

Analisis Model Rasch Menggunakan Ministep Seri-4.8.2. untuk Mengukur Instrumen Penilaian Kinerja Kepala Sekolah (PKKS) Ver. 5.00 Pada SMK Kabupaten Sumenep

Yuni Hartatik, Nunuk Hariyati, Murtadlo

© 2022 JEMS (Jurnal Edukasi Matematika dan Sains)

This is an open access article under the CC-BY-SA license (<https://creativecommons.org/licenses/by-sa/4.0/>) ISSN 2337-9049 (print), ISSN 2502-4671 (online)

Abstrak:

Penilaian Kinerja Kepala Sekolah (PPKS) adalah bentuk evaluasi kualitas pelaksanaan tugas pokok dan fungsi kepala sekolah dalam melaksanakan fungsi-fungsi manajerial dan supervisi/pengawasan pada lembaga yang dipimpinnya. PPKS dilakukan oleh Pengawas Pembina yang telah ditunjuk oleh satuan kerja (Dinas Pendidikan). Kualitas PPKS tergantung kualitas instrumen yang digunakan, sehingga memberikan hasil maksimal. Penelitian ini adalah penelitian survei bertujuan mendeskripsikan kualitas Instrumen PPKS yang dikembangkan oleh Tim Korwas Dinas Pendidikan Provinsi Jawa Timur tahun 2020, menggunakan teknik analisis model Rasch. Data penelitian adalah data sekunder yang diperoleh dari tim Pengawas Dinas Pendidikan Provinsi Jawa Timur Wilayah Kabupaten Sumenep. Hasil analisis menunjukkan Instrumen berkualitas layak, baik reliabilitas, validitas, dan obyektivitas. Namun terdapat beberapa item Sasaran Kerja yang memiliki pilihan (distraktor) yang membingungkan, disebabkan karena rubrik kurang detail, dan penilai hanya satu asesor.

Kata kunci: PPKS, Model Rasch, SMK

Abstract:

Principal Performance Assessment (PPKS) is a format that assesses the quality of implementation of a client's key tasks and features when performing management and oversight / supervision functions in a client-managed institution. PPKS is carried out by supervisors appointed by the unit of work (office education). The quality of PPKS depends on the quality of the equipment used for maximum results. This study is a research study aimed at explaining the quality of PPKS equipment developed by the Korwas team of the East Java Department of Education in 2020 using the Rasch model analysis method. The survey data is secondary data obtained from the supervisory team of the East Java State Department of Education, Sumenep Regency. The results of the analysis show that the equipment is of decent quality in terms of reliability, validity and objectivity. However, there are some work goals (distracting) that can be confusing to choose because the rubrics are not well refined and there is only one evaluator.

Keywords: PPKS, Rasch Model, SMK

Pendahuluan

Menurut Undang Undang RI No. 20 Tahun 2003, dalam rangka menjalankan fungsinya sebagai lembaga pendidikan sekolah harus memenuhi kriteria standar minimal, meliputi standar: isi, proses, kompetensi lulusan, tenaga pendidik dan kependidikan, sarana dan prasarana, pengelolaan, pembiayaan, dan penilaian pendidikan. Pencapaian standarisasi di atas tersebut adalah mutlak, dalam rangka kepastian penjaminan layanan serta pengendalian mutu pendidikan yang dilakukan sekolah.

Yuni Hartatik, Universitas Negeri Surabaya
yuni.20019@mhs.unesa.ac.id

Nunuk Hariyati, Universitas Negeri Surabaya
nunukhariyati@unesa.ac.id

Murtadlo, Universitas Negeri Surabaya
murtadlo@unesa.ac.id

Standarisasi pengelolaan (manajemen) lembaga pendidikan merupakan inti dari strategi pencapaian standarisasi pendidikan, sebab di dalamnya memuat: (1) teknik pengorganisasian, seperti menyusun perencanaan, membangun organisasi dan teknik pengorganisasiannya; (2) teknik pergerakan, yaitu mengoptimalkan kinerja lembaga; (3) teknik pengendalian atau pengawasan, evaluasi, serta (4) restrukturisasi atau pengembangan lembaga untuk menjadi lebih baik (Haq, 2017). Selain itu, dalam Peraturan Menteri Pendidikan Nasional RI Nomor 19 Tahun 2007 tentang Standar Pengelolaan Pendidikan oleh Satuan Pendidikan Dasar dan Menengah, pengelolaan sekolah yang baik juga harus memiliki: (1) Perencanaan Program, (2) Pelaksanaan Program, (3) Pengawasan dan Evaluasi, Kepemimpinan Sekolah, serta (5) Sistem Informasi dan Manajemen.

Semua standar pengelolaan di atas membutuhkan kepemimpinan lembaga (Kepala Sekolah) yang kompeten, terutama pada: (1) kompetensi manajerial, yaitu perencanaan, pengelolaan, monitoring dan mengevaluasi, serta memanfaatkan teknologi informasi dalam pengambilan keputusan, (2) kompetensi supervisi, yaitu dalam rangka menjamin keterlaksanaan program serta memberikan tindak lanjutnya, (3) kompetensi kewirausahaan dalam rangka berinovasi dalam mengembangkan sekolah, (4) kompetensi kepribadian, baik secara diri maupun kedinasan sehingga menjadi teladan bagi orang lain, serta (5) kompetensi sosial, bekerja sama dengan semua pihak dalam rangka mengembangkan lembaga (Permendiknas RI Nomor 13 Tahun 2007 tentang Standar Kepala Sekolah/Madrasah).

