

COMPARATIVE JUDGEMENT (CJ)

N. K. Riantini¹ dan R. A. Apsari²

¹Program Studi Pendidikan Matematika, FMIPA, Universitas Pendidikan Ganesha,
Jalan Udayana No. 11, Singaraja, e-mail:komangrian999@gmail.com

²Program Studi Pendidikan Matematika, FKIP, Universitas Mataram, Jalan Majapahit
No.62, Gomong, Mataram, e-mail:ra.apsari@unram.ac.id

Abstrak. *Comparative Judgement (CJ)*, merupakan suatu penilaian yang secara penuh hanya dilakukan oleh ahli di bidang keahliannya. Dalam penilaian CJ pengaturan ahli tidak menggunakan rubrik untuk menghitung skor untuk setiap siswa, sebaliknya ahli hanya memandang sekumpulan pekerjaan siswa dan mengidentifikasi mana yang lebih baik berdasarkan rubrik yang telah ditentukan dan mengidentifikasi jawaban siswa yang lebih baik dalam satu tes dengan membandingkan dua jawaban siswa kemudian mengurutkan peringkat dari yang terbaik sampai yang sangat kurang. Dalam makalah ini dipilih tiga ahli untuk membandingkan kedua jawaban siswa, dengan tes yang digunakan berupa tes essay. Subjek yang digunakan dalam makalah ini adalah siswa kelas XI AP 4 SMK Negeri 2 Singaraja dengan jumlah siswa 40 orang yang terdiri dari 25 siswa laki-laki dan 15 siswa perempuan. Potensi untuk CJ sebagai alat untuk kolaborasi internasional dalam penilaian, mengevaluasi pembelajaran siswa disekolah, keterlibatan, dan pengalaman siswa. Hasil yang didapatkan menunjukkan bahwa penilaian dengan menggunakan CJ menunjukkan hasil yang sangat positif dan efektif dalam memfasilitasi penilaian siswa untuk mencapai solusinya dan meningkatkan pemahaman konsep matematika siswa.

Kata kunci: Comparative Judgement (CJ), Penilaian Ahli

Abstract. *Comparative Judgement* is a holistic evaluation performed by experts on the field. In CJ, the experts are not using conventional score rubric to check the correctness of the students, but they will compare the work of students in solving the same problem and make the rank from it. The expert will use their expertise to judge which work contains better ideas among others. The data were gathered from 40 students of vocational school XI AP 4 SMK Negeri 2 Singaraja. The data was analyzed using expert evaluation. In this paper, three experts of mathematics education have been chosen to compare the students' answer in solving test on conceptual understanding in mathematical topic. The result showed that CJ gives positive remarks to compare the work of students.

Key Words: Comparative Judgement, Expert Evaluation

I. PENDAHULUAN

Pengambilan keputusan merupakan sesuatu yang dianggap sebagai suatu hasil atau keluaran dari proses mental atau kognitif yang membawa pada pemilihan suatu jalur tindakan di antara beberapa alternatif yang tersedia. Setiap proses pengambilan keputusan selalu menghasilkan satu pilihan final. Keputusan dibuat untuk mencapai tujuan melalui pelaksanaan atau tindakan. Dalam mengambil suatu keputusan tentu banyak faktor yang akan dipertimbangkan untuk mendapatkan suatu keputusan yang akan menghasilkan suatu solusi yang benar. Didunia pendidikan pengambilan keputusan dalam hal penilaian sangat berperan penting untuk skor penilaian terutama dalam bidang matematika. Untuk memudahkan dalam penilaian maka ada suatu model sebagai alternative untuk melakukan penilaian dengan pengambilan keputusan dari para ahli dalam bidang keahliannya yaitu dengan menggunakan model *Comparative Judgement* (CJ). Baru-baru ini model penilaian CJ sangat populer digunakan dalam penilaian karena sudah terbukti menjadi model yang sangat dapat diandalkan dalam proses penilaian, dimana penggunaannya berpusat pada ahli membuat penilaian antara lembar kerja siswa saat membandingkannya satu sama lain.

Comparative Judgement (CJ) adalah penilaian yang dilakukan berdasarkan kepentingan relatif dua elemen pada suatu tingkat tertentu dalam kaitannya dengan tingkatan di atasnya. Secara teoritis kelebihan dari pendekatan CJ adalah penilaian secara penuh hanya dilakukan oleh ahli di bidang tersebut tanpa menggunakan kriteria penilaian dan rubrik penilaian. Dengan kata lain, validitas dari tanggapan tes berdasarkan atas pemahaman yang dimiliki oleh para ahli sesuai dengan bidangnya. Dengan mempertimbangkan isi dari tes yang akan digunakan. Hal ini membuat CJ sangat menjanjikan untuk menilai konstruksi yang diakui dan dianggap penting oleh pakar pendidikan, seperti pemecahan masalah dan pemahaman konsep, tetapi mustahil untuk menentukan rubrik penilaian secara komprehensif [1]. Ini dikarenakan kompetensi peserta didik tidak hanya dilihat pada akhir proses saja, tetapi juga pada saat proses berlangsung. Karenanya rubrik disini dapat berfungsi ganda, satu sisi sebagai penuntun kerja dan satu sisi sebagai sebagai instrument evaluasi. Rubrik juga sangat cocok digunakan pada era yang sangat mengedepankan kompetensi/kinerja seperti sekarang ini. Selain itu, konstruksi seperti pemahaman konsep itu memungkinkan untuk

MANDALIKA

Mathematics and Educations Journal

Volume 1 Nomor 1, Juni 2019

memberikan tes open-ended.

