

PEMANFAATAN APLIKASI WINSTEPS UNTUK ANALISIS TES KINERJA PADA GURU-GURU SEKOLAH MENENGAH KEJURUAN

Ni Made Sri Mertasari¹, I Made Candiasa², Ida Bagus Gede Purwa³

^{1,2} Jurusan Matematika FMIPA UNDIKSHA; ³ Perpustakaan Pusat UNDIKSHA

Email: srimertasari@undiksha.ac.id

ABSTRACT

Vocational schools are programmed to produce skilled workers in various fields to support national development. The learning process in vocational schools emphasizes competence on the skill side through practical activities. Evaluation is carried out by performance tests, in order to obtain authentic student competencies. However, teachers still have difficulty in validating performance tests. This service program is intended to improve the competence of vocational school teachers in validating performance tests by training approach. The training begins with a discussion on the application of the Partial Credit Model (PCM) technique for testing performance tests followed by introducing Winsteps Applications. Monitoring and evaluation were carried out through observation. Evaluation results showed that participants were able to master PCM and were able to use the Winsteps Application. Participants are even able to interpret the output of Winsteps, so that the conclusion process regarding the quality of the test can be known quickly.

Key words: *skills, performance test, partial credit model, winsteps.*

ABSTRAK

Sekolah kejuruan diprogramkan untuk dapat menghasilkan tenaga kerja terampil di berbagai bidang untuk mendukung pembangunan nasional. Proses pembelajaran di sekolah kejuruan lebih menekankan kompetensi pada sisi keterampilan melalui kegiatan praktikum. Evaluasi dilakukan dengan tes kinerja, agar diperoleh informasi tentang kompetensi peserta didik yang otentik. Namun masih ada kesulitan di pihak guru dalam menguji tes kinerja. Program pengabdian ini ditujukan untuk meningkatkan kompetensi guru-guru sekolah kejuruan dalam menguji kualitas tes kinerja. Pengabdian menggunakan pendekatan pelatihan dengan memanfaatkan Aplikasi Winsteps. Pelatihan diawali dengan diskusi tentang penerapan teknik Partial Credit Model (PCM) untuk pengujian tes kinerja. Menghindari perhitungan yang rumit, selanjutnya dilakukan pengenalan Aplikasi Winsteps untuk mendukung PCM. Setelah pelatihan dilakukan pendampingan dan evaluasi melalui observasi. Hasil evaluasi menunjukkan bahwa peserta mampu menguasai PCM serta memanfaatkan Aplikasi Winsteps untuk pengujian tes kinerja. Peserta bahkan sudah mampu melakukan interpretasi terhadap output Winsteps, sehingga proses kesimpulan terkait kualitas tes dapat diketahui dengan cepat.

Kata-kata kunci: *keterampilan, tes kinerja, partial credit model, winsteps.*

PENDAHULUAN

Pendidikan kejuruan merupakan kekuatan pendorong pertumbuhan sosial, ekonomi, dan teknologi (Agrawal, 2013; Goel, 2009). Dunia

usaha dan dunia industri (DUDI) memerlukan tenaga terampil yang siap pakai. Tenaga terampil yang memiliki daya saing dihasilkan melalui pendidikan kejuruan (Oliver, 2010; Newaz, dkk., 2013; Cedefop, 2017). Tenaga

kerja yang sudah siap pakai akan meningkatkan efisiensi DUDI karena tidak lagi diperlukan waktu, biaya, dan tenaga untuk melatih tenaga kerja. Oleh karena itu, secara tidak langsung pendidikan kejuruan juga berkontribusi dalam meningkatkan produktivitas perusahaan (Agrawal, 2013; Goel, 2009). Tidak sedikit pendidikan kejuruan yang mampu menghasilkan inovasi yang dapat dijadikan produk unggulan bagi DUDI. Program-program unggulan yang ditawarkan pendidikan kejuruan merupakan bidang yang menjadi idola anak muda, seperti komputer, elektronika, otomotif. Kecerdasan daya belajar tinggi yang dimiliki anak muda menjadikan mereka mampu menguasai keterampilan di bidang-bidang tersebut dengan lebih mudah.