Data BPK tahun 2019 (<https://jatim.bpk.go.id>) menyebutkan, Indeks Pembangunan Manusia (IPM) di semua wilayah Madura (termasuk Kabupaten Sumenep) masuk kategori merah. Khusus Kabupaten Sumenep menduduki peringkat ke-32 dari 38 Kabupaten/Kota se-Provinsi Jawa Timur (<https://sumenepkab.bps.go.id>) dengan nilai 66.22%, Indeks Pendidikan hanya 65.15%, serta kecenderungan menurun dari tiga tahun terakhir sebelumnya. Fenomena tersebut sangat menarik, sebab berbagai program telah dicanangkan oleh Pemerintah (Pusat dan Provinsi) untuk mendongkrak kualitas pendidikan, misalnya dengan: (1) pengembangan program *double track* pada SMA di Jawa Timur, (2) peningkatan dana BOS, (3) pemberian dana BPOPP, (4) peningkatan kesejahteraan GTT/PTT, (5) penyelenggaraan sekolah berbasis pesantren, dan sebagainya. Faktanya berbanding terbalik terhadap nilai Indeks Pembangunan Manusia.

Karakteristik unik Kabupaten Sumenep (data BPS Kab. Sumenep 2021) yang memiliki wilayah geografis yang sangat luas (2093.59 km²), serta memiliki 126 pulau dengan rata-rata jarak ke pemerintah pusat kabupaten 128 km, merupakan kesulitan tersendiri dalam membangun kualitas SDM berstandar nasional. Untuk itu, Pemerintah Daerah membuka akses pendirian lembaga pendidikan yang dikelola masyarakat (Swasta). Data Dinas Pendidikan Provinsi Jawa Timur Wilayah Kabupaten Sumenep (2021) menyebutkan, terdapat total 147 buah SMA/SMK, meliputi 84 buah SMA (12 SMA berstatus negeri dan 72 SMA berstatus swasta), serta 66 buah SMK (3 SMK berstatus negeri dan 63 SMK berstatus swasta). Sementara itu, berdasar jumlah siswa, terdapat 6 862 siswa SMA Negeri; 8 246 siswa SMA Swasta; 2 777 siswa SMK negeri; dan 4 567 siswa SMK swasta.

Khusus SMK, jumlah SMK berstatus negeri hanya 4.3% dari total jumlah SMK se-Kabupaten Sumenep. Sementara jumlah siswa SMK berstatus negeri 37.8% dibandingkan dengan total jumlah seluruh siswa SMA/SMK. Artinya, jumlah total rata-rata siswa di SMK berstatus swasta hanya 72.5 siswa, atau jika menggunakan klasifikasi jumlah rombongan belajar (1 rombel berisi ± 24 siswa), maka SMA/SMK swasta hanya memiliki satu rombongan belajar per-jenjang kelas untuk semua Program Kompetensi Keahlian (Prokomli). Fakta ini

menunjukkan ada rentang yang cukup jauh antara kualitas pendidikan SMA/SMK berstatus negeri dan swasta, terutama dalam standar pengelolaan (baik pengelolaan sumber daya manusia dan material), sehingga dimungkinkan menjadi penyebab rendahnya Indeks Pembangunan Manusia dan Pendidikan di Kabupaten Sumenep.

Kepala Sekolah sebagai *top manager* pada tingkat satuan kerja memiliki tantangan tugas luar biasa (terutama SMK berstatus swasta), di mana ujung dari tugas pokok di atas adalah Kepala sekolah dikatakan berhasil jika mereka mampu membawa lembaga yang dipimpinnya mencapai Standar Nasional Pendidikan. Agar Kepala Sekolah senantiasa menjalankan tugas pokok dan fungsinya, diperlukan kontroling melalui Penilaian Kinerja Kepala Sekolah (PKKS). Tujuan PKKS adalah: (1) Memperoleh data tentang pelaksanaan tugas pokok, fungsi dan tanggung jawab kepala sekolah dalam melaksanakan fungsi-fungsi manajerial dan supervisi/pengawasan pada sekolah yang dipimpinnya; (2) Memperoleh data hasil pelaksanaan tugas dan tanggung jawabnya sebagai pemimpin sekolah; (3) Menentukan kualitas kerja kepala sekolah sebagai dasar dalam promosi dan penghargaan yang diberikan kepadanya; (4) Menentukan program peningkatan kemampuan profesional kepala sekolah dalam konteks peningkatan mutu pendidikan pada sekolah yang dipimpinnya; serta (5) Menentukan program umpan balik bagi peningkatan dan pengembangan diri dan karyanya dalam konteks pengembangan karir dan profesinya (Direktorat Tenaga Kependidikan Direktorat Jenderal Peningkatan Mutu Pendidik dan Tenaga Kependidikan Departemen Pendidikan Nasional, 2008).

Pelaksanaan PKKS SMK tersebut dilakukan oleh Pengawas Pembina (dilakukan 2 orang assesor) dari Cabang Dinas Propinsi (menurut satuan kerja masing-masing), yang dilaksanakan berkala setahun sekali (dalam penelitian adalah penilaian akhir tahun) dan secara kumulatif dilakukan setiap empat tahun, serta hasil penilaian kinerja dikategorikan dalam klasifikasi amat baik, baik, cukup, sedang atau kurang. Penilaian ini menggunakan Instrumen yang mengacu pada Buku Pedoman Penilaian Kepala Sekolah/Madrasah (Pengembangan Sumber Daya Manusia Pendidikan dan Kebudayaan dan Badan Penjaminan Mutu Pendidikan Pusat Pengembangan Tenaga Kependidikan, Kementerian Pendidikan dan Kebudayaan RI, 2012), dengan memegang teguh prinsip: (1) sahih/validitas; (2) Objektif/relibilitas; (3) Adil/tidak bias, (4) Terpadu, (5) Terbuka, (6) Menyeluruh dan berkesinambungan, (7) Sistematis, (8) Beracuan pada kriteria, dan (9) Akuntabel.

