

ANALISIS RASCH DALAM *UTRECHT WORK ENGAGEMENT SCALE-9* (UWES-9) VERSI BAHASA INDONESIA

Ika Febrian Kristiana^{1,2}, Fajrianthi², Urip Purwono³

¹Fakultas Psikologi Universitas Diponegoro
Jl. Prof. Soedarto, SH, Kampus Undip Tembalang, Semarang, Indonesia 50275
²Program Doktor Psikologi Universitas Airlangga
Jl. Airlangga 4-6, Kampus B Unair, Surabaya, Indonesia, Indonesia 60286
³Fakultas Psikologi Universitas Padjadjaran
Jl. Raya Bandung Sumedang KM 21, Jatinangor, Indonesia 45363

ika.f.kristiana@gmail.com

Abstract

Using a scale in different cultural contexts requires an adaptation process for the sake of fair comparison between obtained scores. The 9-item version of Utrecht Work Engagement Scale (UWES-9) has been adapted in various countries but Indonesia. This study aimed to make a scale adaptation of the UWES-9 in Indonesia. A hundred teachers of inclusive schools in Solo, Semarang, Surabaya, and Malang were involved as participants. Data were collected both online and paper-based. Confirmatory Factor Analysis (CFA) and Rasch test were used to analyze the data. Result showed that the UWES-9 had high values of: overall person and item interaction ($\alpha = .85$); person reliability ($= .71$); and item reliability ($= .95$). The person measure value ($= 1.65$ logit) and separation value 1.57 indicated two respondent groups based on the degree of work engagement. The item 5 had outfit value $MNSQ > 14$, meaning that it was unfit. Some items were found biased toward teaching duration in inclusive school. Internal structure analysis using CFA showed that UWES-9 is a unidimensional instrument with established goodness of fit value (RMSEA = .049; CFI, NFI, & GFI $> .95$). To conclude, UWES-9 has adequate psychometric properties, so it can be used for research or assessment.

Keywords: work engagement; UWES-9; scale adaptation; validity; Rasch model

Abstrak

Penggunaan skala dalam konteks budaya yang berbeda membutuhkan proses adaptasi agar skor yang diperoleh dapat diperbandingkan secara *fair*. *Utrecht Work Engagement Scale* dalam versi 9 item (UWES-9) telah diadaptasi di berbagai negara, namun belum dilakukan di Indonesia. Studi ini bertujuan untuk melakukan adaptasi terhadap UWES-9 di Indonesia. Seratus orang guru sekolah inklusi yang berasal dari Solo, Semarang, Surabaya dan Malang menjadi partisipan. Pengambilan data dilakukan secara online dan berbasis kertas. Analisis data menggunakan *Confirmatory Factor Analysis* (CFA) dan uji Rasch. Hasil analisis menunjukkan UWES-9 memiliki nilai yang tinggi dalam: interaksi antara person dan item secara keseluruhan ($\alpha = 0,85$); reliabilitas *person* ($= 0,71$); dan reliabilitas item ($= 0,95$). Nilai *person measure* ($= 1,65$ logit) dan nilai separation 1,57 menunjukkan terdapat 2 kelompok responden berdasarkan tingkat keterikatan kerja. Item no 5 memiliki nilai *outfit* $MNSQ > 1,4$ yang berarti kurang *fit*. Terdapat beberapa item yang bias terhadap lama mengajar di sekolah inklusi. Pengujian struktur internal dengan CFA menunjukkan bahwa UWES-9 merupakan instrumen yang bersifat unidimensional dengan nilai *goodness of fit* yang baik (RMSEA = 0,049; CFI, NFI, & GFI $> 0,95$). Dapat disimpulkan UWES-9 memiliki properti psikometris yang baik sehingga dapat digunakan dalam penelitian maupun asesmen.

Kata kunci: keterikatan kerja; UWES-9; adaptasi skala; validitas; model rasch

PENDAHULUAN

Keterikatan kerja merupakan hal yang menarik untuk dipelajari dalam konteks

kerja, pendidikan, kesehatan dan lainnya. Dalam dua dekade terakhir studi tentang keterikatan menjadi salah satu tren dalam perspektif psikologi positif, yaitu sebuah

pandangan yang berupaya untuk berfokus pada hal-hal positif dari diri individu daripada berfokus pada hal yang negatif. Schaufeli, Salanova, Gonzalez-Roma, dan Bakker (2002) mendefinisikan keterikatan kerja sebagai kondisi pikiran yang positif dan berhubungan dengan pekerjaan yang dicirikan oleh adanya *vigor*, *dedication*, dan *absorption*. Individu yang terikat akan menginvestasikan banyak energi personal dalam menjalankan peran di pekerjaan sebagaimana hal tersebut menjadi bagian penting dari identitas diri mereka. Sebaliknya, individu yang tidak terikat hanya hadir secara fisik dalam pekerjaan namun tidak menunjukkan adanya pengerahan emosi, energi, dan *passion* dalam menjalankan tugas. *Vigor* dicirikan oleh tingkat energi dan ketahanan mental yang tinggi dalam bekerja, kemauan untuk menginvestasikan usaha dalam bekerja, dan ketekunan dalam menghadapi kesulitan. *Dedication* mengacu pada keterikatan yang sangat kuat dalam pekerjaan dan mengalami rasa penting, antusiasme, inspirasi, kebanggaan, dan tantangan. *Absorption* ditandai dengan sepenuhnya terkonsentrasi dan senang dalam bekerja, dimana waktu berlalu dengan cepat dan memiliki kesulitan dengan memisahkan diri dari pekerjaan.