Lulusan pendidikan kejuruan dituntut tidak hanya memiliki keahlian, melainkan juga mampu belajar apa yang harus dilakukan (Modie, 2006). Oleh karena itu, pendidikan kejuruan memerlukan model pembelajaran yang dapat menghasilkan kompetensi atau keahlian yang relevan dengan kebutuhan dunia kerja dan mampu berkembang sesuai perubahan. Dengan demikian mereka tidak lagi perlu mengikuti pelatihan, yang tentunya memerlukan tenaga, waktu, dan biaya. Di lain sisi evaluasi pembelajaran juga harus mampu memberikan informasi tentang hasil belajar peserta didik yang otentik. Informasi otentik kemajuan belajar peserta didik yang otentik menjadi informasi yang amat berguna dalam meningkatkan kualitas pembelajaran (Heritage, 2010; Serafini, 2010). Tes sebagai instrumen evaluasi harus dipersiapkan dengan baik, agar mampu mengukur hasil belajar dengan baik.

Tes kinerja dipandang dapat mengukur keterampilan dengan lebih otentik karena mampu mengukur kinerja tertentu sesuai dengan kompetensi pribadi peserta didik (Brown, 2004; Linn, dkk., 1991). Oleh karena itu, tepat kiranya kalau tes kinerja diterapkan pada pembelajaran di sekolah kejuruan. Hal yang juga mendukung adalah karena tes kinerja melibatkan pengamatan perilaku yang

ditunjukkan atau produk dari perilaku tersebut (Callahan and Clark, 1977). Pemberian skor pada tes kinerja didasarkan pada kesamaan kinerja yang ditunjukkan peserta didik dan kinerja yang sebenarnya diinginkan sebagai kriteria (Palm, 2008). Akibatnya, penyiapan tes kinerja beserta rubrik penilaiannya harus dilakukan dengan baik. Analisis situasi menunjukkan bahwa guru-guru di sekolah kejuruan sudah mampu menyusun tes kinerja dan rubriknya dengan baik. Permasalahannya ada pada pengujian tes kinerja yang berhasil disusun.

Salah satu teknik yang dapat diterapkan dalam pengujian tes kinerja adalah Partial Credit Model (PCM). PCM awalnya dikembangkan pertamakali oleh Master (1988) sebagai model analisis yang dapat diterapkan untuk tes politomi. PCM merupakan model untuk menganalisis solusi bertahap dari butir tes politomi dengan hanya mempertimbangkan parameter tingkat kesulitan pada tiap-tiap langkah penyelesaian (Verhelst, 2008). Tujuan PCM adalah menggambarkan kompetensi peserta didik dari yang paling sederhana sampai dengan yang paling kompleks secara berkelanjutan (Eggert & Geholz, 2009). Mirip dengan model Rasch, PCM mengasumsikan bahwa semua langkah penyelesaian memiliki daya beda yang sama, sehingga parameter tingkat kesukaran merupakan satu-satunya karakteristik yang dianggap mempengaruhi hasil belajar. Artinya, peluang siswa untuk berhasil dalam menjawab butir soal hanya tergantung pada kemampuan siswa dan tingkat kesukaran butir tes.