Pelaksanaan di lapangan, PPKS memfokuskan dimensi pengukuran pada aspek: (1) **Perilaku kerja**, meliputi: orientasi pelayanan, integritas, komitmen, disiplin, dan kerjasama; (2) **Sasaran kerja**, meliputi: manajerial, kewirausahaan, supervisi, tugas tambahan, Penilaian Kinerja Berkelanjutan/PKB dan kegiatan penunjang ; serta (3) **Kehadiran**. Agar memberikan hasil yang obyektif, diperlukan alat ukur yang modern dan terpercaya, sehingga hasil PKKS dapat digunakan sebagai dasar merumuskan rekomendasi tindak lanjut pembinaan, dan sebagai dasar pengembangan kebijakan perbaikan mutu (Permendiknas RI No. 13 Tahun 2007).

Penelitian ini akan menganalisis tentang kinerja kepala SMA/SMK dalam mengelola lembaga pendidikan mereka, menggunakan teknik analisis Rasch Model. Rasch Model, merupakan teknik pemodelan (statistik) yang menggunakan data mentah untuk menghasilkan skala pengukuran interval yang sama (Sumintono & Widhiarso, 2015). Model Rasch merupakan teori evaluasi modern yang dapat mengklasifikasikan perhitungan barang dan orang dalam suatu peta distribusi (Sumintono & Widhiarso, 2014). Pendekatan model Rasch selain memperhatikan item dan person, juga memperhatikan aspek respon dan korelasi (Ardiyanti,

2016), sehingga diharapkan melalui analisis Rasch, akan didapat hasil pengukuran yang lebih akurat dan memberikan informasi yang tepat mengenai instrumen pengukuran serta kinerja (Sumintono & Widhiarso, 2015) Kepala Sekolah. Pada akhirnya, diskusi penelitian akan didapatkan gambaran tentang penyebab penurunan kualitas indeks pembangunan manusia yang disebabkan kualitas pengelolaan lembaga pendidikan (khususnya SMA/SMK), serta memberikan peta solusi alternatif menyelesaikan permasalahan pendidikan di kabupaten Sumenep, serta pendidikan di Indonesia yang memiliki kasus yang relatif sama.

Model Rasch adalah merupakan pemodelan (statistik) modern yang menggunakan data mentah untuk menghasilkan skala pengukuran pada interval yang sama (Sumintono & Widhiarso, 2015), dapat digunakan untuk mengklasifikasikan pengukuran barang dan orang dalam suatu peta distribusi (Sumintono & Widhiarso, 2014), artinya menjelaskan keterkaitan antara responden dan instrumen, sehingga lebih menjelaskan apa yang terjadi antara responden dan aksi yang dilakukan. Salah satu keistimewaan pengukuran model Rasch adalah tidak tergantung pada sampel yang digunakan (Kubinger, K. D., Rasch, D., & Yanagida, T., 2011), karena pada model Rasch, data yang digunakan adalah skor peluang (P), yaitu perbandingan antara jawaban benar dan jumlah soal yang diberikan (*Odds-Ratio*). Ini berarti, SMK paling bermutu memiliki peluang nilai paling tinggi dibandingkan SMK kurang bermutu. Sebaran data tersebut dapat dilihat pada Peta Wright, yang menunjukkan probabilitas SMK untuk meningkatkan dirinya. Melalui matriks Guttman, dapat diperoleh pola jawaban responden (pengawas penilai) terhadap butir. Munculnya inkonsistensi jawaban responden (*misfit*) atau pola yang tidak umum (*outlier*) dapat terdeteksi. Menurut Lord (dalam Nurcahyo, 2016), model Rasch mengasumsikan bahwa semua butir didiskriminasikan secara sama, sehingga butir tidak dapat dijawab dengan benar berdasarkan tebakan (asal menjawab).

Metode

Jenis penelitian ini adalah non-eksperimen model penelitian survei pendekatan deskriptif kuantitatif. Data yang digunakan adalah data sekunder (karena keterbatasan demografi dan lokasi sekolah yang jauh), diperoleh dari data Pengawas Pembina SMA/SMK pada Dinas Pendidikan Provinsi Jawa Timur Kantor Cabang Wilayah Kabupaten Sumenep (selanjutnya disebut sebagai responden atau assesor) tahun 2021, menggunakan aplikasi **PKKS Ver. 5.00** yang dikembangkan oleh Tim Korwas (Koordinator Pengawas) Prov. Jatim 2020.

Tabel-1. Rentang skor dan sebutan Penilaian Kinerja Kepala Sekolah

Rentang Skor	Sebutan
00.00 s.d. 20.00	Sangat kurang
20.01 s.d. 40.00	Kurang
40.01 s.d. 60.00	Cukup
60.01 s.d. 80.00	Baik
80.01 s.d. 100.0	Sangat Baik

Sumber: Tim Korwas Prov. Jatim, 2020

Tabel-2. Rincian aspek, indikator, dan kode butir Penilaian Kinerja Kepala Sekolah

Unsur Penilaian	Rentang skor	Indikator Penilaian Kinerja	Kode Butir
Perilaku kerja (PK)	0 (Tidak memenuhi) 1 (Baik) 2 (Amat baik)	Orientasi Pelayanan	PK1
		Integritas.	PK2
		Komitmen	PK3
		Disiplin	PK4
		Kerjasama	PK5
Sasaran kerja (SK)	0 (Sangat kurang) 1 (kurang) 2 (cukup) 3 (baik) 4 (Sangat baik)	Manajerial	SK1
		Kewirausahaan	SK2
		Supervisi	SK3
		Tugas Tambahan	SK4
		PKB	SK5
		Kegiatan Penunjang	SK6

Sumber: Tim Korwas Prov. Jatim, 2020

Analisis data penelitian momfokuskan pada dua aspek PKKS, yaitu: (1) **Perilaku Kerja/PK**, diuraikan menjadi 5 indikator kinerja, di mana penilaian pada Instrumen menggunakan skor: 0 (Tidak memenuhi), 1 (Baik), dan 2 (Amat baik), serta (2) **Sasaran Kerja/SK**, diuraikan menjadi 6 kinerja, dimana menggunakan skor: 0 (sangat kurang), 1 (kurang), 2 (cukup), 3 (baik), dan 4 (Sangat baik), dengan asumsi reliabilitas kesepakatan penilaian antarpengawas atau responden memenuhi. Skor di atas selanjutnya dikonversi dalam data interval dengan nilai 0 s.d. 100. Kriteria penilaian dapat dilihat pada Tabel-1, sementara aspek kinerja, indikator kinerja dan pengkodean kinerja dapat dilihat pada Tabel-2 .