Sampai saat ini instrumen yang paling banyak digunakan untuk mengukur keterikatan kerja adalah *Utrecht Work Engagement Scale* (UWES) yang dikembangkan oleh Schaufeli dan Bakker. UWES dikembangkan untuk mengukur keterikatan kerja yang terdiri dari tiga aspek yaitu: *vigor*, *dedication*, dan *absorption*. Studi tentang validitas UWES telah banyak dilakukan, misalnya pengujian validitas antara keterikatan kerja dengan variabel yang menjadi antitesisnya maupun variabel yang memiliki kedekatan makna. Misalnya, pengujian validitas diskriminan antara keterikatan kerja dengan *burnout*. Hasilnya, ketiga aspek *burnout* yang diukur dengan *Maslach Burnout Inventory* (MBI) berhubungan negatif dengan tiga aspek dari keterikatan kerja. Indikasi keterikatan kerja

ditunjukkan oleh skor yang rendah pada skala kelelahan dan skala sinisme serta skor yang tinggi pada skala *professional efficacy* dari MBI (Demerouti, Bakker, Janssen, & Schaufeli, 2001; Schaufeli dkk., 2002; Montgomery, Peeters, Schaufeli, & Den-Ouden, 2003). Selain dengan *burnout*, uji validitas diskriminan keterikatan kerja juga telah dilakukan terhadap *workaholic*. Berdasarkan studi tersebut diketahui bahwa keterikatan dan *workaholic* adalah dua konsep yang sama sekali berbeda (Schaufeli, Taris, & Van Rhenen, 2003). Karyawan *workaholic* akan mengorbankan kesehatan mental dan kontak sosialnya di luar pekerjaan, sedangkan pekerja yang terikat merasa cukup baik secara mental maupun sosial (Schaufeli dkk., 2003).

Berbagai hasil studi validasi terdahulu menunjukkan bahwa UWES-9 memiliki properti psikometrik yang baik dan valid untuk digunakan dalam beberapa jenis pekerjaan dan budaya yang berbeda (Schaufeli & Bakker, 2004). Studi psikometri UWES dengan menggunakan database internasional (*cross-national study*) pada lebih dari 25 negara dan lebih dari 17 jenis pekerjaan menunjukkan nilai konsistensi internal (Cronbach's α) yang tergolong cukup bagus untuk versi UWES-17, 15, dan 9 yaitu di atas kriteria nilai minimal 0,60 (Nunnally & Bernstein, 1994). Studi psikometri UWES-9 melaporkan validitas meliputi struktur internal dengan nilai indeks *goodness of fit* rata-rata (AGFI, GFI, NFI, NNFI, dan CFI > 0,80), sedangkan untuk RMSEA pada analisis terpisah menunjukkan nilai di atas 0,05. Pada analisis *multiple group* menunjukkan nilai di bawah 0,05. Nilai rata-rata korelasi variabel laten dan manifes dalam UWES-9 adalah di atas 0,70 (Schaufeli & Bakker, 2004).

Terdapat beberapa versi dari UWES berdasarkan jumlah item dalam skala, yaitu 24, 17, 15, 9, dan 3 item. Mempersingkat versi asli UWES bertujuan untuk memenuhi kebutuhan para peneliti yang menghadapi keterbatasan waktu saat melakukan survei

keterikatan kerja. Versi yang singkat juga mengurangi efek keletihan, frustrasi dan penolakan partisipan untuk mengerjakan item yang terlalu banyak dan memakan waktu. Kondisi ini mendorong peneliti untuk mengembangkan UWES versi singkat yang valid dan reliabel (Schaufeli, Shimazu, Hakanen, Salanova, & De Witte, 2017). Kualitas psikometrik dari UWES-17 yang terkait dengan konsistensi internal, stabilitas maupun validitas konstruk telah dibuktikan dalam sejumlah besar studi. Proses iteratif yang digunakan dalam mengurangi jumlah item yang terdapat pada UWES-17 dijelaskan dengan detail dalam Schaufeli, Bakker, & Salanova (2006). Proses pengembangan UWES versi singkat menghasilkan UWES-9 yang memiliki kualitas psikometrik sebaik versi originalnya. Misalnya, pengujian konsistensi internal UWES-9 di 10 negara menunjukkan hasil berkisar antara 0,85–0,92 (Schaufeli dkk., 2006). Koefisien stabilitas UWES-9 dalam rentang waktu pengujian 16–18 bulan berkisar 0,70 (de Lange, De Witte, & Notelaers, 2008). Berdasarkan perbandingan yang sistematis melalui sejumlah studi psikometrik antara UWES-9 dengan UWES-17, Mills, Culbertson, dan Fullagar (2011) menyimpulkan bahwa UWES-9 layak bahkan lebih disukai daripada UWES-17, karena UWES-9 dinilai sebagai versi sederhana atau parsimoni dari UWES-17 dimana memiliki reliabilitas dan validitas yang serupa dengan UWES-17.

Terkait struktur faktor UWES terdapat bukti empiris yang sama-sama mendukung baik model hipotesis yang terdiri dari tiga faktor maupun model satu faktor. UWES-9 merupakan suatu kesatuan konstruk yang dibentuk oleh tiga aspek yang berhubungan kuat (Schaufeli dkk., 2017). Keterikatan kerja yang diukur dengan UWES dapat dinilai sebagai konstruk satu maupun tiga dimensi. Korelasi yang kuat antar ketiga dimensi dan nilai Cronbach yang tinggi pada skala total mendukung model satu dimensi, sementara nilai *fit* pada model tiga dimensi mendukung keberadaan tiga

subskala atau keterikatan sebagai konstruk multidimensional (setidaknya untuk UWES-15 dan UWES-17). Pada kasus di mana peneliti tertarik meneliti dimensi-dimensi yang berbeda dalam keterikatan kerja, maka instrumen multidimensional harus digunakan. Sebaliknya jika peneliti bermaksud mengukur konsep keterlibatan kerja sebagai suatu kesatuan maka dapat menggunakan skor total, dalam hal ini UWES-9 merupakan pilihan yang lebih baik. Karena ketiga subskala dari UWES memiliki korelasi yang kuat, maka tidak disarankan untuk memasukkan ketiga subskala tersebut secara simultan dalam analisis regresi multivariat. Hal ini dilakukan untuk menghindari terjadinya masalah multikolinieritas. Dalam kasus yang demikian lebih baik menggunakan UWES sebagai skor total (Schaufeli & Bakker, 2004). Hasil studi yang dilakukan oleh Seppala dkk. (2009) tentang validitas konstruk UWES-9 menunjukkan bahwa UWES-9 merupakan instrumen unidimensional.