Skor kategori pada PCM menunjukkan banyaknya langkah untuk menyelesaikan dengan benar butir tersebut. Skor pada kategori yang lebih tinggi menuntut kemampuan yang lebih besar daripada skor pada kategori yang lebih rendah. Jika i adalah butir politomi dengan kategori skor x sebesar $0, 1, 2, \dots, m_i$. Probabilitas respons individu n dengan tingkat kemampuan θ memperoleh skor kategori j

pada butir soal i , dinyatakan dengan persamaan, sebagai berikut (Master, 1988).

$$P_{ni}(\theta) = \frac{\exp \sum_{j=0}^x (\theta - \delta_j)}{\sum_{h=0}^m \exp \sum_{j=0}^h (\theta - \delta_j)} \quad \text{untuk } x = 0, 1, 2, \dots, m$$

Dalam formula (1) di atas: $P_{ni}(\theta)$ = peluang peserta ke- n dengan tingkat kemampuan θ memperoleh skor x pada butir i ; x = skor peserta; j = Kategori/tahap dalam butir tes; i = butir soal; n = peserta; θ = tingkat kemampuan peserta ke n ; δ_j = tingkat kesukaran tahap (j) pada butir (i).

PCM merupakan pengembangan dari model Rasch butir dikotomi yang diterapkan pada butir politomi. Asumsi yang berlaku pada model Rasch juga berlaku pada model PCM, seperti berikut (Hambleton, Swaminathan, dan Rogers, 1991). 1) Unidimensi, setiap butir hanya mengukur satu ciri peserta didik. 2) Independensi lokal, respon pada butir yang satu bebas dari pengaruh respon pada butir lain jika kemampuan yang mempengaruhi performansi dibuat konstan. 3) Fungsi karakteristik butir atau kurva karakteristik butir merefleksikan hubungan yang sebenarnya antara kemampuan dan respon peserta terhadap butir tes. PCM termasuk teori tes modern yang membebaskan pengaruh kelompok responden terhadap hasil analisis. Perhitungan pada PCM cukup rumit. Menghindari faktor kesulitan yang dapat berimplikasi pada pemborosan waktu dan tenaga, pengujian tes kinerja dengan model PCM dibantu dengan paket Aplikasi Winsteps.

METODE

Tujuan pengabdian kepada masyarakat yang dilakukan adalah peningkatan kemampuan guru-guru sekolah menengah kejuruan (SMK) dalam analisis tes kinerja. Pertimbangannya adalah bahwa guru-guru SMK lebih banyak menerapkan tes kinerja dalam pembelajaran karena pendidikan di SMK lebih menekankan pada aspek keterampilan. Khalayak sarannya

adalah guru-guru di SMK Negeri-2 Tabanan. Upaya peningkatan kemampuan guru sekolah kejuruan dalam analisis tes kinerja dilakukan dengan metode pelatihan disertai pendampingan dan monitoring. Kegiatan diawali dengan pengenalan model PCM untuk menganalisis tes kinerja.

Disadari bahwa perhitungan pada PCM cukup rumit, seperti tampak pada penerapan formula (1). Selain itu, perhitungan pada model PCM juga harus didahului dengan perhitungan indeks kesukaran butir. Bila perhitungan dilakukan secara manual, maka dikhawatirkan banyak menghabiskan waktu dan tenaga. Mempertimbangkan kondisi tersebut, sesuai perencanaan, analisis dibantu dengan aplikasi Winsteps. Pada tahap berikutnya dilakukan pelatihan mengoperasikan paket aplikasi Winsteps untuk analisis butir tes. Langkah-langkah operasional Winsteps diawali dengan membuat data pada Aplikasi *Microsoft Excel*. Data tersebut kemudian disimpan dan dilanjutkan dengan membuka aplikasi *Winsteps*. Data yang sudah dibuat pada aplikasi *Microsoft Excel* kemudian diimpor. Pada saat impor ada beberapa pilihan yang muncul, seperti tampak pada Gambar-1 dan dalam hal ini dipilih *Excel* karena data dibuat dengan aplikasi *Microsoft Excel*.



Gambar-1: Menu Impor Data

Data yang sudah diimpor selanjutnya disimpan dalam bentuk dokumen teks. Pada langkah analisis, file data berbentuk teks dokumen tersebut harus dibuka terlebih dahulu, dan selanjutnya dianalisis. Hasil analisis dapat ditampilkan dengan memilih *output tables*. Output menampilkan beberapa tabel, antara lain *Separation Table*, *Item Measure*, *Item Fit Order*, *Person Measure*, dan *Person Fit Order*.