Secara praktis untuk mempelajari penggunaan CJ untuk menilai pemahaman konsep adalah potensi efisiensinya. Tidak seperti instrumen psikometrik (berkaitan dengan pengukuran kepribadian seseorang) yang dikembangkan dengan saksama, CJ dapat diterapkan dengan cepat ke konsep sasaran apa pun dengan sedikit usaha di luar merekrut ahli dengan bidang keahliannya masing-masing sesuai dengan keperluan. Tidak seperti wawancara klinis, CJ menggunakan prinsip psikologis yang telah lama ditetapkan dalam perbandingan berpasangan dan menghasilkan validitas dan reliabilitas yang tinggi dengan pelatihan yang minimal [2].

Dalam penilaian CJ pengaturan ahli tidak menggunakan rubrik untuk menghitung skor untuk setiap siswa, sebaliknya ahli hanya memandang sekumpulan pekerjaan siswa dan mengidentifikasi barang mana yang “lebih baik” berdasarkan rubrik yang telah ditentukan dan ke profesionalan ahli sendiri. Saat ahli melakukan proses ini mengidentifikasi yang lebih baik dalam satu tes dengan membandingkan 2 jawaban siswa kemudian mengurutkan peringkat dari yang terbaik sampai yang sangat kurang.

Penilaian dengan CJ sangat tepat digunakan untuk mengukur hasil dari tes masalah terbuka, misalnya dalam mencermati pemahaman konsep siswa. Pemahaman konsep adalah kemampuan yang dimiliki seseorang untuk mengemukakan kembali ilmu yang diperolehnya baik dalam bentuk ucapan maupun tulisan kepada orang sehingga orang lain tersebut benar-benar mengerti apa yang disampaikan. Menurut Bloom [3], pemahaman adalah tingkatan yang paling rendah dalam aspek kognitif yang berhubungan dengan penguasaan sesuatu. Paling tidak ada 5 hal yang dilakukan siswa jika siswa tersebut telah memahami suatu konsep, yaitu dapat menyatakan konsep dengan kata-kata sendiri, menyebutkan ciri dan contoh dari konsep, membedakan contoh dan bukan contoh dari konsep serta memecahkan masalah yang berkenaan dengan konsep tersebut [4&5].

Sementara itu, menurut Bloom [6], terdapat tiga macam indikator pemahaman yaitu: pengubahan (*translation*), pemberian arti (*intepretation*) dan ekstrapolasi (*ekstrapolation*). Dalam belajar matematika, proses pengubahan dapat dilihat dari kemampuan siswa untuk mengubah soal atau permasalahan dalam bentuk kalimat biasa menjadi kalimat matematika dalam hal ini notasi/symbol dan sebaliknya siswa dapat menerjemahkan bahasa matematika menjadi bentuk kalimat biasa. Untuk proses

interpretasi dapat dilihat dari kemampuan siswa dalam memberi arti terhadap suatu konsep yang sedang dipelajari. Seorang dapat dikatakan telah dapat menginterpretasikan tentang suatu konsep atau prinsip tertentu jika dia telah mampu membedakan, membandingkan atau mempertentangkannya dengan suatu yang lain. Sementara itu, untuk ekstrapolasi dapat dilihat dari kemampuan siswa dalam membuat langkah-langkah yang mencerminkan suatu interpretasi konsep dengan tujuan untuk mengembangkan pemikiran siswa sebagai persiapan untuk menghadapi suatu permasalahan.

Ini berarti bahwa pemahaman konsep yang kuat dalam pembelajaran matematika merupakan tonggak utama dan sangat membantu siswa dalam memahami suatu materi matematika. Jadi secara operasional yang dimaksud pemahaman konsep matematika yaitu dapatnya siswa memecahkan suatu soal atau masalah matematika dengan menggunakan konsep yang benar dengan membuat perubahan, interpretasi, dan ekstrapolasi.