Proses adaptasi bertujuan agar sebuah instrumen bersifat ekuivalen saat digunakan dalam konteks budaya yang berbeda. Proses adaptasi melibatkan metode kualitatif dalam memastikan ekuivalensi secara konseptual dan idiomatik. Selanjutnya dibutuhkan analisis statistik untuk memberikan informasi tentang properti psikometrik dan bukti validitas instrumen saat digunakan pada suatu budaya. Adaptasi dan validasi instrumen, merupakan proses yang berbeda namun bersifat komplementer (Borsa, Damásio, Bandeira, 2012). Pada proses adaptasi, peneliti harus memberikan bukti tentang ekuivalensi item secara semantik dan properti psikometrik yang adekuat dari versi baru instrumen yang diadaptasi (International Test Commission, 2016). Instrumen yang telah diadaptasi memberikan keuntungan karena memungkinkan peneliti untuk membandingkan data dari sampel dan latar belakang budaya yang berbeda. Penilaian dengan menggunakan instrumen hasil adaptasi menjadi lebih *fair* karena mengases

konstruk berdasarkan perspektif teoritis dan metode yang sama. Hal ini memberi peluang untuk dapat melakukan generalisasi sekaligus menyelidiki perbedaan dalam berbagai populasi (Borsa dkk., 2012).

Menurut Messick (1996) validitas bukan properti sebuah tes atau asesmen melainkan makna dari skor tes, yaitu kesimpulan tentang makna skor atau interpretasi dan implikasi tindakan yang terkait dari interpretasi skor tersebut. Karena itu, tingkat kepercayaan dari makna dan interpretasi skor sebuah skala/ tes merupakan aspek penting dalam validitas. Prinsip pada model Rasch terkait dengan validitas konstruk yang dikemukakan oleh Messick (Baghaei, 2008). Terdapat dua hal yang “mengancam” validitas konstruk, yaitu *construct under-representation* dan *construct irrelevant variance*. *Construct under-representation* merupakan ketidaksempurnaan (*imperfectness*) sebuah tes dalam mengases keseluruhan fitur sebuah konstruk, sedangkan *construct irrelevant variance*: terdapatnya sub-dimensi yang tidak relevan dengan konstruk, menyusup dan mengontaminasi dalam pengukuran (Messick, 1996). Analisis Rasch merupakan metode yang *powerful* untuk menguji validitas konstruk. Statistik kesesuaian (*fit statistic*) pada model Rasch menunjukkan indikasi terjadinya *construct irrelevant variance* sedangkan jarak (*gap*) pada Rasch *item-person map* merupakan penanda adanya *construct under-representation* (Baghaei, 2008).

Pengukuran dengan menggunakan skala umumnya menghasikan skor nominal atau ordinal. Metode analisis dengan menggunakan teori tes klasik memperlakukan skor mentah seolah-olah berskala interval atau rasio, sehingga memberikan hasil yang kurang akurat. Kelemahan lain pada penggunaan teori tes klasik adalah skor yang dipengaruhi oleh sampel soal dan karakteristik peserta tes. Model rasch memiliki keunggulan yang dapat mengatasi kelemahan yang terdapat

pada teori tes klasik. Dalam model rasch, skor mentah dikonversi ke dalam bentuk *odds ratio* dan selanjutnya melalui transformasi logaritma menjadi unit logit yang berskala interval dan bersifat linear. Dengan demikian tingginya *trait* yang dimiliki individu maupun tingkat kesulitan item akan dapat diestimasi dengan lebih tepat dan dapat diperbandingkan, karena memiliki satuan yang sama yaitu logit (Sumintono & Widhiarso, 2014). Meski demikian, studi-studi psikometrik UWES-9 umumnya menggunakan prosedur Teori Tes Klasik.

Studi ini bertujuan untuk melakukan adaptasi dan menguji validitas berdasarkan *evidence based on test content* dan *evidence based on internal structure* (American Educational Research Association, 2014) terhadap UWES-9. Model Rasch akan digunakan dalam analisis dimensi, invarians, keandalan skor, dan keberfungsian item diferensial (berdasarkan jenis kelamin dan lama mengajar inklusi).

METODE

Subjek

Data demografis responden disajikan pada Tabel 1.

Tabel 1.
Data Demografis Responden

Kategori	(N = 100)
<i>Gender</i>	
Laki-laki	15
Perempuan	85
<i>Usia</i>	
25 – 30 th	23
31 – 40 th	35
>40 th	42
<i>Lama mengajar inklusi</i>	
< 5 th	44
> 5 th	56
<i>Pendidikan</i>	
PLB	21
Non-PLB	79

Keterangan: PLB = Pendidikan Luar Biasa

Sebanyak 100 guru kelas di sekolah inklusi menjadi responden. Sekolah inklusi yang dimaksud adalah sekolah yang melibatkan siswa non berkebutuhan khusus dan siswa berkebutuhan khusus untuk belajar bersama dalam satu kelas. Guru-guru kelas di sekolah inklusi yang terlibat berasal dari berbagai sekolah di Kota Solo, Semarang, Surabaya, dan Malang yang mengisi skala secara daring melalui tautan bit.ly/keterikatanguru maupun secara langsung berbasis kertas.

Prosedur penelitian

Proses adaptasi dilakukan melalui serangkaian tahapan dengan mengacu pada panduan adaptasi tes dari *International Test Committee (ITC) guidelines for translating and adapting test* (2016). Berikut tahapan-tahapan dalam adaptasi skala yang dilakukan:

- a. *Pre-condition (PC) test* yaitu memperoleh izin yang diperlukan dari pemegang hak kekayaan intelektual berkaitan dengan tes sebelum melakukan adaptasi. Format asli skala UWES-9 diperoleh dari website tokoh keterikatan kerja Schaufeli (<http://www.wilmarschaufeli.nl>)
- b. Mengevaluasi konstruk: Pada tahapan ini, peneliti menggunakan *subject matter expert* dengan karakteristik: memiliki pengalaman dan pemahaman tentang konstruk keterikatan maupun konstruk yang berdekatan serta memiliki kemampuan bahasa Inggris yang baik. Evaluasi dilakukan secara kualitatif dengan mencermati apakah definisi dan isi konstruk cukup jelas, tidak *overlapping*, dan tidak ambigu pada konteks atau budaya dimana penelitian akan dilakukan.. Equivalensi/ kesetaraan konstruk, dimensi dan indikator UWES-9 dalam budaya Indonesia dilakukan melalui analisis semantik dan idiomatik. Item-item dalam UWES-9 terdistribusi sebagai berikut: item 1, 2, 5 mengukur *vigor*; item 3, 4, 7 mengukur *dedication*; item 6,8,9 mengukur *absorption*. Rating dilakukan dengan skala Likert: tidak pernah (0), hampir tidak pernah atau

beberapa kali atau jarang atau kurang dalam setahun (1), jarang/ sekali atau kurang dalam sebulan (2), kadang-kadang atau beberapa kali dalam sebulan (3), sering atau sekali seminggu (4), sangat sering atau beberapa kali dalam seminggu (5), dan selalu atau setiap hari (6).