Setelah pelatihan operasional Winsteps peserta sudah mampu melakukan analisis butir

tes dan mereka mencoba sesuai kemampuan. Selanjutnya dilakukan pendampingan dan monitoring. Sampai saat ini pendampingan dan monitoring masih berlangsung. Kendati demikian, evaluasi sudah dilakukan untuk mendapatkan umpan balik dari program pengabdian. Evaluasi dilakukan melalui observasi disertai dengan wawancara.

HASIL DAN PEMBAHASAN

Peserta pelatihan sangat bervariasi dari sisi kompetensi dan pengalaman. Beberapa peserta sudah memiliki kompetensi dalam hal analisis butir dengan teori tes modern dengan menggunakan model Rasch atau Item Response Theory (IRT). Bahkan beberapa sudah memahami PCM, walaupun masih dalam tahap teori. Di lain sisi, mayoritas peserta belum memiliki kompetensi di bidang teori tes modern. Khusus untuk pemanfaatan aplikasi Winsteps dalam analisis butir tes memang semua peserta belum memiliki pengalaman. Situasi seperti ini dimanfaatkan dengan baik melalui forum diskusi kelompok dengan memanfaatkan tutor sebaya. Pendekatan tersebut ternyata dapat mempercepat penguasaan peserta terhadap penerapan model PCM untuk analisis butir tes. Sesuai dugaan awal, mayoritas peserta menyatakan bahwa perhitungan PCM sangat rumit, sehingga banyak menyita waktu dan tenaga. Oleh karena itu, pemanfaatan aplikasi pendukung menjadi solusi terbaik. Aplikasi yang sudah disiapkan adalah Winsteps dan semua peserta menyatakan setuju untuk memanfaatkan aplikasi Winsteps untuk analisis tes kinerja. Untuk itu persiapan segera dilakukan dengan memasang aplikasi Winsteps pada masing-masing komputer milik peserta. Sebagai langkah awal, peserta diberikan tutorial prosedur operasional aplikasi Winsteps, termasuk pemanfaatan *Microsoft Excel* sebagai aplikasi untuk memasukkan data. Memang aplikasi Winsteps mampu mengimpor data dari berbagai aplikasi selain *Microsoft Excel*, seperti *R*, *SAS*, *SPSS*,

Observasi dilakukan untuk mendapatkan informasi tentang tingkat penguasaan peserta. Selanjutnya, wawancara dilakukan untuk menggali lebih lanjut permasalahan yang dialami peserta. Dengan demikian, dapat dirumuskan langkah tindak lanjut untuk mendapatkan hasil yang lebih optimal.

STATA, *Tabbed Text*. Namun dengan mempertimbangkan pengalaman peserta dalam mengoperasikan aplikasi-aplikasi tersebut, pilihan jatuh pada *Microsoft Excel*. Alasannya sangat sederhana, yakni hampir semua peserta sudah berpengalaman menggunakan *Microsoft Excel*, kendati tingkat penguasaannya bervariasi. Sebagai tindak lanjut kondisi tersebut, peserta hanya diperkenalkan teknik impor data dari *Microsoft Excel* ke aplikasi Winsteps. Sudah barang tentu sistematika data sudah diperkenalkan terlebih dahulu.