II. METODE PENELITIAN

Data dikumpulkan dengan memberikan soal pemahaman konsep dalam bentuk soal uraian bagi 40 orang siswa kelas XI AP4 SMK Negeri 2 Singaraja. Jawaban siswa kemudian dianalisis oleh ahli pendidikan matematika untuk memutuskan gagasan mana yang lebih baik dari kedua jawaban yang diberikan. Adapun soal itu sebagai berikut

Rumah Andi berada di pusat kota pada peta. Jika Andi ingin ke rumah Budi dengan menggunakan peta, ia harus berjalan memutar. Pada peta, Andi harus ke arah timur 2 cm dilanjutkan ke arah utara 3 cm. Jika skala pada peta menyatakan bahwa 1 cm = 2 km pada jarak sesungguhnya. Tentukan jarak yang di tempuh Andi untuk sampai kerumah Budi!

Tiga orang ahli pendidikan matematika yang melakukan penilaian dengan CJ pada penelitian ini dapat dicermati pada Tabel 1 berikut.

Tabel 1: Ahli Pendidikan Matematika Pada Penelitian Ini

No.	Nama	Institusi
1.	Muhammad Husnul	Dosen Prodi S1 Pendidikan Matematika

No.	Nama	Institusi
	Khuluq, S.Pd. M.Sc	Universitas Negeri Makasar, Sulawesi Selatan
2.	Nur Chasanah, S.Pd. M.Sc	Guru Matematika SMA Bina Citra Utama, Supervisor Guru Matematika SD- SMA Bina Citra Utama, Kalimantan Tengah
3.	Andi Harpeni Dewantara, M.Pd.	Dosen Jurusan Pendidikan Guru Sekolah Dasar IAIN Bone, Sulawesi Selatan

III. HASIL DAN PEMBAHASAN

Awal mulanya model CJ ini telah digunakan pada studi komparatif sejak tahun 1996 [7]. Pollitt adalah orang pertama yang mengusulkan model penilaian komparatif (CJ) sebagai alternatif untuk penilaian. Dia menyatakan bahwa, penilaian secara holistik lebih baik dari “mikro-penilaian” tingkat pertanyaan yang lebih cenderung mendukung reabilitasnya dengan mengorbankan validitas. Sebuah pendekatan holistik memiliki potensi untuk kembali fokus pada penilaian keterampilan demonstrasi. Hal ini khususnya terjadi karena kinerja ahli cenderung memiliki subjektivitas yang tinggi. Namun di sisi lain, salah satu kelemahan penilaian holistik adalah kurangnya kriteria penilaian yang terlihat [9]. Sehingga mustahil untuk menentukan rubrik penilaian secara komprehensif

Penilaian ini dilakukan secara penuh oleh ahli di bidang tersebut tanpa menggunakan kriteria penilaian dan rubrik penilaian. Prinsip ini berasal dari prinsip psikologis bahwa manusia lebih baik dalam membandingkan dua objek daripada mereka membandingkan satu objek dengan kriteria yang ditentukan [10]. Berdasarkan penelitian yang telah dilakukan oleh Whitehouse Claire [11], ahli yang mengajar pada bidang yang sama dengan apa yang dinilainya akan membuat penilaian yang dihasilkan semakin konsisten. Tugas ahli disini adalah memutuskan yang manakah dari dua bagian jawaban siswa yang pemaparannya lebih baik. Ahli juga bisa memberikan komentar tambahan apabila diperlukan.

CJ sangat cocok digunakan dalam penilaian sumatif karena semakin besar jumlah penilaian, maka semakin tinggi reabilitas urutan peringkatnya. Ini juga bermanfaat untuk meningkatkan jumlah ahli, semakin banyak ahli yang digunakan dalam penilaian maka semakin kuat tingkat reabilitas yang didapatkan.

Pendekatan CJ digunakan untuk mengukur pemahaman konsep matematika pada penelitian ini melibatkan empat langkah, sebagai berikut.

MANDALIKA

Mathematics and Educations Journal

Volume 1 Nomor 1, Juni 2019

1. Langkah pertama peserta menyelesaikan semua pertanyaan di dalam tes yang disusun untuk menggali pemahaman mereka terhadap suatu konsep tertentu. Pertanyaan tes bersifat open-ended yang memungkinkan para peserta untuk memberikan berbagai jenis tanggapan.
2. Tahap kedua dari pendekatan CJ membutuhkan ahli pendidikan matematika untuk melakukan penilaian terhadap sepasang kualitas dari jawaban tes yang diberikan. Masing-masing ahli diberikan jawaban dari 2 siswa. Kemudian masing-masing ahli diminta untuk memutuskan gagasan mana yang lebih baik dari kedua jawaban yang diberikan. Tidak ada kriteria penilaian secara terperinci atau rubrik penilaian dan sebaliknya setiap penilaian bersifat holistik dan hanya didasarkan pada penilaian ahli. Dengan cara yang sama hasil jawaban dari tes tersebut disajikan berpasang-pasangan dan diberikan kepada beberapa ahli kemudian keputusannya dimodelkan secara statistik untuk menghasilkan urutan skala tanggapan tes dari 'terburuk' hingga 'terbaik'.
3. Tahap ketiga dari pendekatan CJ yaitu Konsistensi internal. Wiersma [12] menyatakan konsistensi internal berarti "konsistensi dari tes dalam mengukur apa yang seharusnya diukur". Kemudian, Gay [13] menyatakan konsistensi internal adalah derajat konsistensi pengukuran yang ditampilkan oleh tes terhadap apa yang ingin diukur. Jadi konsistensi internal berkenaan dengan tingkatan atau derajat yang menunjukkan seberapa jauh tes dapat mengukur secara konsisten apa yang seharusnya diukur. Kemudian selanjutnya perhitungan akan dilakukan dengan menggunakan rumus Spearman-Brown sebagai berikut.

