



ANALISIS BUTIR SOAL MATA KULIAH ALGORITMA PEMROGRAMAN

Yasinta Lisa¹

¹Prodi Pendidikan Komputer, STKIP Persada Khatulistiwa Sintang

Email: yasintalisa@gmail.com

Informasi Artikel	ABSTRACT
<p>Riwayat artikel: Disubmit : 5 Mei 2021 Direvisi : 20 Mei 2021 Diterima : 5 Juni 2021 Dipublikasi : 12 Juni 2021</p> <p>Keywords: <i>validity, reliability, level of difficulty, discriminatory power, cognitive level, item analysis.</i></p>	<p><i>This study aims to describe item analysis based on Bloom's taxonomy on the mid-semester test instrument for Programming Algorithms and Data Structures and its relation to student perceptions. The sample of the trial target is the Mid-Semester Test Question for the 2020/2021 Academic Year of the STKIP Persada Khatulistiwa Sintang Computer Education Study Program. The research method uses descriptive quantitative research. The results of the data analysis showed that, the results of the validity test stated that of the 30 questions (90%) there were 27 questions which were declared valid and 3 (10) questions invalid. Test the reliability of the test items obtained in the high category. The difficulty level of the questions ranges from 83.3% (Easy Questions), 10% (Medium Questions), and 6.7% (Difficult Questions). The Distinguishing Power of Questions in the moderate category and the Detractor Index have mostly functioned well. The distribution of cognitive levels is mostly at C1 (50%), C2 (23.3%), C3 (6.7%), C4 (10%), C5 (6.7%), and C6 (3.35) . This result is not proportional for the higher education level.</i></p>
<p>Kata Kunci: validitas, reliabilitas, tingkat kesukaran, daya pembeda, tingkatan kognitif, analisis butir soal.</p>	<p>ABSTRAK</p> <p>Penelitian ini bertujuan untuk mendeskripsikan analisis butir soal berdasarkan taksonomi bloom pada instrumen tes tengah semester mata kuliah Algoritma Pemrograman dan Struktur Data dan kaitannya dengan persepsi mahasiswa. Sampel sasaran uji coba yaitu Soal Tes Tengah Semester Tahun Akademik 2020/2021 Prodi Pendidikan komputer STKIP Persada khatulistiwa sintang. Metode penelitian menggunakan penelitian deskriptif kuantitatif. Hasil analisis data menunjukkan bahwa, hasil uji coba validitas menyatakan bahwa dari 30 soal (90%) terdapat 27 soal yang dinyatakan valid dan 3 (10) soal tidak valid. Uji Reliabilitas soal tes diperoleh pada kategori tinggi. Tingkat Kesukaran Soal berkisar 83,3% (Soal Mudah), 10% (Soal Sedang), dan 6,7% (Soal Sukar). Daya Pembeda Soal pada kategori cukup dan Indeks Pengecoh sebagian besar telah berfungsi dengan baik. Sebaran tingkatan kognitif sebagian besar berada pada C1 (50%), C2 (23,3%), C3 (6,7%), C4 (10%), C5 (6,7%), dan C6 (3,35). Hasil ini belum proporsional untuk tingkatan Perguruan Tinggi.</p>



PENDAHULUAN

Hasil belajar merupakan unsur penting dalam mengukur kemampuan peserta didik. Menurut Gronlund yang dikutip oleh Purwanto (2011) menyatakan bahwa hasil belajar yang diukur merefleksikan tujuan pengajaran. Selain menguasai materi dan metode mengajar, Guru juga dituntut untuk mampu menguasai teknik evaluasi yang benar karena guru memiliki tugas untuk melakukan penilaian terhadap proses dan hasil belajar peserta didik yang mempengaruhi keputusan tentang peningkatan kualitas pencapaian belajarnya.

Evaluasi pembelajaran merupakan proses mengumpulkan, menganalisa dan menginterpretasikan informasi yang dijadikan dasar untuk mengetahui kemajuan, perkembangan, dan pencapaian belajar siswa serta keefektifan pengajaran oleh guru. Secara umum tujuan evaluasi pembelajaran (Ratnawulan & Rusdiana, 2014) adalah : a) untuk mengetahui keefektifan dan efisiensi sistem pembelajaran baik berupa tujuan, materi, metode, media, sumber belajar, lingkungan maupun sistem penilaian itu sendiri, serta b) untuk menghimpun bahan keterangan (data) yang dijadikan sebagai bukti mengenai taraf kemajuan anak didik dalam mengalami proses pendidikan selama jangka waktu tertentu.