Pelatihan penerapan aplikasi Winsteps untuk analisis butir berjalan dengan lancar. Faktor pendukungnya adalah semua peserta sudah mampu memanfaatkan aplikasi *Microsoft Excel*, dan bahkan beberapa diantaranya sudah mahir, sehingga penekanan hanya pada impor data, analisis, dan interpretasi hasil. Pendekatan asistensi selama pelatihan memberi manfaat yang sangat positif dalam mengoperasikan aplikasi Winsteps. Hanya dalam satu hari semua peserta sudah mampu mengoperasikan Winsteps, sekalipun data masih disiapkan panitia penyelenggara. Pada hari-hari berikutnya peserta sudah mampu melakukan analisis data masing-masing dengan memanfaatkan aplikasi Winsteps. Semua output Winsteps sudah mampu ditampilkan dengan baik dan sudah mampu diinterpretasikan dengan tepat. Sebagai informasi tambahan, pada kesempatan ini disajikan tiga tabel yang merupakan output penting dari Winsteps yang berhasil

ditampilkan peserta, yakni tabel pemisah (*separation table*) seperti tampak pada Tabel-1, tabel ukuran butir (*item measure*)

yang tampak pada Tabel-2 dan tabel ukuran peserta (*Person Measure*) yang tampak pada Tabel-3.

Tabel-1: Tabel Pemisah (*Separation Table*)

52-638WS - Notepad
 File Edit Format View Help
 TABLE 3.1 data psilitomi.xlsx
 20U638WS.TXT Jul 9 2021 9: 3
 INPUT: 20 PERSON 10 ITEM REPORTED: 20 PERSON 10 ITEM 5 CATS MINISTEP 4.8.2.0

SUMMARY OF 20 MEASURED PERSON

	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT		OUTFIT	
					MNSQ	ZSTD	MNSQ	ZSTD
MEAN	29.3	10.0	-.12	.35	.98	.03	.98	.03
SEM	1.7	.0	.19	.01	.08	.20	.08	.19
P.SD	7.6	.0	.85	.04	.34	.85	.33	.84
S.SD	7.8	.0	.87	.04	.35	.87	.34	.86
MAX.	42.0	10.0	1.38	.48	1.58	1.43	1.57	1.42
MIN.	15.0	10.0	-1.93	.31	.45	-1.46	.47	-1.45
REAL RMSE	.37	TRUE SD	.76	SEPARATION	2.07	PERSON RELIABILITY	.81	
MODEL RMSE	.35	TRUE SD	.77	SEPARATION	2.21	PERSON RELIABILITY	.83	
S.E. OF PERSON MEAN = .19								
PERSON RAW SCORE-TO-MEASURE CORRELATION = 1.00								
CRONBACH ALPHA (KR-20) PERSON RAW SCORE "TEST" RELIABILITY = .83 SEM = 3.12								
STANDARDIZED (50 ITEM) RELIABILITY = .96								

SUMMARY OF 10 MEASURED ITEM

	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT		OUTFIT	
					MNSQ	ZSTD	MNSQ	ZSTD
MEAN	58.7	20.0	.00	.24	1.00	.00	.98	-.05
SEM	1.0	.0	.06	.00	.10	.35	.09	.33
P.SD	3.0	.0	.17	.00	.29	1.04	.28	.99
S.SD	3.1	.0	.18	.00	.30	1.10	.29	1.04
MAX.	63.0	20.0	.33	.24	1.40	1.32	1.34	1.14
MIN.	53.0	20.0	-.25	.24	.60	-1.53	.60	-1.50
REAL RMSE	.26	TRUE SD	.00	SEPARATION	.00	ITEM RELIABILITY	.00	
MODEL RMSE	.24	TRUE SD	.00	SEPARATION	.00	ITEM RELIABILITY	.00	
S.E. OF ITEM MEAN = .06								
ITEM RAW SCORE-TO-MEASURE CORRELATION = -1.00								
Global statistics: please see Table 44.								
UMEAN=.0000 USCALE=1.0000								

Komponen-komponen yang tampak pada Tabel-1 di atas adalah seperti berikut. 1. Menunjukkan banyak data peserta (*person*) dan butir (*item*) yang diolah. 2. Nilai rata-rata logit untuk peserta.. 3. Pengelompokan peserta. 4. Nilai reliabilitas peserta untuk memberi petunjuk bahwa tes yang diberikan

dapat mengelompokkan peserta menjadi beberapa tingkatan sesuai dengan tujuan pemberian tes. 5. Nilai *Alpha Cronbach*. 6. Standar kesalahan (*error*) butir. 7. Pengelompokan butir. 8. Nilai reliabilitas butir tes.