$$\frac{j \cdot icc}{1 + (j-1)icc}$$

Keterangan:

J = Jumlah ahli

Icc = Intraclass koefisien korelasi

Kriteria yang dapat diacu adalah koefisien reliabilitas 0,80 menyatakan tes tersebut *acceptable* [14]. Oleh karena koefisien reliabilitas secara wajar bergerak pada interval 0,00-1,00, maka kriteria-kriterianya yaitu: Sangat Kurang (0,00-0,20), Kurang (0,20-0,40), Cukup (0,40-0,60), Baik (0,60-0,80), Sangat Baik (0,80-1,00)

MANDALIKA

Mathematics and Educations Journal

Volume 1 Nomor 1, Juni 2019

Tes hasil belajar dengan indeks reliabilitas berada pada kategori cukup, baik, dan sangat baik ditoleransi untuk diterima sebagai perangkat tes yang relatif baku.

4. Tahap terakhir dari pendekatan CJ adalah meranking dengan menggunakan matriks berpasangan. Misalkan kita bertujuan untuk membandingkan n kondisi O_1, \dots, O_n . (misal n gambar, yang umumnya menggunakan konten yang sama, masing-masing diproses dengan algoritma yang berbeda) dengan skor kualitas sebenarnya yang tidak diketahui $O = (O_1, \dots, O_n)$, $O_i \in R$. Tujuan dari analisis ini adalah untuk memperkirakan skor $O = (O_1, \dots, O_n)$, yang mendekati skor kualitas Q sebenarnya. Karena perbandingan berpasangan bersifat relatif, ketika $O_i = O_j$ maka akan bernilai 0. Perbandingan berpasangan biasanya direpresentasikan dalam bentuk matriks hitungan C , di mana setiap elemen C_{ij} mengukur jumlah kasus di mana kondisi O_i telah dipilih lebih baik dari kondisi O_j . perhatikan matriks berikut.

$$C = \begin{bmatrix} C_{11} & C_{12} & C_{13} \\ C_{21} & C_{22} & C_{23} \\ C_{31} & C_{32} & C_{33} \end{bmatrix}$$

Setelah melihat data dalam penelitian yang dibandingkan maka kita bisa membentuk matriks berikut.

$$C = \begin{bmatrix} 0 & 3 & 0 \\ 27 & 0 & 7 \\ 30 & 23 & 0 \end{bmatrix}$$

Ini, $C_{12} = 3$ memberitahu kita bahwa kondisi O_1 telah dipilih tiga kali lebih baik daripada kondisi O_2 , dan $C_{21} = 27$ memberitahu kita bahwa kondisi O_2 telah dipilih 27 kali lebih baik daripada O_1 begitu seterusnya. Itu probabilitas bahwa satu kondisi dipilih lebih baik dari yang lain (dilambangkan sebagai P_{ij} untuk O_i dan O_j) diperkirakan menggunakan informasi empiris dalam matriks C dan berikut adalah rumus untuk mencari probabilitasnya [15].

$$P_{ij} = \frac{C_{ij}}{C_{ij} + C_{ji}}, i \neq j$$

Analisis Data

Pada analisis konvensional jawaban siswa dinilai berdasarkan rubrik penilaian. Tetapi ini memiliki kelemahan sehingga harus sangat berhati-hati, memerlukan proses intensif yang harus berulang-ulang untuk setiap konsep yang penting kemudian merekam satu per satu wawancara klinis (wawancara untuk mempermudah penilaian) dan menganalisis atau menilai kualitas pemahaman dari masing-masing peserta didik. Tetapi wawancara klinis memiliki kekurangan karena membutuhkan keterampilan dan

konsistensi dari pewawancara dan penilai, dan tidak selalu mengarah pada hasil yang dapat dipercaya. Biaya, durasi dan kesulitan mengukur pemahaman konsep adalah hambatan dalam kemajuan pendidikan matematika.

Dalam makalah ini penulis menguji metode baru yang dirancang untuk mengukur pemahaman konsep, yang disebut dengan perbandingan ahli CJ. Oleh karena banyak subjek ketika komparasi yang dilakukan terlalu banyak maka pada makalah ini digunakan modifikasi dengan menggabungkan rubrik konvensional dengan CJ yang diharapkan dapat mengatasi kelemahan dari metode yang sudah ada.

