *Journal of Career Assessment*, 15(1), 23–43.
- Guan, Y., Zhou, W., Ye, L., Jiang, P., & Zhou, Y. (2015). Perceived organizational career management and career adaptability as predictors of success and turnover intention among Chinese employees. *Journal of Vocational Behavior*, 88, 230–237.
- Haenggli, M., & Hirschi, A. (2020). Career adaptability and career success in the context of a broader career resources framework. *Journal of Vocational Behavior*, 119, 103414. Retrieved from <https://doi.org/10.1016/j.jvb.2020.103414>
- Hirschi, A., Freund, P. A., & Herrmann, A. (2014). The Career Engagement Scale: Development and validation of a measure of proactive career behaviors. *Journal of Career Assessment*, 22, 575–594.
- Hofmans, Joeri, Theuns, P., & Mairesse, O. (2007). Impact of the Number of Response Categories on Linearity and Sensitivity of Self-Anchoring Scales: A Functional Measurement Approach. *Methodology*, 3(4), 160–169.
- Iliescu, D., Ispas, D., Ilie, A., & Ion, A. (2013). The structure of vocational interests in Romania. *Journal of Counseling Psychology*, 60(2), 294–302.
- Jung, Y.-M., & Yoo, I.-Y. (2020). The effectiveness of a career efficacy enhancement program for KOREAN nursing students: A quasi-experimental study. *Nurse Education Today*, 89, 104423.
- Linacre, J. M. (2011). *A User's Guide to WINSTEPS Ministeps Rasch-Model Computer Programs*. <https://doi.org/ISBN-0-941938-03-4>
- Luoma, S. (2004). *Assessing speaking*. Cambridge: Cambridge University Press.
- Marsinun, R., Erwinda, L., & Syahputra, Y. (2020). Homosexual and Transgender Tendencies in Terms of Gender: A Rasch Perspective. In *International Conference on Progressive Education (ICOPE 2019)* Atlantis Press, 422, 358–361.
- Mazahreh, L. G., Stoltz, K. B., & Wolff, L. A. (2017). Development Validation of the Career Adapt-Abilities Scale in the Hashemite Kingdom of Jordan Validation of the Career Adapt-Abilities Scale in the Hashemite Kingdom of Jordan ABSTRACT. *Measurement and Evaluation in Counseling and Development*, 1756(December). <https://doi.org/10.1080/07481756.2017.1358058>
- Morgan, B., & de Bruin, G. P. (2017). Structural validity of Holland's circumplex model of vocational personality types in Africa. *Journal of Career Assessment*, 25(1), 1–17.
- Nagy, G., Trautwein, U., & Lüdtke, O. (2010). The structure of vocational interests in Germany: Different methodologies, different conclusions. *Journal of Vocational Behavior*, 76(2), 153–169.
- Pan, M. (2016). *Nonverbal Delivery in Speaking Assessment From an Argument to a Rating Scale Formulation and Validation*. China: Springer.
- Park, I.-J., Han, K., & Ryu, K. (2019). Development and Validation of a Career Future Time Perspective Scale. *Journal of Career Development*, 1–14.
- Park, I. J., & Jung, H. (2015). Relationships among future time perspective, career and organizational commitment, occupational self-efficacy, and turnover intention.

-
- Social Behavior and Personality: An International Journal*, 43, 1547–1561.
- Pignault, A., & Houssemand, C. (2013). Work Context and Career Development Construction of a Work Context Inventory. *Procedia - Social and Behavioral Sciences*, 82(2013), 169 – 175.
- Porfeli, E. J., & Savickas, M. . (2012). Career Adapt-Abilities Scale-USA Form: Psychometric properties and relation to vocational identity. *Journal of Vocational Behavior*, 80, 748–753.
- Savickas, M. L. (1997). Career adaptability: An integrative construct for life-span, life-space theory. *The Career Development Quarterly*, 45, 247–259.
- Savickas, M. L. (2013). *Career construction theory and practice*. In R. W. Lent, & S. D. Brown (Eds.). *Career development and counseling: Putting theory and research to work* (pp. 147–183). (2nd ed.). New Jersey: John Wiley & Sons, Inc.
- Savickas, M. L., & Porfeli, E. J. (2011). Revision of the Career Maturity Inventory: The Adaptability Form. *Journal of Career Assessment*, 19, 355–374.
- Savickas, M. L., & Porfeli, E. J. (2012). Career Adapt-Abilities Scale: Construction, reliability, and measurement equivalence across 13 countries. *Journal of Vocational Behavior of Vocational*, 80, 661–673.
- Sorkkila, M., Ryba, T. V., Aunola, K., Selänne, H., & Salmela-Aro, K. (2017). Sport burnout inventory–Dual career form for student-athletes: Assessing validity and reliability in a Finnish sample of adolescent athletes. *Journal of Sport and Health Science*, 10(January), 1–9.
- Stambulova, N. (2010). Counseling athletes in career transitions: The five-step career planning strategy. *Journal of Sport Psychology in Action*, 1(2), 95–105. <https://doi.org/10.1080/21520704.2010.528829>
- Stiggins, R., Arter, J., Chappuis, J., & Chappuis, S. (2006). *Classroom assessment for student learning –Doing it right, using it well*. Portland. OR: Educational Testing Service.
- Sumintono, B., & Widhiarso, W. (2015). *Aplikasi Pemodelan Rasch pada Assessment Pendidikan*. Bandung: Trim Komunikata.
- Syahputra, Y., Sandjaja, S. S., Afdal, A., & Ardi, Z. (2019). Development an inventory of homosexuality and transgender exposure (IHTE): A Rasch analysis. *Konselor*, 8(4), 120–133.
- Tierney, R., & Simon, M. (2004). What’s still wrong with rubrics: focusing on the consistency of performance criteria across scale levels. *Practical Assessment, Research & Evaluation*, 9(2), 1–10.
- Tsai, C. S., Hsu, H., & Hsu, Y. (2017). Tourism and Hospitality College Students ’ Career Anxiety : Scale Development and Validation Tourism and Hospitality College Students ’ Career Anxiety : Scale Development. *Journal of Hospitality & Tourism Education*, 0(0), 1–8. <https://doi.org/10.1080/10963758.2017.1382365>
- Vagle, N. . (2014). *Design in Five: Essential Phases to Create Engaging Assessment Practice*. Citizen: Solution Tree Press.
- Wegemer, C. M., & Eccles, J. S. (2019). Gendered STEM career choices: Altruistic values, beliefs, and identity. *Journal of Vocational Behavior*, 110(2019), 28–42.
- Xu, H., & Tracey, T. J. G. (2015). Career Decision Ambiguity Tolerance Scale: Construction and initial validations. *Journal of Vocational Behavior*, 88, 1–9. <https://doi.org/10.1016/j.jvb.2015.01.006>
- Yu, H., & Zheng, X. (2013). The impact of employee career adaptability: Multilevel



analysis. *Acta Psychologica Sinica*, 45, 680–693.

Zacher, H. (2014). Career adaptability predicts subjective career success above and beyond personality traits and core self-evaluations. *Journal of Vocational Behavior*, 84, 21–30.