Tabel-2: Ukuran Butir (*Item Measure*)

TABLE 13.1 data politomi.xlsx													
ZOU638WS.TXT Jul 9 2021 9: 3													
INPUT: 20 PERSON 10 ITEM REPORTED: 20 PERSON 10 ITEM 5 CATS MINISTEP 4.8.2.0													

PERSON: REAL SEP.: 2.07 REL.: .81 ... ITEM: REAL SEP.: .00 REL.: .00													
ITEM STATISTICS: MEASURE ORDER													

ENTRY	TOTAL	TOTAL		MODEL	INFIT		OUTFIT		PTMEASUR-AL	EXACT MATCH			
NUMBER	SCORE	COUNT	MEASURE	S.E.	MNSQ	ZSTD	MNSQ	ZSTD	CORR.	EXP.	OBS%	EXP%	ITEM

10	53	20	.33	.24	.68	-1.15	.63	-1.33	.76	.62	60.0	39.3	10
9	56	20	.15	.24	1.10	.43	1.06	.31	.65	.62	30.0	38.5	9
5	57	20	.10	.24	1.40	1.32	1.34	1.14	.64	.63	35.0	39.1	5
7	57	20	.10	.24	1.29	1.00	1.26	.90	.56	.63	25.0	39.1	7
8	58	20	.04	.24	.60	-1.53	.60	-1.50	.83	.63	55.0	38.7	8
6	59	20	-.02	.24	.67	-1.19	.64	-1.30	.74	.63	45.0	37.7	6
1	60	20	-.07	.24	1.14	.56	1.14	.56	.44	.63	40.0	37.8	1
2	62	20	-.19	.24	1.27	.95	1.30	1.02	.50	.63	35.0	38.6	2
3	62	20	-.19	.24	.70	-1.05	.79	-.66	.60	.63	35.0	38.6	3
4	63	20	-.25	.24	1.18	.68	1.07	.34	.58	.64	35.0	38.9	4

MEAN	58.7	20.0	.00	.24	1.00	.0	.98	-.1			39.5	38.6	
P.SD	3.0	.0	.17	.00	.29	1.0	.28	1.0			10.4	.5	

Tabel-3: Ukuran Peserta (*Person Measure*)