Jika penilaian dilakukan secara konvensional dengan menggunakan rubrik penilaian maka hasil yang didapatkan siswa adalah sebagai berikut.

Tabel 2 Hasil Jawaban Siswa

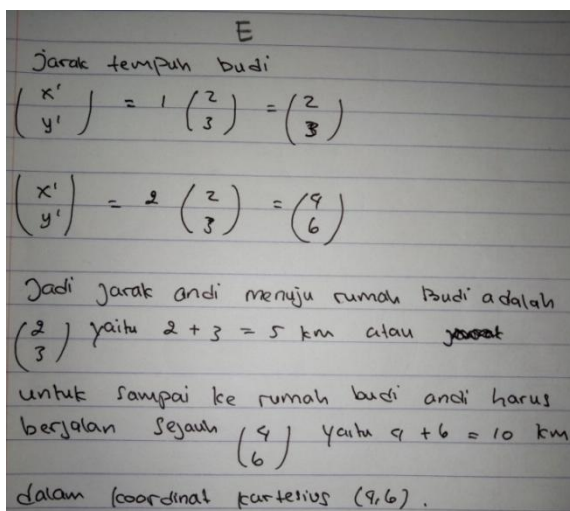
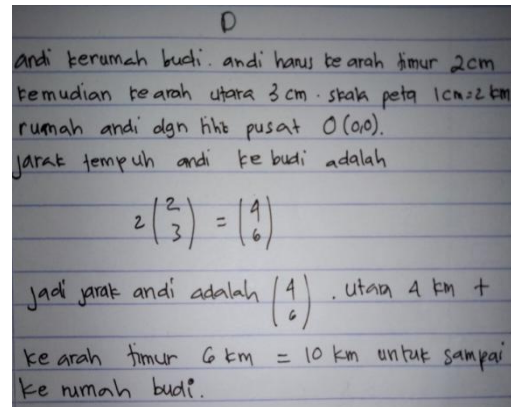
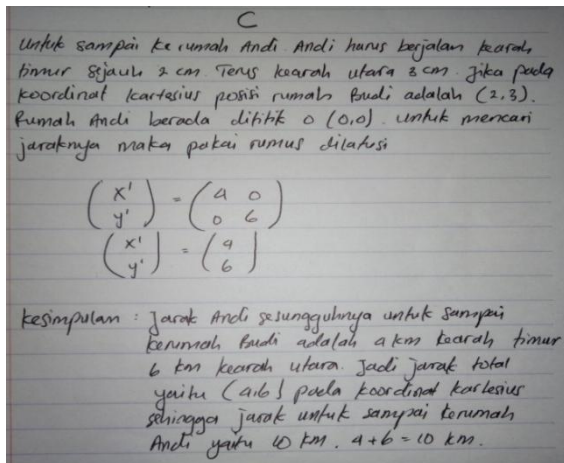
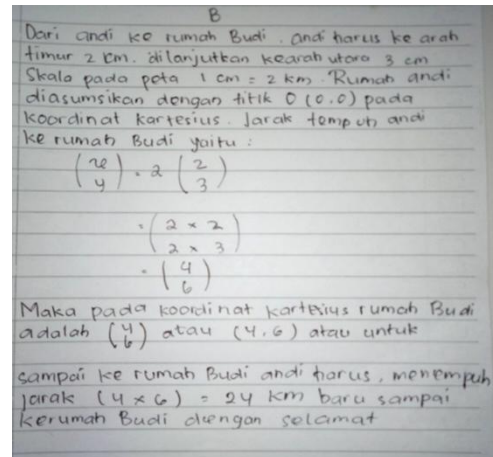
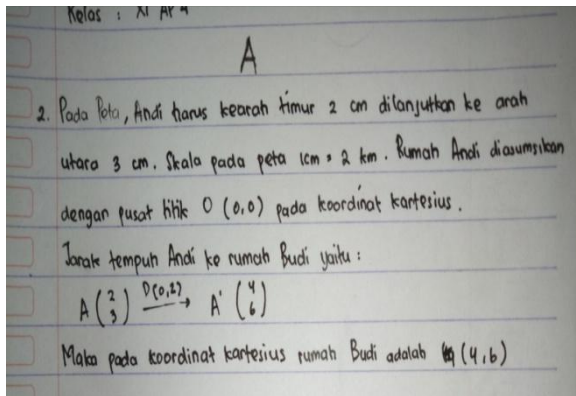
No	Skor	Frekuensi
1	2	5
2	3	5
3	4	8
4	6	9
5	8	5
6	9	4
7	10	4
Jumlah		40

Setelah melihat hasil yang didapatkan oleh masing-masing siswa, maka dipilihlah siswa yang mendapatkan skor 8 dimana siswa yang mendapatkan skor tersebut berjumlah 5 orang. Dari hasil tes yang telah dilakukan maka berikut adalah ke-5 jawaban siswa tersebut. Gambar 1 berikut adalah contoh dari jawaban siswa.