- c. Proses penerjemahan (*translation and back translation*) dengan memastikan bahwa proses penerjemahan dan adaptasi mempertimbangkan perbedaan bahasa, psikologis, dan budaya dalam populasi yang dituju melalui pilihan ahli dengan keahlian yang relevan.
- d. Penilaian oleh *subject matter expert*, pada tahap ini ahli melakukan penilaian terhadap tata bahasa dan pilihan kata (*proofread*) terhadap hasil terjemahan.
- e. Skala siap uji dan *pilot testing* (lampiran)
Langkah ini sekaligus untuk menguji validitas skala berdasarkan *evidence based on test content* dan *evidence based on internal structure*.

Analisis Data

Estimasi reliabilitas dan validitas (struktur internal) dilakukan menggunakan pemodelan Rasch (Sumintono & Widhiarso, 2014). Model analisis rasch merupakan model matematika pengukuran yang menggambarkan bagaimana respon unik setiap individu terhadap item tertentu dalam instrumen. Model Rasch adalah bagian dari teori item respons. Dalam teori tes modern pemodelan rasch bisa memberikan informasi reliabilitas lebih spesifik ke tingkat item dan *person* (Verhelst, 1993). Analisis Rasch dilakukan menggunakan perangkat lunak Winsteps versi 3.75.0.

Sebelum melakukan analisis, transformasi data terlebih dahulu dilakukan dengan menggunakan scalogram. Dengan scalogram data disusun berdasar dari tingkat mudah ke tingkat tersulit, demikian juga secara bersamaan responden diurutkan dari peringkat yang terendah ke yang paling mampu (Sumintono & Widhiarso, 2014). Scalogram dilakukan dengan menggunakan

bantuan program winstep. Analisis Rasch dilakukan untuk menguji validitas struktur internal meliputi: unidimensionalitas (indeks *raw variance explained* harus >20%, indeks *raw variance unexplained* harus < 15%), *item fit* (indeks *item fit* dengan kriteria item yang *misfit* atau *outliers* adalah nilai *outfit* MNSQ yang diterima $0,5 < \text{MNSQ} < 1,5$, keandalan skor (*alpha cronbach* item yang bagus > 0,8), dan keberfungsian item diferensial (kriteria probabilitas item bias <5%).

HASIL DAN PEMBAHASAN

Hasil dari serangkaian proses adaptasi skala berupa skala UWES-9 siap uji dengan penyesuaian beberapa istilah atau kata yang memiliki kesepadanan makna namun tidak

ambigu bagi partisipan (dapat dilihat pada bagian lampiran).

Selanjutnya proses pengujian unidimensionalitas yang pertama dilakukan menggunakan CFA dengan model satu faktor dan tiga faktor berdasarkan pada konstruksi teoretis skala keterikatan kerja (UWES-9) meliputi: *vigor* (diwakili oleh item 1, 2, 5); *dedication* (diwakili oleh item 3, 4, 7); dan *absorption* (diwakili oleh item 6, 8, 9). Untuk menginterpretasikan kesesuaian model, indikator fit yang digunakan: *Root Mean Squared Error of Approximation* (RMSEA) < 0,05; *Comparative Fit Index* (CFI), NFI, & GFI > 0,95; maka menunjukkan kecocokan yang dapat diterima (Hu & Bentler, 1995). Hasil uji CFA dapat dilihat pada Tabel 2.

Tabel 2.
Fit Statistik dari Pengujian CFA Model Keterikatan Kerja

Model	λ^2 (df)	RMSEA	CFI	NFI	GFI
Unidimensi (1 faktor)	28,58 (18)	0,049	0,99	0,99	0,97
Multidimensi (3 faktor)	2296,23 (24)	0,10	0,96	0,96	0,95

Keterangan: $\lambda^2 = \text{chi-square}$, $df = \text{degrees of freedom}$, RMSEA = *Root mean square error of approximation*, CFI = *Comparative fit index*, NFI = *Normed fit index*; GFI = *Goodness of fit index*

Berdasarkan uji CFA, model unidimensi keterikatan kerja memenuhi semua kriteria nilai *fit* dibandingkan dengan model multidimensi dimana nilai RMSEA > 0,05 yang dapat dikatakan tidak fit. Dengan demikian maka UWES 9 dapat dikatakan sebagai instrumen keterikatan kerja yang bersifat unidimensional.

Pemodelan Rasch dilakukan untuk menguji validitas berdasar struktur internal meliputi: unidimensionalitas, *item fit*, keandalan skor (reliabilitas item dan *person*), dan keberfungsian item diferensial (DIF). Hasil analisis masing-masing disajikan pada tabel-Tabel 3.