Pada dasarnya kegiatan analisis butir soal sangat penting dilakukan oleh guru. Hal tersebut bertujuan agar Guru mampu mengetahui kehandalan dan validitas dari soal yang telah mereka susun. Selain itu, melalui analisis butir soal akan mendapatkan umpan balik yang nantinya dapat dijadikan acuan dalam pengambilan keputusan. Soal yang masuk kategori layak dapat langsung digunakan, soal dengan kategori kurang layak dapat direvisi, sedangkan soal dengan kategori tidak layak dapat digugurkan sehingga harus dihapus atau diganti dengan soal lainnya (Karkal & Kundapur, 2016), (Odukoya, Adekeye, & Igbino, May 2018). Namun sayangnya, pengetahuan dan keterampilan guru dalam melakukan analisis butir soal masih tergolong rendah (Sumiati, Widiastuti, & Suhud, Juli 2018). Sebagian besar guru cenderung membuat soal sesuai dengan kompetensi inti dan kompetensi dasar yang telah ditentukan tanpa mengukur terlebih dahulu apakah siswa telah

Penelitian ini difokuskan pada penyusunan dan analisis butir soal. Penelitian dibatasi hanya pada aspek kognitif dengan instrumen berupa tes pilihan ganda. Aspek kognitif dipilih karena mata kuliah Algoritma Pemrograman dan Struktur Data memerlukan pemahaman konsep yang baik tentang teori-teori dasar pemrograman. Tes pilihan ganda dipilih karena dapat mewakili banyak indikator dan waktu mengerjakan tes yang singkat.



METODE PENELITIAN

Metode penelitian yang digunakan adalah metode penelitian deskriptif. Metode penelitian deskriptif adalah metode penelitian yang digunakan untuk menggambarkan maupun menganalisis suatu hasil penelitian tetapi tidak digunakan untuk membuat kesimpulan yang luas. Penyusunan kisi-kisi alat tes diawali dengan mengkaji indikator pencapaian materi ajar. Informasi-informasi yang dikumpulkan dijadikan bahan untuk menyusun alat tes. Penyusunan kisi-kisi alat tes mengacu pada indikator yang akan dicapai dan tingkat kognitif soal. Langkah selanjutnya menulis soal pilihan ganda yang disesuaikan dengan kisi-kisi yang telah dibuat. Selanjutnya dilakukan uji coba terhadap alat tes yang hasilnya digunakan sebagai sarana untuk memperoleh data-data terkait alat tes yang akan dianalisis. Instrumen penelitian meliputi soal tes pilihan ganda ujian tengah semester, kunci jawaban. Dan lembar jawaban hasil pekerjaan mahasiswa. Pelaksanaan tes dilakukan selama 50 menit. Hasil penelitian dianalisis validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecoh menggunakan bantuan *software* Anates V4 dengan taraf signifikansi sebesar 5%.

HASIL DAN PEMBAHASAN

Hasil penyusunan butir soal berdasarkan kompetensi/ materi disajikan pada Tabel 1 berikut :

Tabel 1. Rekapitulasi Butir Soal per Materi

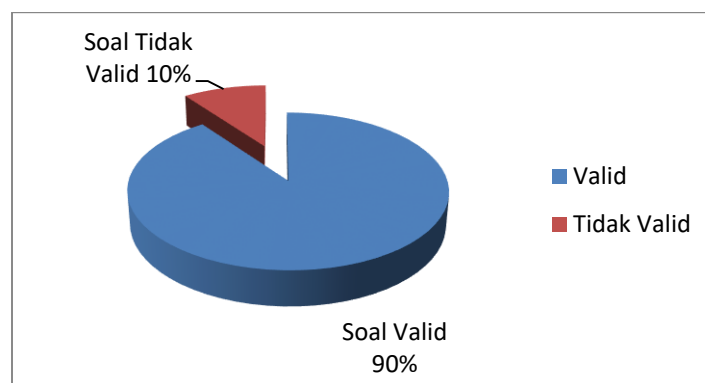
No	Kompetensi/Materi	Nomor Soal	Jumlah	Prosentase
1	Konsep Dasar Algoritma	1, 2, 4, 5, 7, 11.	6	20%
2	Konsep Dasar Struktur Data	6, 7, 9, 10, 13, 19, 24, 25, 26, 27.	10	33,3%
3	Mengenal Bahasa C++	3, 12, 14, 15, 16, 17, 18, 20, 21, 22, 23, 30.	12	40%
4	Penggunaan Seleksi kondisi	28, 29.	2	6,7%
Total			30	100%