PERSON: REAL SEP.: 2.07 REL.: .81 ... ITEM: REAL SEP.: .00 REL.: .00

PERSON STATISTICS: MEASURE ORDER

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S.E.	INFIT		OUTFIT		PTMEASUR-AL		EXACT MATCH		PERSON
					MNSQ	ZSTD	MNSQ	ZSTD	CORR.	EXP.	OBS%	EXP%	
9	42	10	1.38	.41	1.00	.16	1.02	.20	-.15	.13	40.0	45.0	9
19	39	10	.93	.36	.45	-1.39	.47	-1.32	-.27	.15	70.0	46.2	19
4	38	10	.81	.35	1.12	.42	1.16	.50	-.56	.15	50.0	45.4	4
15	38	10	.81	.35	.69	-.66	.70	-.62	.11	.15	40.0	45.4	15
1	36	10	.57	.33	1.11	.39	1.11	.40	.19	.16	50.0	40.9	1
11	35	10	.46	.33	.49	-1.46	.49	-1.45	.07	.16	50.0	36.9	11
20	34	10	.36	.32	.96	.02	.95	-.01	.40	.17	40.0	34.5	20
5	33	10	.26	.32	1.08	.32	1.07	.30	.43	.17	20.0	32.2	5
18	32	10	.16	.31	1.58	1.43	1.57	1.42	-.04	.17	30.0	31.6	18
12	30	10	-.03	.31	1.11	.40	1.11	.40	.66	.17	20.0	32.2	12
16	30	10	-.03	.31	1.43	1.15	1.44	1.17	-.14	.17	20.0	32.2	16
6	28	10	-.23	.31	.66	-.92	.66	-.94	.32	.17	60.0	32.6	6
17	27	10	-.32	.31	1.02	.18	1.01	.16	.49	.17	40.0	32.2	17
13	26	10	-.42	.31	.62	-1.06	.62	-1.07	.16	.17	50.0	34.4	13
8	25	10	-.52	.32	1.38	1.02	1.38	1.03	-.23	.17	10.0	35.4	8
7	23	10	-.73	.33	1.25	.72	1.25	.70	.24	.16	20.0	39.9	7
2	19	10	-1.22	.38	.68	-.64	.69	-.61	.18	.14	50.0	42.4	2
3	19	10	-1.22	.38	1.49	1.08	1.47	1.04	.18	.14	20.0	42.4	3
10	18	10	-1.37	.39	1.04	.25	.98	.13	.52	.13	50.0	41.5	10
14	15	10	-1.93	.48	.51	-.89	.53	-.86	.44	.11	60.0	49.6	14
MEAN	29.3	10.0	-.12	.35	.98	.0	.98	.0			39.5	38.6	
P.SD	7.6	.0	.85	.04	.34	.9	.33	.8			16.3	5.7	

Tabel-1 yang memuat *item measure* menyajikan informasi tentang butir, yakni tingkat kesukaran atau kesulitan tiap-tiap butir tes. Pada kolom *Measure* disajikan tingkat kesukaran butir. Tampak pada tabel bahwa butir 10 dengan tingkat kesukaran +0,33 merupakan tes yang paling sukar dijawab oleh peserta tes. Sebaliknya, butir 4 dengan tingkat kesukaran -0,25 merupakan butir yang paling mudah. Artinya, sesuai asumsi teori respons butir, semua butir layak untuk dianalisis lebih lanjut.

Tabel *person measure* menampilkan informasi mengenai kemampuan peserta tes

dalam menyelesaikan setiap butir tes yang diberikan. Pada kolom *Measure* tertera nilai *logit person*, untuk peserta tes nomor 9 dengan kemampuan +1,38 logit menjadi peserta tes yang punya kecerdasan tertinggi dibandingkan dengan yang lainnya. Di lain sisi, peserta tes nomor 14 dengan kemampuan -1,93 logit menjadi peserta yang memiliki kemampuan paling rendah. Secara umum, sesuai batasan teori respons butir, semua peserta layak untuk dilibatkan sebagai responden dalam pengujian tes karena memiliki kemampuan di antara -3 sampai +3 logit.

SIMPULAN

Sekolah menengah kejuruan (SMK) diprogramkan untuk menghasilkan tenaga terampil di berbagai bidang sesuai dengan kebutuhan pembangunan. Keterampilan menjadi kompetensi yang lebih ditekankan daripada kompetensi di bidang teori. Oleh karena itu pembelajaran di SMK

didominasi kegiatan praktikum, baik di laboratorium, di bengkel kerja, maupun di dunia usaha dan industri (DUDI). Alat evaluasi pembelajaran yang ditekankan adalah tes kinerja yang mampu menghasilkan informasi tentang kompetensi peserta didik yang otentik.

Pengabdian kepada masyarakat ini menyoar peningkatan kemampuan guru SMK untuk menganalisis tes kinerja yang masih menjadi selama ini. Pengabdian dilakukan di SMK Negeri 2 Tabanan dengan pendekatan yang dilanjutkan dengan pendampingan, monitoring, dan evaluasi.