Tabel 3.
Dimensi dan Invarians dalam Uji Dimensionalitas dengan Rasch

<i>Raw variance explained</i>		<i>Raw unexplained variance</i>	
<i>By measure</i>	53,4%	<i>In 1st contrast</i>	9,6%
<i>By person</i>	29,1%	<i>In 2nd contrast</i>	7,3%
<i>By items</i>	24,3%	<i>In 3rd contrast</i>	6,9%
		<i>In 4th contrast</i>	6,2%
		<i>In 5th contrast</i>	5,2%

Estimasi validitas pada pemodelan Rasch dilakukan melalui uji unidimensionalitas yang menggunakan prinsip analisis

komponen utama (*Principal Component Analysis*) dari residual yaitu mengukur sejauh mana instrumen mengukur apa yang

yang seharusnya diukur. Dari Tabel 3, diketahui hasil *pengukuran raw variance* data sebesar 53,4% (>20%) yang menunjukkan bahwa minimal persyaratan unidimensionalitas terpenuhi dengan kategori nilai sangat bagus (Sumintono & Widhiarso, 2014; Linacre, 2007). Varians yang tidak dapat dijelaskan oleh instrumen

tidak boleh melebihi 15% dan hasil estimasi menunjukkan nilai sebesar 9,6%. Instrumen UWES-9 dapat dikatakan mengukur konstruk keterikatan kerja sebagai konstruk unidimensional dan mengukur apa yang seharusnya diukur (dengan tidak ada varians lain diluar konstruk).

Tabel 4.
Pengukuran Item

Item	Nilai logit	S.E	Infit		Outfit	
			MNSQ	ZSTD	MNSQ	ZSTD
1	0,09	0,11	1,05	0,3	1,05	0,3
2	-0,06	0,12	0,87	-0,7	0,91	-0,3
3	-0,39	0,14	0,77	-1,1	0,70	-1,4
4	-0,15	0,12	0,81	-1,0	0,92	-0,3
5	0,71	0,10	1,76	3,9	1,83	3,9
6	-0,33	0,13	0,99	0,0	0,89	-0,4
7	-1,26	0,20	0,91	-0,3	0,69	-1,2
8	0,95	0,09	0,90	-0,61	1,00	0,1
9	0,44	0,10	1,01	0,1	1,08	0,5

Keterangan: S.E = Standar Error measurement, MNSQ = mean square, ZSTD = Z-standard

Pengujian terhadap setiap item dalam UWES-9 dilakukan untuk mengetahui informasi mengenai item. Berdasarkan nilai logit dapat diketahui bahwa item 9 dengan nilai logit tertinggi 0,95 yang menunjukkan bahwa item tersebut merupakan item yang paling sulit disetujui oleh responden. Item 9 berisi pernyataan "Saya merasa terbawa dengan pekerjaan ketika saya sedang bekerja". Sebaliknya, item yang paling mudah disetujui oleh responden adalah item 7 dengan nilai logit (-1,26). Pernyataan dalam item 7 adalah "Saya bangga dengan pekerjaan yang saya lakukan".

Selanjutnya *item fit* dalam pemodelan Rasch digunakan untuk memeriksa item-item mana yang sesuai dan yang tidak (*misfit* atau *outliers*). Kriteria yang digunakan untuk memeriksa item yang tidak sesuai antara lain (Sumintono & Widhiarso, 2010): 1) Nilai outfit MNSQ yang diterima adalah $0,5 < \text{MNSQ} < 1,5$ dan 2) Nilai outfit ZSTD yang diterima adalah $-2,0 < \text{ZSTD} < 2,0$. Berdasarkan kriteria dari Linacre (2007) dimana nilai MNSQ infit atau outfit item lebih dari 1,4 logit dikatakan sebagai item

yang *misfit* yang bermakna bahwa item tersebut membingungkan (Linacre, 2007). Dari Tabel 4 diketahui, item yang *misfit* adalah item nomor 5 dengan nilai MNSQ infit dan outfit $> 1,5$. Item no 5 dalam UWES-9 tertulis: "Saat bangun tidur di pagi hari, saya merasa bersemangat untuk pergi bekerja".

Karakteristik responden merupakan guru inklusi yaitu guru yang mengajar kelas yang melibatkan siswa berkebutuhan khusus dan non-berkebutuhan khusus secara bersama dalam pembelajaran. Pekerjaan guru dinilai sebagai pekerjaan yang penuh tekanan apalagi bekerja dengan siswa berkebutuhan khusus (Robertson, Chamberlain, & Kasari, 2003; Yell, Katsiyannis, Drasgow, & Herbst, 2003). Guru inklusi dituntut untuk sabar, penuh toleransi dan empati, penerimaan, dan sikap-sikap positif lainnya. Selain sikap positif dari guru, banyak tanggung jawab yang harus dipenuhi oleh guru inklusi antara lain: menyiapkan kurikulum, asesmen untuk siswa, mengajar sesuai dengan tingkat pemahaman dan perkembangan siswa, mensosialisasikan nilai-nilai inklusi dan

toleransi pada siswa non berkebutuhan khusus, orangtua, dan masyarakat (Hettiarachchi & Das, 2014; Lehohla & Hlalele, 2012). Selain itu, guru yang mengajar siswa berkebutuhan khusus sering dihadapkan dengan situasi yang tidak bisa diprediksi misalnya dengan tiba-tiba siswa akan mengamuk, menentang, tantrum, atau bahkan berperilaku agresif dan menyerang guru sedangkan guru harus menahan emosi-emosi negatifnya (Hochschild, 2012; Johnson dkk., 2005; Stansfeld, Rasul, Head, & Singleton, 2011).

Kondisi di Indonesia, stres yang dialami guru inklusi bisa diperparah dengan kurangnya dukungan dari pemerintah serta rendahnya kondisi ekonomi atau finansial yang menyebabkan layanan terhadap siswa menjadi tidak maksimal, memunculkan sikap negatif dan penolakan terhadap inklusi, hingga penolakan untuk berpartisipasi dan memenuhi tanggung jawab pekerjaan (Lewis & Doorlag, 2006; Harijanto, 2011; Pratiwi, 2016). Guru yang mengalami stres kerja tingkat tinggi berdampak negatif terhadap kebermaknaan dan keterikatan kerja mereka (George, Louw & Badenhorst, 2008; Schulze & Steyn, 2007). Gambaran kondisi guru inklusi khususnya di Indonesia dimungkinkan berelevansi dengan item yang sulit disetujui maupun *misfit* atau membingungkan dari UWES-9. Selain itu, pemilihan kata dalam item juga perlu diperhatikan dan tentu memiliki relevansi pula pada item *misfit*.