MANDALIKA

Mathematics and Educations Journal

Volume 1 Nomor 1, Juni 2019



Gambar 1. Jawaban 5 Siswa yang Diteliti Lebih Lanjut

Setelah itu, semua jawaban dipasangkan. Misalnya, jawaban siswa A dipasangkan dengan siswa B, siswa A dengan C dan seterusnya sehingga terbentuk 10 kelompok

MANDALIKA

Mathematics and Educations Journal

Volume 1 Nomor 1, Juni 2019

dengan kode X1, X2, X3, X4, X5, X6, X7, X8, X9, X10 berturut-turut. Selanjutnya ahli pendidikan matematika mengidentifikasi jawaban mana yang menunjukkan kinerja yang lebih kuat dari masing-masing kelompok itu kemudian akan dipilih satu jawaban yang memiliki kriteria jawaban yang kuat disetiap kelompok tersebut, begitu selanjutnya sampe ke-10 kelompok. Berikut adalah hasil rekapitan penilaian jawaban siswa yang lebih baik dari masing-masing kelompok dengan menggunakan CJ dari 3 pakar pendidikan matematika.

Tabel 3. Rekapitulasi Penilaian Ahli

No	Nama Kelompok	Jawaban terbaik dari Penilai		
		I	II	III
1	X1	B	A	B
2	X2	C	C	C
3	X3	D	D	D
4	X4	E	E	E
5	X5	C	C	C
6	X6	D	D	D
7	X7	E	E	E
8	X8	D	D	C
9	X9	E	C	C
10	X10	E	D	D

Dari masing-masing penilaian ahli maka digabungkan menjadi satu tabel untuk menilai kesepakatan antar-penilai sebagai berikut.

Tabel 4. Kesepakatan Antar Penilai

Subjek	A	B	C	D	E
Judges I	5	4	3	2	1
Judges II	4	5	2	1	3
Judges III	5	4	1	2	3

Korelasi antara ahli I dan III identik, tetapi ahli II kurang setuju dengan penilaian dari ahli I dan II. Maka salah satu solusi untuk masalah ini adalah menghitung koefisien korelasi intraclass karena ahli yang digunakan lebih dari dua. Untuk data dari tiga ahli

MANDALIKA

Mathematics and Educations Journal

Volume 1 Nomor 1, Juni 2019

dibawah ini, koefisien korelasi intraclass adalah 0,714 yang didapatkan dengan menghitung koefisien korelasi intraclass dengan menggunakan SPSS sebagai berikut:

	Intraclass Correlation ^b	95% Confidence Interval		F Test with True Value 0			
		Lower Bound	Upper Bound	Value	df1	df2	Sig
Single Measures	,714 ^a	,138	,963	7,000	4	8	,010
Average Measures	,882 ^c	,325	,987	7,000	4	8	,010

Kemudian setelah didapatkan koefisien korelasi intraclass (icc) maka dilakukan perhitungan secara manual yaitu:

$$\frac{j \cdot icc}{1 + (j-1)icc} = \frac{3 \cdot (0,714)}{1 + (3-1)(0,714)} = 0,882$$

Sedemikian sehingga didapatkan hasil perhitungan akhir adalah 0,882 yang artinya penilaian tersebut berada dalam kriteria realibilitas yang sangat baik.

Langkah terakhir dari CJ yaitu membuat perankingan dari matriks perbandingan berpasangan seperti pada Tabel 5 berikut.

Tabel 5. Perbandingan Data

No	Data yang dibandingkan		Pemenang		
			I	II	III
1	A	B	B	A	B
2	A	C	C	C	C
3	A	D	D	D	D
4	A	E	E	E	E
5	B	C	C	C	C
6	B	D	D	D	D
7	B	E	E	E	E
8	C	D	D	D	C
9	C	E	E	C	C
10	D	E	E	D	D

Dari Tabel 5 dapat dilihat bahwa A menang 1 kali, B 2 kali, C tiga kali, D 4 kali dan E 5 kali. Oleh karena ada 5 data yang dibandingkan, susun Matriks perbandingan berpasangan 5×5 dengan isian sebagai berikut.