Program pengabdian memang belum tuntas dilakukan. Kendati demikian, sampai dengan 70% kegiatan berlangsung sudah diperoleh hasil yang cukup membanggakan. Peserta sudah mampu menganalisis tes kinerja menggunakan *Partial Credit Model (PCM)*. Selanjutnya, peserta juga sudah mampu mengoperasikan aplikasi Winsteps untuk

mendukung analisis dengan PCM yang melibatkan perhitungan-perhitungan yang rumit. Peserta sudah mampu impor data dari *Microsoft Excel* dan selanjutnya menganalisis dengan aplikasi Winsteps. Dalam analisis, peserta sudah mampu menampilkan semua output dari Winsteps. Lebih dari itu, peserta juga sudah mampu menginterpretasikan output dari Winsteps, sehingga mampu mengambil kesimpulan tentang kualitas tes. Pada sisa 30% kegiatan pengabdian peserta diharapkan mampu memperdalam penguasaan terhadap PCM dan Winsteps, sehingga kompetensi untuk menganalisis tes kinerja yang diperoleh lebih optimal.

DAFTAR RUJUKAN

- Agrawal, Tushar. Vocational education and training programs (VET): An Asian perspective. (2014). *Asia-Pacific Journal of Cooperative Education*, 2013 14(1), 15-26.
- Brown, J. D. (2004). Performance assessment: Existing literature and directions for research. *Second Language Studies*, 22 (2), 91–139.
- Callahan, Joseph F. and Leonard H. Clark. (1977). *Teaching in the Middle and Secondary Schools*. New York: Macmillan Publishing Co., Inc.
- Cedefop (2017). *The changing nature and role of vocational education and training in Europe. Volume 1: conceptions of vocational education and training: an analytical framework*. Luxembourg: Publications Office. Cedefop research paper; No 63. <http://dx.doi.org/10.2801/532605>
- Eggert, Sabina & Susanne Bo Geholz. (2009). Students' Use of Decision-Making Strategies With Regard to Socioscientific Issues: An Application of the Rasch Partial Credit Model, *Published online 27 July 2009 in Wiley InterScience* (www.interscience.wiley.com), DOI 10.1002/sce.20358.
- Goel, V. P. (2009). *Technical and vocational education and training (TVET) system in India for Sustainable Development*. New Delhi: Ministry of HRD, Government of India.
- Hambleton R.K., Swaminathan H., and Rogers H.J. (1991). *Fundamental of Item Response Theory*. Sage Publication Inc., Newbury Park, CA.
- Heritage, Margaret. (2010). Formative Assessment and Next-Generation Assessment Systems: Are We Losing an Opportunity? *Paper prepared for the Council of Chief State School Officers*, September 2010.
- Linn, RL, Baker, EL, Dunbar, SB Complex. (1991). Performance-based assessment: Expectations and Validation Criteria,

- Educational Researcher, November 1991, 20(8): 15-21
- Master, Geofferey N. (1988). The Analysis of Partial Credit Scoring. *Applied Measurement in Education*, 1(4), 279-297.
- Modie, Gavin. (2006). Identifying Vocational Education and Training. *Journal of Vocational Education and Training*. December 2006 DOI: 10.1080/13636820200200197. Source: OAI.
- Newaz, Mohammad Tanvi, Faruquee, Murtaza, Farha, Sadia. (2013). Vocational education and training in Bangladesh: Why it is not working? *International Journal of Research Studies in Education*. 2013 October, Volume 2 Number 4, 29-40.
- Oliver, Damian. (2013). Complexity in Vocational Education and Training Governance. *Research in Comparative and International Education, Volume 5 Number 3 2010*.
- Palm, Torulf. (2008). "Performance Assessment and Authentic Assessment: A Conceptual Analysis of the Literature," *Practical Assessment, Research, and Evaluation*: Vol. 13, Article 4.
- Serafini, F. (2010). *Classroom reading assessments: More efficient ways to view and evaluate your readers*. Portsmouth, NH: Heinemann.
- Verhelst, N.D. and H.H.F.M. Verstralen. (2008). Some Considerations on the Partial Credit Model, *Psicologica*, 2008, 29, 229-254.