Berdasarkan Tabel 5, nilai *person reliability* 0,71 sedangkan nilai *item reliability* 0,95 dapat dikatakan bahwa konsistensi jawaban dari responden cukup bagus dimana kualitas item-item dalam instrumen juga istimewa (Bond & Fox, 2007). Nilai *item separation* = 4,58 (> 3) menunjukkan tingkat pengukuran tergolong sangat baik (Linacre, 2007). Nilai *separation* 1,57 atau dibulatkan menjadi 2 memenuhi batas minimum nilai *separation* dari Linacre (2007) sehingga dapat diterima dan menunjukkan tingkat pengukuran yang baik. Nilai *separation* 1,57 digunakan dalam

pengelompokan yang lebih detail melalui rumus pemisahan strata yang dapat dilihat pada (Sumintono & Widhiarso, 2014).

Tabel 5.
Nilai Reliabilitas dan *Separation*

	Item	Person	Alpha Cronbach (interaksi person-item)
Reliabilitas	0,95	0,71	0,85
<i>Separation</i>	4,58	1,57	

Nilai pemisahan strata 2,42 atau dapat dibulatkan menjadi 2 menunjukkan terdapat 2 kelompok responden yang dapat dibedakan oleh UWES-9 yaitu responden yang memiliki keterikatan kerja tinggi dan rendah. Selain nilai reliabilitas item dan *person*, nilai yang penting dan ditunjukkan dari pemodelan rasch yaitu nilai alpha cronbach (mengukur reliabilitas yaitu interaksi antara *person* dan item secara keseluruhan) sebesar 0,85 dan termasuk dalam kategori bagus sekali.

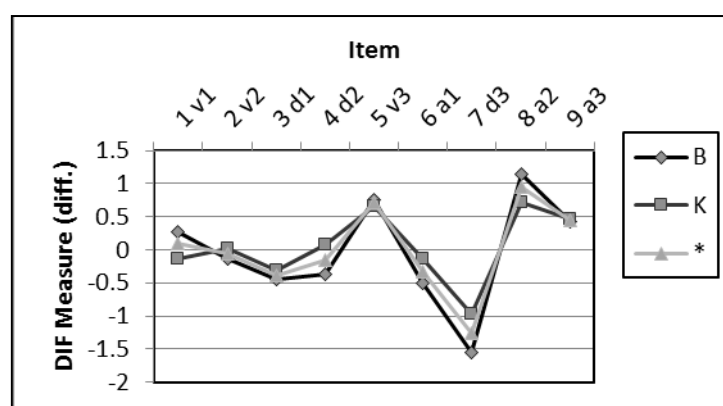
Untuk menguji item yang memiliki bias, digunakan indeks keberfungsian item diferensial (indeks DIF) dengan kriteria item dikatakan bias apabila nilai probabilitas item <5% (Linacre, 2012). Hasil uji DIF tersaji dalam Tabel 6.

Berdasarkan Tabel 6 UWES-9 tidak memiliki item yang bias terhadap gender maupun latar belakang pendidikan, namun terdapat item yang bias terhadap lama mengajar pada responden guru inklusi. Lama mengajar pada responden penelitian dikategorikan menjadi 2 kelompok yaitu lebih dari 5 tahun dan kurang dari 5 tahun. Berdasarkan kriteria indeks DIF (nilai D.F prob <5%), item yang bias terhadap lama mengajar adalah item 8. Pernyataan dalam item yang bias terhadap lama mengajar yaitu: "Saya larut dalam pekerjaan saya". Guru yang memiliki pengalaman mengajar > 5 tahun (B) menunjukkan respon yang berbeda dengan guru yang memiliki pengalaman mengajar < 5 tahun (K) sebagaimana ditunjukkan oleh Gambar 1.

Tabel 6.
Uji DIF Berdasarkan Gender dan Lama Mengajar

Item	Gender		Lama Mengajar		Latar Belakang Pendidikan	
	Ringkasan DIF chi-square	D.F prob.	Ringkasan DIF chi-square	D.F prob.	Ringkasan DIF chi-square	D.F prob.
1	0,0209	0,8850	2,8264	0,0927	0,0000	1,0000
2	0,2163	0,6419	0,4033	0,5254	0,0000	1,0000
3	0,0035	0,9529	0,2249	0,6353	0,0000	1,0000
4	2,1892	0,1390	3,2765	0,0703	0,0000	1,0000
5	2,5299	0,1117	0,2352	0,6277	0,0000	1,0000
6	0,0060	0,9383	2,0147	0,1558	0,0000	1,0000
7	0,0000	1,0000	2,2467	0,1339	0,0000	1,0000
8	1,4395	0,2302	5,0102	0,0252	0,0000	1,0000
9	0,0455	0,8312	0,0822	0,7744	0,0000	1,0000

Keterangan: item bias jika D.F prob < 0,500



Gambar 1. Person DIF Plot Berdasarkan Lama Mengajar

Keterangan: B= lama mengajar > 5 tahun, K= lama mengajar < 5 tahun, *= kontras

Pembahasan perbedaan respon berdasarkan lama mengajar guru didukung oleh hasil penelitian terdahulu. Pengalaman dan pelatihan telah diteliti berhubungan dan turut menentukan sikap positif guru terhadap inklusi (Avramidis & Kalyva, 2007; MacFarlane & Woolfson, 2013). Sikap positif guru terhadap inklusi dapat mengarahkan intensi untuk berperilaku kerja yang positif termasuk perilaku terikat terhadap inklusi (Ajzen, 2005). Lebih spesifik, pelatihan dan pengembangan telah diteliti turut menentukan keterikatan guru (Altunela, Kocak, & Cankir, 2015; Bakker & Bal, 2010).

UWES merupakan skala pertama dan paling sering digunakan untuk mengukur

keterikatan kerja. Dalam perkembangannya, UWES mengalami beberapa kali perubahan versi meskipun hanya pada jumlah itemnya yaitu dari versi 24, 17, 15, 9, bahkan versi terbaru yaitu 3 item. Hal ini menjadi tantangan tersendiri bagi peneliti yang tertarik untuk menggunakan UWES terutama dalam versi singkat. Tantangannya adalah dalam upaya memenuhi kriteria atau kualitas psikometri yang ketat terkait dengan pengukuran *conjoint* aditif sebagaimana ditawarkan oleh model rasch. Model pengukuran rasch memberikan dasar ilmiah untuk membangun pendekatan pengukuran linier yang sekarang ini banyak digunakan dalam ilmu kesehatan dan sosial (Keenan, Redmond, Horton, Conaghan, & Tennant, 2007; Kyriakides, Kaloyirou, Lindsay,

2006). Hasil analisis menunjukkan bahwa 9 item dari skala UWES selaras dengan ekspektasi model rasch dan memberikan skor mentah yang reliabel pada level item maupun person yang ditunjukkan dengan nilai reliabilitas interaksi person dengan total item = 0,85.