Obyek penelitian adalah 30 SMK Negeri/Swasta, dengan asumsi jumlah sampel yang layak untuk model Rasch pada tingkat kepercayaan 95% (Sumintono dan Widhiarso, 2013), diambil secara acak berdasarkan hasil penilaian assesor tahun 2021. Rincian SMK dan pengkodean dapat dilihat pada Tabel-3.

Tabel-3. Kode nama SMK di Kabupaten Sumenep dalam survei

Nama SMK	Status	Akre	Lokasi	Kode SMK
SMKN 1 Kalianget	Negeri	A	Kertasada, Kalianget	KAN
SMKN 1 Sumenep	Negeri	A	Pabian, Batuan	SUN
SMKS Al Azhar	Swasta	B	Moncek Tengah, Lenteng	AZS
SMKS Al Azis	Swasta	B	Aeng Dake, Bluto	ZHS
SMKS Al-Amien Putri	Swasta	B	Prenduan, Pragaan	MPS
SMKS Al-Amien Tegal	Swasta	C	Prenduan, Pragaan	ALS
SMKS Al-Bukhari	Swasta	C	Jungkat, Raas	BHS
SMKS Al-Imron	Swasta	B	Pakamban, Pragaan	IMS
SMKS Al-Ittihad Lenteng	Swasta	B	Lembung Timur, Lenteng	ITS
SMKS Al-Karimiyah	Swasta	C	Braji, Gapura	KRS
SMKS Al-Marsufie	Swasta	C	Kapedi, Bluto	MSS
SMKS Al-Ittihad Ra as	Swasta	C	Ketupat, Ra as	ITR

Nama SMK	Status	Akre	Lokasi	Kode SMK
SMKS Anwarul Ma'arif	Swasta	B	Tambak Agung Barat, Abunten	AMS
SMKS At-Taufiqiyah	Swasta	B	Aeng Baja, Bluto	ATS
SMKS Darul Ulum (A)	Swasta	C	Dasuk, Dasuk	DAS
SMKS Darul Ulum (B)	Swasta	-	Bumbungan, Bluto	DUS
SMKS Al Ghafur	Swasta	B	Karang Sokon, Guluk Guluk	AGF
SMKS Isl Terp Miftahul Ulum	Swasta	B	Kapedi, Bluto	MLS
SMKS Islam Darus Salam	Swasta	B	Bilapora Temor, Ganding	DSS
SMKS Mambaul Ulum	Swasta	C	Guluk Guluk, Guluk Guluk	MUS
SMKS Nurud Dhalam	Swasta	C	Pasongsongan, Pasongsongan	NDS
SMKS Nurul Huda	Swasta	C	Moncek Barat, Lenteng	NUS
SMKS Nurul Islam	Swasta	C	Karang Cempaka, Bluto	NIS
SMKS Nurul Muclisin	Swasta	C	Pakondang, Rubaru	NMS
SMKS Sidingpuri	Swasta	B	Poreh, Lenteng	SPS
SMKS Al Itqan	Swasta	B	Larangan, Ganding	AIT
SMKS Sumber Mas	Swasta	C	Rombiya Barat, Ganding	SMS
SMKS Al Fata	Swasta	C	Bilapora Rebba, Lenteng	AFT
SMKS Bustanul Ulum	Swasta	C	Sonok, Nonggunong	BSU
SMKS Raudatul Iman	Swasta	C	Gadu Barat, Ganding	RDI

Analisis data menggunakan model Rasch, didukung *software* **Ministep Seri-4.8.2**. Realiabilitas SMK dan butir di dapat dari nilai logit yang diuji, yaitu *person*/SMK dan *item*/butir. *Person Reliability* menunjukkan konsistensi penilaian responden terhadap SMK, serta *Item Reliability* menunjukkan konsistensi butir penilaian. Nilai reliabilitas *person* dan *item* yang diterima adalah > 0.67 . Sementara reliabilitas interaksi antara *person* dan *item* secara keseluruhan digunakan nilai *Alhpa Cronbach*, di mana diterima jika > 0.6 .

Validitas butir dapat dilihat dari nilai *item fit order* (kesesuaian atau ketepatan butir), terdiri dari: *Outfit Mean Square* (MNSQ) dan *Outfit Z-Standart* (ZSTD). Butir *fit* (sesuai) bermakna butir tersebut secara konsisten dengan yang diharapkan pada model, artinya tidak ada miskonsepsi responden terhadap butir (Akhtar, 2017). Nilai *Outfit* MNSQ yang diterima adalah antara 0.5 s.d. 1.5, artinya nilai *Outfit* MNSQ yang *fit* jika nilainya mendekati 1. Sementara itu, *Outfit* ZSTD yang diterima adalah antara (-2.0) s.d. 2.0, artinya nilai *Outfit* ZSTD yang *fit* jika nilainya mendekati 0 (Sumintono & Widhiarso, 2014).

Daya diskrimasi diukur melalui *Pt. Mean Corr*, di mana nilai *Pt. Mean Corr* butir diterima antara 0.4 s.d. 0.85 (Sumintono & Widhiarso, 2014), Jika *Pt. Measure Corr* bernilai positif, maka SMK yang kualitas tinggi akan memiliki probabilitas memperoleh nilai butir tinggi, atau SMK yang kualitas rendah akan memiliki probabilitas memperoleh nilai butir rendah pula. Namun, jika nilai *Pt. Measure Corr* bernilai negatif, terindikasi SMK berkualitas rendah malah memperoleh nilai tinggi, atau SMK berkualitas tinggi malah memperoleh nilai rendah. Butir yang tidak *fit* ini dimungkinkan adanya: (1) penilaian oleh responden yang asal-asalan, dan atau (2) butir memiliki daya pembeda rendah.