MANDALIKA

Mathematics and Educations Journal

Volume 1 Nomor 1, Juni 2019

$$C = \begin{bmatrix} C_{11} & C_{12} & C_{13} & C_{14} & C_{15} \\ C_{21} & C_{22} & C_{23} & C_{24} & C_{25} \\ C_{31} & C_{32} & C_{33} & C_{34} & C_{35} \\ C_{41} & C_{42} & C_{43} & C_{44} & C_{45} \\ C_{51} & C_{52} & C_{53} & C_{54} & C_{55} \end{bmatrix} \quad C = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 \\ 3 & 3 & 0 & 1 & 2 \\ 3 & 3 & 2 & 0 & 2 \\ 3 & 3 & 1 & 1 & 2 \end{bmatrix}$$

Lalu cari probabilitas dengan menggunakan rumus:

$$P_{ij} = \frac{C_{ij}}{C_{ij} + C_{ji}}, i \neq j$$

- $P_{12} = \frac{C_{12}}{C_{12} + C_{21}} = \frac{1}{1+2} = \frac{1}{3} = 0,3$ $P_{42} = \frac{C_{42}}{C_{42} + C_{24}} = \frac{3}{3+0} = \frac{3}{3} = 1$
- $P_{21} = \frac{C_{21}}{C_{21} + C_{12}} = \frac{2}{2+1} = \frac{2}{3} = 0,6$ $P_{43} = \frac{C_{43}}{C_{43} + C_{34}} = \frac{2}{2+1} = \frac{2}{3} = 0,6$
- $P_{31} = \frac{C_{31}}{C_{31} + C_{13}} = \frac{3}{3+0} = \frac{3}{3} = 1$ $P_{45} = \frac{C_{45}}{C_{45} + C_{54}} = \frac{2}{2+1} = \frac{2}{3} = 0,6$
- $P_{32} = \frac{C_{32}}{C_{32} + C_{23}} = \frac{3}{3+0} = \frac{3}{3} = 1$ $P_{51} = \frac{C_{51}}{C_{51} + C_{15}} = \frac{3}{3+0} = \frac{3}{3} = 1$
- $P_{34} = \frac{C_{34}}{C_{34} + C_{43}} = \frac{1}{1+2} = \frac{1}{3} = 0,3$ $P_{52} = \frac{C_{52}}{C_{52} + C_{25}} = \frac{3}{3+0} = \frac{3}{3} = 1$
- $P_{35} = \frac{C_{35}}{C_{35} + C_{53}} = \frac{2}{2+1} = \frac{2}{3} = 0,6$ $P_{53} = \frac{C_{53}}{C_{53} + C_{35}} = \frac{1}{1+2} = \frac{1}{3} = 0,3$
- $P_{41} = \frac{C_{41}}{C_{41} + C_{14}} = \frac{3}{3+0} = \frac{3}{3} = 1$ $P_{54} = \frac{C_{54}}{C_{54} + C_{45}} = \frac{1}{1+2} = \frac{1}{3} = 0,3$
- $P_{11} = \frac{C_{11}}{C_{11} + C_{11}} = \frac{0}{0+0} = 0$ $P_{13} = \frac{C_{13}}{C_{13} + C_{31}} = \frac{0}{0+3} = \frac{0}{3} = 0$
- $P_{14} = \frac{C_{14}}{C_{14} + C_{41}} = \frac{0}{0+3} = \frac{0}{3} = 0$ $P_{15} = \frac{C_{15}}{C_{15} + C_{51}} = \frac{0}{0+3} = \frac{0}{3} = 0$
- $P_{22} = \frac{C_{22}}{C_{22} + C_{22}} = \frac{0}{0+0} = 0$ $P_{23} = \frac{C_{23}}{C_{23} + C_{32}} = \frac{0}{0+3} = \frac{0}{3} = 0$
- $P_{24} = \frac{C_{24}}{C_{24} + C_{42}} = \frac{0}{0+3} = \frac{0}{3} = 0$ $P_{25} = \frac{C_{25}}{C_{25} + C_{52}} = \frac{0}{0+3} = \frac{0}{3} = 0$
- $P_{33} = \frac{C_{33}}{C_{33} + C_{33}} = \frac{0}{0+0} = 0$ $P_{44} = \frac{C_{44}}{C_{44} + C_{44}} = \frac{0}{0+0} = 0$
- $P_{55} = \frac{C_{55}}{C_{55} + C_{55}} = \frac{0}{0+0} = 0$

Setelah didapatkan hasil probabilitasnya maka langkah selanjutnya mengurutkan ranking hasilnya berdasarkan peluang dari yang paling tinggi sampai yang rendah. Hasil akhir penilaian dapat dilihat pada Tabel 6 berikut.

Tabel 6. Ranking Akhir

Ranking	Subjek
1	D
2	C
3	E
4	B
5	A

Ini merupakan hasil akhir dari langkah CJ yang telah diurutkan dari ranking satu sampai terakhir. Sehingga jawaban terbaik diberikan oleh Siswa D, diikuti C, E, B dan terakhir A. Dari sana dapat dilihat walaupun dengan penggunaan rubrik penilaian konvensional, kelima memperoleh nilai yang sama, tapi jawaban tersebut masih dapat diurutkan kualitasnya.