SIMPULAN

Skala UWES-9 meskipun menunjukkan sifat unidimensional dalam mengukur keterikatan kerja namun mengandung bias lama mengajar. Pengujian validitas dengan menggunakan model rasch menunjukkan bahwa skala ini memiliki nilai reliabilitas item maupun *person* yang tergolong tinggi tanpa tergantung pada jumlah dan karakteristik responden sebagaimana pada pendekatan teori tes klasik. Dengan demikian UWES-9 dapat direkomendasikan untuk mengukur keterikatan kerja. Hasil penelitian ini menunjukkan bahwa item nomor 5 perlu mendapatkan perhatian khusus karena berdasarkan uji CFA dan rasch, item tersebut teridentifikasi sebagai item yang membingungkan bagi responden dan memiliki nilai outfit MNSQ > 1,4 yang berarti kurang sesuai mengukur *vigor* sebagai aspek keterikatan kerja.

UCAPAN TERIMA KASIH

Peneliti menyampaikan terima kasih pada Lembaga Pengelola Dana Pendidikan sebagai pemberi dana. Ucapan terima kasih juga disampaikan pada Ketua Program Studi S3 Fakultas Psikologi Universitas Airlangga dan Dekan Fakultas Psikologi Universitas Diponegoro atas dukungan dan motivasi yang diberikan pada peneliti.

DAFTAR PUSTAKA

Altunela, M., C., Kocak, O., E., & Cankir, B. 2015. The Effect of Job Resources on Work Engagement: A Study on Academicians in Turkey. *Educational Sciences: Theory and practice*. 15, 2,

409-417. Retrieved in September 21st 2017. doi: 10.12738/estp.2015.2.2349.

American Educational Research Association. (2014). *Standards for educational and psychological testing/ American Educational Research Association, American Psychological Association, National Council on Measurement in Education*. Washington DC: American Educational Research Association.

Ajzen, I. (2005). *Attitudes, Personality and Behavior (Second Edition)*. New York: Open University Press.

Avramidis, E., & Kalyva, E. (2007). The influence of teaching experience and professional development on Greek teachers' attitudes towards inclusion. *European Journal of Special Needs Education*, 22(4), 367-389.

Baghaei, P. (2008). The Rasch model as a construct validation tool. *Rasch Measurement Transactions*, 22, 1145-1146.

Bakker, A. B., & Bal, P. M. (2010). Weekly work engagement and performance: A study among starting teachers. *Journal of Occupational and Organizational Psychology*, 83(1), 189-206. <http://doi.org/10.1348/096317909X402596>.

Borsa, J. B., Damásio, B. F., & Bandeira, D. R. (2012). Cross-cultural adaptation and validation of psychological instruments: Some considerations. *Paidéia set.-dec*, 22(53), 423-432. doi:<http://dx.doi.org/10.1590/1982-43272253201314>.

Bond, T. G., & Fox, C. M. (2007). *Applying the rasch model: Fundamental measurement in the human sciences*. New York: Springer Dordrecht Heidelberg.

- de Lange, A. H., De Witte, H., & Notelaers, G. (2008). Should I stay or should I go? Examining longitudinal relations among job resources and work engagement for stayers versus movers. *Work & Stress*, 22, 201–223. doi:10.1080/02678370802390132
- Demerouti, E., Bakker, A. B., Janssen, P. P., M. & Schaufeli, W. B. (2001). Burnout and engagement at work as a function of demands and control. *Scandinavian Journal of Work, Environment & Health*, 27, 279-286.
- George, E., Louw, D., & Badenhorst, G. (2008). Job satisfaction among urban secondary-school teachers in Namibia. *South African Journal of Education*, 28(2), 135-154.
- Harijanto, S. (2011) *The Indonesian government policy on special education in global perspective*. Artikel dipresentasikan pada International seminar on Special Education, 19 Maret 2011.
- Hettiarachchi, S., & Das, A. (2014). Perceptions of inclusion and perceived preparedness among school teachers in Sri Lanka. *Teaching and Teacher Education*, 43, 143–153. <http://doi.org/10.1016/j.tate.2014.07.003>.
- Hochschild, A. R. (2012). *The managed heart: commercialization of human feeling*. Updated with a New Preface edition. (3rd ed.). Berkeley: University of California Press.
- Hu, L. T., & Bentler, P. M. (1995). Evaluating model fit. Dalam R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications*. London: Sage.
- International Test Commission. (2016). *ITC guidelines for translating and adapting tests (second edition)*. ITC, www.InTestCom.org.
- Johnson, S., Cooper, C., Cartwright, S., Donald, I., Taylor, P., & Millet, C. (2005). The experience of work-related stress across occupations. *Journal of Managerial Psychology*, 20(2), 178-187. doi: 10.1108/02683940510579803
- Keenan, A. M., Redmond, A., Horton, M., Conaghan P., & Tennant, A. (2007). The foot posture index: Rasch analysis of a novel, foot specific outcome measure. *Archives Physical Medicine and Rehabilitation*, 88, 88-93.
- Kyriakides, L., Kaloyirou, C., & Lindsay, G. (2006). An analysis of the revised olweus bully/ victim questionnaire using the rasch measurement model. *British Journal of Educational Psychology*, 76(4), 781-801.
- Lehohla, M., & Hlalele, D. (2012). Inclusive Classrooms: An Ecosystemic Perspective. *Journal of Human Ecology*, 37(3), 189–201. <http://doi.org/10.1080/09709274.2012.11906464>.
- Lewis, R. B., & Doorlag, D. H. (2006). *Teaching special students in general education classroom*. New Jersey: Pearson Education.
- Linacre, J. M. (2007). *A user's guide to WINTEPS Rasch-model computer programs*. Chicago, Illinois: MESA Press.
- Linacre, J. M. (2012). *Winsteps Rasch measurement computer program: User's guide*. Beaverton: Winsteps.com.
- MacFarlane, K., & Woolfson, L. M. (2013). Teacher attitudes and behavior toward the inclusion of children with social, emotional and behavioral difficulties in