Untuk menguji butir tidak mengandung unsur bias (sehingga tidak ada SMK dan atau butir yang lebih diuntungkan dibandingkan yang lain), yang mempengaruhi akurasi pengukuran, dapat dilihat melalui nilai *Differential Person Functioning (DPF, Between/Within)* dan *Differential Item Functioning/DIF (DPF, Between/Within)* dengan nilai probabilitas butir yang tidak boleh kurang dari 5% (Sumintono & Widhiarso, 2015). Jika SMK yang memiliki nilai *DPF (p < 0,05)* bisa mengindikasikan kecurangan responden (assessor) saat mengisi instrumen (Sumintono & Widhiarso, 2014).

Analisis validitas skala penilaian yang bertujuan untuk memverifikasi pilihan penilaian (jawaban) bahwa butir penilaian tersebut membingungkan atau tidak bagi responden, dapat dilihat pada nilai logit *Rating (Partial Credit) Scale*. Jika nilai logit semakin naik dari yang skala rendah ke skala tinggi secara positif, maka pilihan jawaban tidak membingungkan.

Indeks separasi (kualitas) SMK dan separasi butir menunjukkan peta strata SMK dan butir, serta juga dapat menunjukkan konsistensi penilaian (SMK dan butir), artinya nilai indeks separasi dapat menunjukkan tingkat reliabilitas. Menurut Wright & Stone (1999), indeks separasi > 5 setara dengan reliabilitas di atas 0.90, sedangkan indeks separasi < 1 setara dengan reliabilitas di bawah 0.50. Peta pengelompokan menggunakan rumus:
$$H = \frac{4(\text{Nilai Separasi})}{3}$$
.

Telaah interaksi antara hasil penilaian SMK terhadap butir, bahwa hanya terdapat satu atribut laten atau tidak adanya residu yang mendasari para responden (assesor) dalam menjawab butir (Lord, 1980 dalam Ridho, 2011) dilakukan uji unidimensional, di mana untuk *raw variance measure* minimal mencapai 20%. Namun, untuk kasus *unexplained variance*, kualitas instrumen dikategorikan baik jika memiliki nilai kurang dari 15% (Sumintono & Widhiarso, 2014).

Sementara itu, tingkat kualitas SMK dan hasil penilaian dilihat dari nilai *Logit*, pada: (1) tingkat penilaian pada butir, (2) kesesuaian butir, dan (3) kualitas SMK. Tingkat kualitas penilaian butir (*item measure*) dihitung dari estimasi logit *odds-ratio*, di mana jika memiliki nilai semakin tinggi, maka dikatakan butir semakin tinggi, dan sebaliknya. Tingkat penilaian butir juga dilihat dari *Wright Map* yang menampilkan grafik vertikal, menjelaskan peta distribusi SMK dan butir yang diuji pada sekala yang sama. Sedangkan untuk mengetahui tingkat abilitas (kemampuan) SMK dapat diukur dari nilai *Person Measure* (Sumintono & Widhiarso, 2015). Rangkuman kriteria uji kualitas instrumen (butir) tes dapat dilihat pada Tabel-4.

Tabel-4. Kriteria uji kualitas butir soal

Kualitas Butir Soal	Jenis Pengukuran	Keterangan
Reliabilitas (Responden/Butir)	<ul style="list-style-type: none"> · <i>Person Reliability</i> · <i>Item reability</i> 	n < 0.67 (Lemah)
		0.67 – 0.80 (Cukup)
		0,81 – 0,90 (Baik)
		0,91 – 0,94 (Sangat Baik)
		n > 0,94 (Istimewa)
Reliabilitas keseluruhan	Alpha Cronbach	n < 0.5 (Sangat buruk)
		0.5 - 0.6 (Buruk)
		0.6 – 0.7 (Cukup)

Kualitas Butir Soal	Jenis Pengukuran	Keterangan
		0.7 – 0.8 (Baik) n > 0.8 (Sangat Baik)
Separasi	<i>Separation Index</i>	n < 2 (Buruk) 2.0 – 3.0 (Cukup) 3.0 – 4.0 (Baik) 4.0 – 5.0 (Sangat baik) n > 5.0 (Istimewa)
Kesesuaian butir soal	<i>Items Fit Order</i>	Outfit MNSQ antara 0.5 s.d. 1.5 (sangat baik jika mendekati 1) Outfit ZSTD antara -2.0 s.d. +2.0 (sangat baik jika mendekati 1) Pt Mean Corr antara 0.4 s.d. 0.85 (positif)
Butir bias	<ul style="list-style-type: none"> • <i>Dif. Person Functioning (DPF)</i> • <i>Dif. Item Functioning (DIF)</i> 	Nilai probabilitas butir soal > 5% atau 0.05.
Uji Unidimensional butir soal	Nilai varians mentah (kasar) Varians yang tidak diketahui	< 20.0% (Jelek) 20% s.d. 39.9% (Cukup) 40% s.d. 59.9 (Baik) >60% (Istimewa) < 15% (baik)
Tingkat kesukaran butir	<i>Item Measure</i>	Measure logit > 1 (Sangat sulit) 0.5 < Measure logit < 1 (Sulit) -0.5 < Measure logit < 0.5 (Menengah) -0.5 < Measure logit < -1 (Mudah) Measure logit < -1 (Sangat mudah)
Tingkat kualitas SMK	<i>Person Measure</i>	Measure logit > 1 (Sangat Tinggi) 0.5 < Measure logit < 1 (Tinggi) -0.5 < Measure logit < 0.5 (Menengah) -0.5 < Measure logit < -1 (Rendah) Measure logit < -1 (Sangat Rendah)
Uji pilihan jawaban	<i>Rating (Partial Credit) Scale</i>	Logit naik secara positif
Peta variabel	<i>Variable Maps Items</i>	Pengelompokan butir penilaian rendah dapat dilihat dari nilai batas bawah yang diperoleh dari rata-rata <i>logit</i> butir pada tabel ukuran

Kualitas Butir Soal	Jenis Pengukuran	Keterangan
		butir. Kelompok butir penilaian berada di antara batas atas dan batas bawah. Pengelompokan butir penilaian rendah dapat dilihat dari nilai batas atas yang diperoleh dari logit rata-rata SMK dari tabel ukuran (SMK yang diuji).