CJ yang penulis bahas disini hanya terbatas pada soal uraian (essay). Soal uraian yang digunakan dalam CJ hanya 1 buah soal dengan mengambil 3 orang ahli dalam bidang matematika. CJ ini juga bisa digunakan pada soal uraian yang terdiri dari banyak butir soal dan banyak ahli untuk melakukan perbandingan pada bidang keahliannya dengan menggunakan bantuan web untuk banyak subjek yang dinilai salah satunya yaitu FACETS software sebagai suatu alternative penilaian untuk memudahkan dalam penilaian dengan banyak butir soal dan banyak ahli yang dipakai. Sedemikian sehingga memudahkan untuk melakukan penilaian pada tes-tes yang lebih besar. Semakin banyak ahli, soal yang diteskan serta siswa yang digunakan dalam penilaian CJ ini maka semakin kuat tingkat validitas dan reabilitas yang didapatkan.

IV. SIMPULAN

Dari pembahasan yang dilakukan dapat ditarik simpulan sebagai berikut. Model CJ mampu menghasilkan suatu alternatif penilaian yang tepat untuk memecahkan permasalahan dalam penilaian. Dengan memakai model CJ kesalahan-kesalahan yang dilakukan dalam penilaian dapat diminimalisir. CJ juga memberikan perbandingan dalam kemampuan pemahaman konsep siswa dengan lebih baik. Hal ini disebabkan karena penilaian yang dilakukan tidak hanya dilihat pada jawaban akhir saja, tetapi juga pada saat proses berlangsung.

Dalam makalah ini hanya membahas mengenai model *Comparative Judgement* (CJ) dalam bentuk tes uraian. Bagi pembaca yang ingin mengembangkan model

MANDALIKA

Mathematics and Educations Journal
Volume 1 Nomor 1, Juni 2019

comparative judgement (CJ) lebih luas lagi dalam mengukur kemampuan pemecahan masalah berfikir kreatif, penugasan proyek atau bentuk assessment lainnya atau dengan menggunakan bantuan software dengan jumlah tes yang lebih besar.

ACKNOWLEDGEMENT

Penelitian ini dilakukan dengan pendanaan mandiri. Penulis mengucapkan terimakasih kepada guru dan siswa kelas XI AP4 SMK Negeri 2 Singaraja serta kepada ahli yang telah membantu proses penilaian CJ ini dilakukan.

DAFTAR PUSTAKA

- [1] Laming, D. (2004). Marking university examinations: some lessons from psychophysics. *Psychophysics Learning and Teaching*, 3, 89-96.
- [2] Jones, I., Swan, M., & Pollit, A. (2014). Assessing mathematical problem solving using comparative judgement. *International Journal of Science and Mathematics Education*, 13, 151-177.
- [3] Suherman, E. (2003). *Strategi Pembelajaran Matematika Kontemporer*. Jakarta: IMSTEP Universitas Pendidikan Ganesha.
- [4] Hamalik, O. (2001). *Perencanaan Pengajaran Berdasarkan Pendekatan Sistem*. Bandung: Bumi Aksara
- [5] NCTM, (2000). Principles and Standards for School Mathematics.. http://www.netm.org/uploadedFile/Standards_and_Positions/PSSM_ExeexutiveSummary.pdf. Diakses pada tanggal 3 juli 2017
- [6] Gulo, W. (2002). *Strategi Belajar Mengajar*. Jakarta: PT Gramedia Widiasarana Indonesia.
- [7] D'Arcy, J. (1997). Comparability Studies between Modular and Non-Modular Syllabuses in GCE Avanced Biology, English Literature and Mathematics in the 1996 summer examination Standing Committen on research on behalf of the Joint Forum for the GCSE and GCE.
- [8] Pollit, A. (2004). *Let's stop marking exams, International Association for educational Assessment Comference*. Philadelphia PA.
- [9] Meadows, M., & Bilington, L. (2005). *A review of the literature on marking reability*. Manchester: AQA. No. RPA_05_MM_WP_05. A report produced for the National Assessment Agency.
- [10] Thurstone, L.L. (1927). *A law of comparative judgement*. Psychological Review, v34 n4 pp. 273-286.
- [11] Whitehouse, C. (2010). *Using Adaptive Comparative Judgement to Obtain a Highly Reliable Rank Order in Summative Assessment*. Guildford: AQA.
- [12] Wiersma, W. (1991). *Research method in education*. Fifth Edition. Boston: Allyn and Bacom.
- [13] Gay, L. (1987). *Education research, Competencies for analysis and application*. Third edition. Colombus: Merrill Publishing Company.

MANDALIKA

Mathematics and Educations Journal

Volume 1 Nomor 1, Juni 2019

- [14] Long, T. J., Convey, J. J., & Chwalek, A. R. (1985). *Completing dissertation in the behavioral sciences and education*. London: Jossey-Bass Publishers.
- [15] Tsukida, K. and Gupta, M. R. (2011). *How to Analyze Paired Comparison Data. Technical Report UWEETR-2011-0004*. Department of Electrical Engineering University of Washington.