- mainstream schools: An application of the theory of planned behavior. *Teaching and teacher education*, 29, 46-52.
- Messick, S. (1996). Validity and washback in language testing. *Language Testing*, 13(3): 241-256.
- Mills, M. J., Culbertson, S. S., & Fullagar, C. J. (2011). Conceptualizing and measuring engagement: An analysis of the Utrecht Work Engagement Scale. *Journal of Happiness Studies*, 13, 519–545. doi: 10.1007/s10902-011-9277-3
- Montgomery, A., Peeters, M. C. W., Schaufeli, W. B., & Den-Ouden, M. (2003). Work-home interference among newspaper managers: Its relationship with Burnout and engagement. *Anxiety, Stress & Coping*, 16, 195-211.
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory (McGraw-Hill Series in Psychology) (Vol. 3)*. New York: Mc Graw Hill.
- Pratiwi, J. C. (2016). Sekolah Inklusi Untuk Anak Berkebutuhan Khusus: Tanggapan Terhadap Tantangan Kedepannya. *Prosiding Ilmu Pendidikan*, 1(2).
- Robertson, K., Chamberlain, B., & Kasari, C. (2003). General education teachers' relationships with included students with autism. *Journal of Autism and Developmental Disorders*, 33, 123–130. doi: 10.1023/A:1022979108096.
- Schaufeli, W. B., Salanova, M., Gonzalez-Roma, V., & Bakker, A. B. (2002). The measurement of engagement and burnout: A two-sample confirmatory factor analytic approach. *Journal of Happiness Studies*, 3, 71–92. doi: 10.1023/A:1015630930326.
- Schaufeli, W. B., Taris, T. W., & Van-Rhenen, W. (2003). *Workaholism, burnout and engagement: Three of a kind or three different kinds of employee well-being?* Submitted for publication.
- Schaufeli, W. B., & Bakker, A. B. (2004). *Utrecht Work Engagement Scale: Preliminary manual*. Utrecht University: Occupational Health Psychology Unit.
- Schaufeli, W. B., Bakker, A. B., & Salanova, M. (2006). The measurement of work engagement with a short questionnaire: A cross-national study. *Educational and Psychological Measurement*, 66, 701–716. doi: 10.1177/0013164405282471.
- Schaufeli, W. B., Shimazu, A., Hakanen, J., Salanova, M., & De Witte, H. (2017). An Ultra-Short Measure for Work Engagement: The UWES-3 Validation Across Five Countries. *European Journal of Psychological Assessment*. Advance online publication. <http://dx.doi.org/10.1027/1015-5759/a000430>.
- Schulze, S., & Steyn, T. (2007). Stressors in the professional lives of South African secondary school learners. *South African Journal of Education*, 27(4), 691-708.
- Seppala, P., Mauno, S., Feldt, T., Hakanen, J., Kinnunen, U., Tolvanen, A., & Schaufeli, W. (2009). The construct validity of Utrecht work engagement scale: Multisample and longitudinal evidence. *Journal Happiness Study*, 10, 459-481. doi: 10.1007/s10902-008-9100-y.
- Stansfeld, S. A., Rasul, F. R., Head, J., & Singleton, N. (2011). Occupation and mental health in a national UK survey. *Social psychiatry and*

- psychiatric epidemiology*, 46(2), 101-110. doi: 10.1007/s00127-009-0173-7
- Sumintono, B. & Widhiarso, W. (2014). *Aplikasi model rasch untuk penelitian ilmu-ilmu sosial, edisi revisi*. Jakarta: Trim Komunikata.
- Verhelst, N., D. (1993). Item response theory (Item response theory). In P. F. Sanders (Ed.), *Psychometrie in de praktijk (Practice of Psychometrics)*. Arnhem, The Netherlands: Cito Institute for Test Development.
- Yell, M. L., Katsiyannis, A., Drasgow, E., & Herbst, M. (2003). Developing legally correct and educationally appropriate programs for students with autism spectrum disorders. *Focus on Autism and Other Developmental Disabilities*, 18, 182–191.

Appendix. UWES-9 Versi Bahasa Indonesia

UWES

Berikut 9 pernyataan tentang perasaan Saudara di tempat kerja. Silakan dibaca dengan cermat dan berilah angka (0-6) pada titik-titik di dalam kolom di depan setiap pernyataan sesuai dengan kondisi yang Saudara rasakan di tempat kerja. Apabila Saudara belum pernah mengalami perasaan tersebut, tuliskan angka 0 (nol). Apabila Saudara pernah mengalami perasaan tersebut, tunjukkan seberapa sering Saudara merasakannya dengan cara menuliskan angka (dari 1 hingga 6) yang menggambarkan seberapa tepat Saudara sering merasakan hal tersebut.

	Hampir tidak pernah	Jarang	Kadang-kadang	Sering	Sangat sering	Selalu
0	1	2	3	4	5	6
Tidak pernah	Beberapa kali atau kurang dalam setahun	Sekali atau kurang dalam sebulan	Beberapa kali dalam sebulan	Sekali seminggu	Beberapa kali dalam seminggu	Setiap hari

1. (.....) Di tempat kerja, saya merasa penuh dengan energi
2. (.....) Dalam bekerja, saya merasa kuat dan bertenaga
3. (.....) Saya antusias dalam melakukan pekerjaan saya
4. (.....) Pekerjaan saya menginspirasi saya
5. (.....) Saat bangun tidur di pagi hari, saya merasa bersemangat untuk pergi bekerja
6. (.....) Saya merasa senang saat saya bekerja secara intensif
7. (.....) Saya bangga dengan pekerjaan yang saya lakukan
8. (.....) Saya larut dalam pekerjaan saya
9. (.....) Saya merasa terbawa dengan pekerjaan ketika saya sedang bekerja