Sumber : (Sumintono & Widhiarso, 2015)

Analisis lebih mendalam dapat dilihat pada tabel skalogram atau matriks Guttman (untuk mengetahui penyebab secara langsung mengapa pola responnya tidak sesuai). Sebagaimana dalam model Rasch, *measure* dilakukan secara simultan, mengurutkan butir penilaian dari paling tinggi sampai paling rendah, dan SMK yang kualitasnya paling tinggi ke paling rendah. Sehingga, muncul pola butir penilaian tingkat kualitas butir terhadap strata SMK. Melalui skalogram dan matriks Guttman, dapat diketahui assesor yang terindikasi curang, dan atau memiliki kemampuan yang sama (Sumintono & Widhiarso, 2014).

Hasil dan Pembahasan

Ringkasan Statistik menunjukkan, reliabilitas data SMK 0.76 (cukup), reliabilitas item penilaian 0.97 (istimewa), serta nilai Alpha Cronbach (KR-20) 0.75 (baik). Artinya konsistensi antara data SMK dan item penilaian, serta interaksi antara data keduanya telah memenuhi syarat. Sementara dari outfit MNSQ SMK dan butir penilaian masing-masing 1.07, dan outfit ZSTD SMK masing masing -0.4, artinya kesesuaian antara data SMK dan item (*fit person* dan *fit item*) telah memenuhi syarat (Tabel-5).

Tabel-5. Ringkasan uji reliabilitas dan validitas responden dan butir

Kriteria	Parameter (N)	
	Responden (49)	Butir (30)
Outfit MNSQ	1.07	1.07
Outfit ZSTD	-0.4	-0.4
Separation	1.77	6.01
Reliabilitas	0.76	0.97
Reliabilitas Alpha Cronbach	0.75	

Tabel-6. Peluang SMK yang memiliki penilaian asal asalan

Aspek Penilaian	No	Kode SMK	Nama SMK	Nilai DPF
Penilaian Kinerja (PK)	1	SUN	SMKN 1 Sumenep	-0.07
	2	IMS	SMKS Al-Imron	-0.04
	3	ITS	SMKS Al-Ittihad Lenteng	-0.03

	4	MLS	SMKS Isl Terp Miftahul Ulum	-0.01
	5	NDS	SMKS Nurud Dhalam	0.02
	6	NUS	SMKS Nurul Huda	0.03
Sasaran Kinerja (SK)	1	MSS	SMKS Al-Marsufie	0.02
	2	NUS	SMKS Nurul Huda	0.03

Pt. Mean Corr menunjukkan, semua bernilai positif dan namun terdapat butir yang nilainya di luar harapan ($0.4 \geq Pt. Corr \leq 0.85$), yaitu SK3 (Supervisi) yang perlu diberikan perhatian, sebab dimungkinkan penilaian oleh responden yang asal-asalan, dan atau butir memiliki daya pembeda rendah.

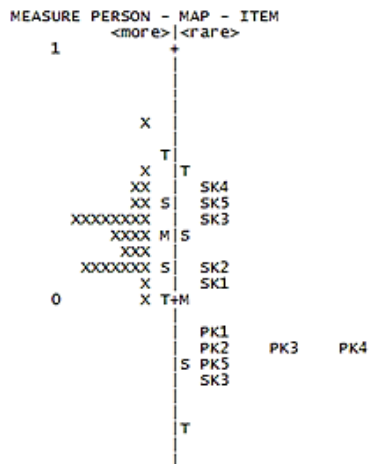
Nilai *Differential Person Functioning* (DPF) menunjukkan terdapat tujuh (23.3%) SMK yang memperoleh nilai tidak sesuai harapan (Tabel-6), bahkan terdapat satu (3.3%) SMK memiliki DPF jelek ($p < 0,05$), yaitu SMKS Nurul Huda. Selain itu, terdapat empat (13.3%) SMK yang memperoleh nilai DPF negatif, yaitu: (1) SMKN 1 Sumenep; (2) SMKS Al-Imron; (3) SMKS Al-Ittihad Lenteng; dan (4) SMKS Islam Terpadu Miftahul Ulum. Fenomena tersebut mengindikasikan kesalahan atau kecurangan responden (asesor) saat mengisi instrumen (Sumintono & Widhiarso, 2014).

Nilai logit *Rating (Partial Credit) Scale* menunjukkan ada dua aspek penilaian yang memiliki pilihan jawaban (distraktor) membingungkan assessor, terutama pada SK2 (Supervisi), dan PK5 (Kerja sama). Nilai Logit menunjukkan, distraktor SK3 (Supervisi) memunculkan miskonsepsi pada asesor antara memberikan nilai 4 (baik) dan 5 (sangat baik). Sementara PK5 (Kerja sama) terjadi 1 (baik) dan 2 (amat baik), namun sangat tipis, karena hasil konversi 3 distraktor menjadi 5 distraktor.

Tabel-7. Nilai logit distraktor item penilaian

Kode Butir	Aspek Penilaian	Keterangan
SK3	Supervisi	Distraktor-5 lebih rendah dari distraktor 4
PK5	Kerjasama	Distraktor-4 lebih rendah dari distraktor 3

Nilai indeks separasi *person* dan *item* menunjukkan, separasi SMK 1.77 (buruk), setara



dengan reliabilitas cukup. Sementara separasi *item* penilaian 6.01 (istimewa), atau setara dengan reliabilitas > 0.90 . Ini menunjukkan terdapat SMK tidak cukup kontinum antara SMK berkualitas tinggi dan rendah, atau tidak adanya nilai bervariasi dari asesor, artinya terdapat SMK yang tidak diukur dengan tepat oleh butir instrumen.

Gambar-1. Wright map level SMK dan butir penilaian

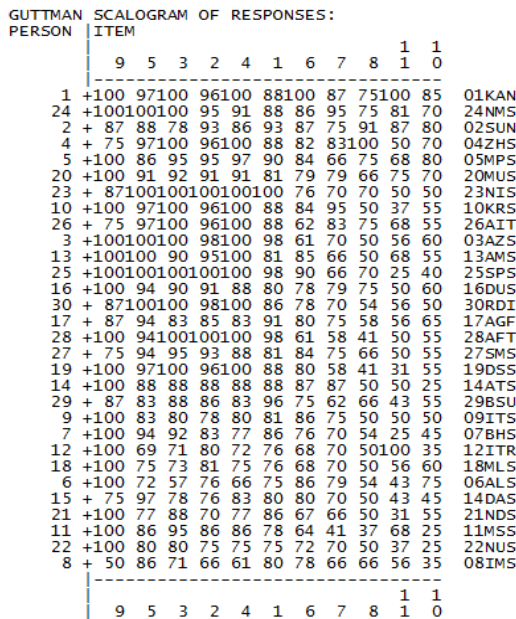
Jika dikelompokkan, maka terdapat tiga kategori SMK, sementara untuk butir penilaian terdapat 8 kategori level butir.

Hasil uji dimensional menunjukkan, untuk *raw variance measure* teramati 71% (istimewa), sedangkan untuk *unexplained variance* tidak ada nilai di atas 15%, artinya instrumen PKKS bersifat unidimensional, bahwa terdapat hanya satu atribut laten yang mendasari para asesor dalam mengisi instrumen, atau dikatakan instrumen PKKS mampu mengukur sesuatu yang seharusnya dapat diukur.

Tabel-8. Level kategori SMK

Level Kategori SMK	Jumlah	Kode/Butir yang tidak dipenuhi
Baik	4	1 SMK sangat tinggi (<i>outlier</i>)
Sedang	24	-
Rendah	2	1 SMK sangat rendah (<i>batas outlier</i>)

Level Kategori SMK	Jumlah	Kode/Butir yang tidak dipenuhi
Baik	4	Tidak ada
Sedang	24	SK4 (tugas tambahan)
Rendah	2	SK4 (tugas tambahan) (kewirausahaan) SK5 (PKB) SK1 (manajerial) SK3 (supervisi)



Gambar-2. Skalogram distribusi penilaian item

mampu memenuhi aspek SK6 (kerja sama). Namun untuk aspek PK (penilaian kinerja), semua SMK memenuhi semua subaspek.

Tampilan *Wright Map* yang merupakan distribusi level SMK dan butir PKKS (Gambar-1) menunjukkan, terdapat satu (3.3%) SMK yang memiliki nilai Logit sangat tinggi (*outlier*). Mengacu indeks separasi, terdapat empat (13.3%) SMK pada kelompok tinggi, 24 (80.0%) SMK kelompok sedang, dua (6.7%) SMK kelompok rendah yang didalamnya terdapat satu SMK sangat rendah tepat di garis batas *outlier*. Namun karena keterbatasan aplikasi **Ministep 5.1** tidak dapat menampilkan identitas SMK yang termaksud.

Untuk kategori *item*, untuk SMK kategori baik (dan sangat baik) dapat memenuhi semua *item* PKKS. Sementara untuk SMK kategori rendah hampir semua tidak dapat memenuhi *item* aspek SK (sasaran kinerja), artinya hanya

Penelusuran tampilan matrik Guttman (Gambar-2) menunjukkan, tidak ada hasil penilaian SMK yang identik, artinya penilaian asesor obyektif atau tidak melakukan *copy paste*. *Item* penilaian yang memiliki nilai kecil atau sulit dipenuhi SMK adalah butir ke-10 atau SK5 (PKB atau penilaian berkelanjutan), sementara butir yang paling mudah untuk dipenuhi adalah butir ke-9 atau SK4 (tugas tambahan). Artinya ada ketidakcocokan dengan *Wright Map*. Nampak pula bahwa, SMK yang memiliki level (sangat) tinggi *outlier* adalah input data pertama, yaitu SMKN 1 Kalianget, sementara SMK level (sangat) rendah *oulier* adalah SMKS Al-Imron.

Instrumen PKKS yang dikembangkan Tim Korwas (Koordinator Pengawas) Prov. Jatim 2020 menggunakan aplikasi **PKKS Ver. 5.00** memiliki Relibilitas yang layak, reliabilitas item 0.97 kategori istimewa, serta memiliki konsistensi data antara SMK dan item yang ditunjukkan dengan nilai Alpha Cronbach 0.75 kategori baik.

Instrumen di atas juga memiliki syarat kesesuaian (*fit*) antara data kualitas SMK dan *item* PKKS, artinya tidak ada miskonsepsi antara pemahaman asesor terhadap butir PKKS, yang dibuktikan dari nilai outfit MNSQ dan outfit ZSTD yang memenuhi syarat. Selain itu, instrumen juga mampu menunjukkan bahwa SMK berkelas tinggi memiliki probabilitas memperoleh nilai tinggi, yang ditunjukkan dengan nilai *Pt. Measure Corr* pada rentang 0.4 s.d. 0.85, serta tidak menunjukkan unsur bias. Namun, khusus butir Aspek Sasaran Kerja (SK) subaspek supervisi harus direviu, sebab memiliki daya beda rendah atau dimungkinkan asesor memberikan penilaian asal-asalan (Sumintono & Widhiarso, 2014).

Hasil penelusuran data menunjukan instrumen Sasaran Kerja (SK) memiliki kelemahan, yaitu: (1) pedoman penilaian (rubrik) kurang detail, serta (2) dilakukan oleh satu asesor sehingga tidak memiliki reliabilitas antar-penilai. Selain itu, terdapat pilihan jawaban dalam instrumen PKKS yang membingungkan, terutama untuk aspek Sasaran Kerja (SK) subaspek supervisi dan aspek Perilaku Kerja (PK) subaspek kerja sama memiliki distraktor membingungkan. Hasil penelusuran data dari ministep menunjukkan terjadi inkonsistensi rubrik penilaian.

Rubrik penilaian detail tersebut penting sebab, ia merupakan standar kinerja untuk kompetensi tertentu (Suwarno, 2017) sehingga hasil penilaian terukur. Sementara reliabilitas antar-*observer* atau *inter-rater* diperlukan agar tercipta reliabilitas berdasarkan konsistensi penilaian dua asesor berbeda terhadap suatu butir, karena belum tentu pengamat manusia menafsirkan jawaban dengan cara yang sama (Sumintono & Widhiarso, 2014). Ketidakkonsistenan pengisian item PKKS dimungkinkan menyebabkan nilai *Differential Person Functioning* tidak sesuai harapan, bahkan terdapat SMK yang berkualitas rendah malah memperoleh hasil skor PKKS tinggi.

Perbandingan gambar *Wright Map* dan Skalogram Guttman menunjukkan hasil yang tidak sama, di mana pada *Wright Map*, butir yang sulit dipenuhi SMK adalah aspek Sasaran Kerja (SK) subaspek Tugas Tambahan, dan yang mudah dipenuhi adalah aspek Sasaran Kerja (SK) sub-aspek supervisi. Sementara pada Skalogram Guttman, butir yang sulit dipenuhi SMK adalah aspek Sasaran Kerja (SK) sub-aspek Penilaian Berkelanjutan, dan yang paling mudah dipenuhi adalah aspek Sasaran Kerja (SK) sub-aspek tugas tambahan. Fenomena di atas perlu dilakukan reviu, bagaimana konsistensi butir PKKS, terutama perbaikan rubrik penilaian.

Simpulan

Instrumen PKKS menggunakan aplikasi **PKKS Ver. 5.00** yang dikembangkan oleh Tim Korwas (Koordinator Pengawas) Prov. Jatim 2020 memiliki kelayakan dalam aspek validitas, obyektifitas, dan reliabilitas. Terdapat item PKKS yang memberikan hasil penilaian inkonsistensi, terutama pada aspek Sasaran Kerja, yang disebabkan karena: (a) kurang detailnya rubrik penilaian, serta (b) penilai hanya dilakukan oleh satu asesor.

Daftar Rujukan

- Akdon. 2006. *Strategis Managemen For Educational Management*, Bandung: Alfabeta.
- Ardiyanti, Yusi. *Berpikir Kritis Siswa Dalam Pembelajaran Berbasis Masalah Berbantuan Kunci Determinasi*. Jurnal Pendidikan Indonesia Program Studi Pendidikan Matematika, Universitas Singaperbangsa Karawang. Vol. 5, No. 2, Oktober 2016 pp. 193
- Badan Pengembangan Sumber daya Manusia Pendidikan dan Kebudayaan dan Penjaminan Mutu Pendidikan. 2012. *Pedoman Penilaian Kinerja Kepala Sekolah/Madrasah*. Jakarta: Kementerian Pendidikan dan Kebudayaan Pusat Pengembangan Tenaga Kependidikan. Buchari Alma. 2018. *Manajemen Pemasaran*. Bandung: Alfabeta
- Faishal Haq, Muhammad. *Analisis Standar Pengelolaan Pendidikan Dasar dan Menengah*. Jurnal Evaluasi, Vol.1, No.1, Maret 2017 pp. 26
- Faizin, Imam. *Strategi Pemasaran Jasa Pendidikan Dalam Meningkatkan Nilai Jual Madrasah Dalam Meningkatkan Nilai Jual Madrasah*. Jurnal Madaniyah, Volume 7 Nomor 2 Edisi Agustus 2017 pp. 261
- Febrianto, Priyono Tri dan Sulaiman. *Kajian strategis dan prioritas pembangunan pendidikan menengah untuk meningkatkan mutu pendidikan di Pulau Madura* Fakultas Ilmu Pendidikan, Universitas Trunojoyo, Bangkalan, Madura
- <https://beritajatim.com/pendidikan-kesehatan/double-track-sma-diikuti-14-ribu-siswa-dari-157-sekolah> Minggu, 17 November 2019
- <https://jatim.bpk.go.id/dari-media/dinas-pendidikan-jawa-timur-fokus-madura-tingkatkan-kualitas-pendidikan> diterbitkan 14 januari 2019 diakses 08 sep 2021
- Pemerintah Kabupaten Sumenep Peraturan Daerah Kabupaten Sumenep Nomor : 7 Tahun 2013 Tentang Pengelolaan dan Penyelenggaraan Pendidikan
- Peraturan Menteri Pendidikan Dan Kebudayaan Republik Indonesia Nomor 6 Tahun 2018 Tentang Penugasan Guru Sebagai Kepala Sekolah
- Sumintono, B., & Widhiarso, W. (2014). *Aplikasi Model Rasch Untuk Penelitian Ilmu-Ilmu Sosial (Edisi Revisi)*. Bandung: Trim Komunikata Publishing House.
- Sumintono, Bambang dan Widhiarso, Wahyu. 2015. *Aplikasi Pemodelan Rasch Pada Assessment Pendidikan*. Cimahi: Trim Komunikata Publishing House.

Suwarno. 2017. *Pentingnya Rubrik Penilaian Dalam Penilaian Keterampilan Vocational Siswa*:
Prosiding Seminar Nasional Pendidikan Vokasi Pusat Pendidikan Vokasi FKIP
Universitas Sebelas Maret