Analisis validitas butir soal dilakukan dengan menggunakan program Anates. Hasil perhitungan validitas butir soal dikorelasikan dengan nilai r_{tabel} . Jika diperoleh $r_{hitung} > r_{tabel}$, maka soal tersebut dinyatakan valid. Jumlah peserta yang mengikuti tes pada mata kuliah Algoritma Pemrograman dan Struktur Data sebanyak 56 orang. Nilai r_{tabel} pada taraf signifikan 5% dengan $N = 56$ adalah 0,2632. Hasil rekapitulasi uji validitas butir soal disajikan pada Tabel 4 berikut ini :



Tabel 2. Klasifikasi Berdasarkan Validitas Butir Soal

No	Validitas	Nomor Soal	Jumlah	Prosentase
1	$\geq 0,2632$ (Valid)	1, 2, 4, 6, 7, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30.	27	90%
2	$< 0,2632$ (Tidak Valid)	3, 5, 8.	3	10%
Total			30	100%



Gambar 1. Diagram distribusi soal berdasarkan Validitas Butir Soal

Berdasarkan Tabel 2 dan Gambar 1, dapat diketahui masing-masing prosentase Validitas Butir soal dengan kategori Valid sebesar 90% dengan jumlah soal 27 butir soal dan kategori Tidak Valid sebesar 10% dengan jumlah soal 3 butir soal. Karakteristik pertama dan memiliki peranan penting dalam instrumen evaluasi, yaitu validitas butir soal (validity) (Sukardi, 2008). Suatu instrumen tes dapat dikatakan valid apabila dapat mengukur dengan tepat dan benar. Oleh karena itu soal yang baik adalah soal yang Valid yaitu soal yang mempunyai dukungan yang besar terhadap skor total. Skor pada item menyebabkan skor total menjadi tinggi atau rendah. Sebuah item memiliki validitas yang tinggi jika skor pada item memiliki kesejajaran dengan skor total. Kesejajaran ini dapat diartikan sebagai korelasi sehingga untuk mengetahui validitas item digunakan rumus korelasi (Arikunto, 2006).

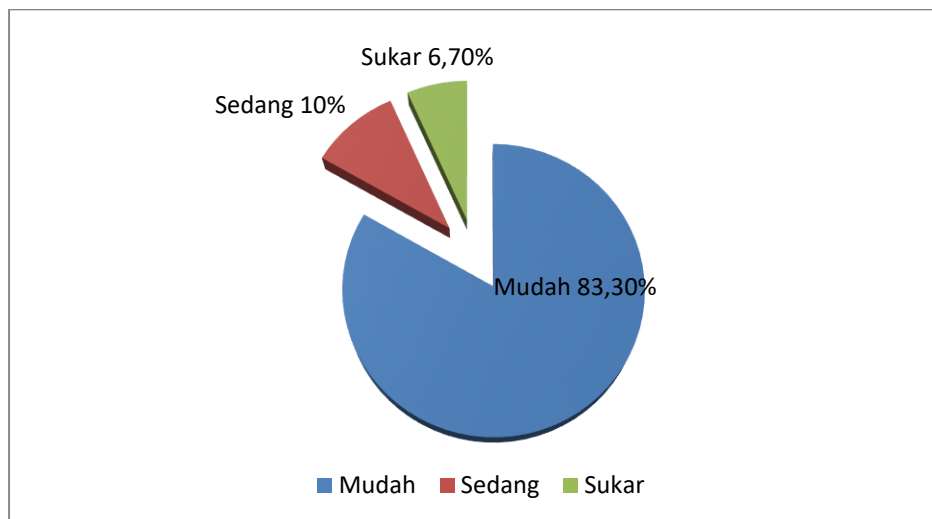
Pengujian reliabilitas tes dilakukan dengan menggunakan program Anates. Jika hasil uji Reliabilitas tes $\geq 0,88$ maka tes tersebut dinyatakan memiliki Reliabilitas yang tinggi. Pada perhitungan Uji Reliabilitas butir soal menggunakan program Anates diperoleh nilai sebesar 0,94 jika dibandingkan dengan batas ambang Reliabilitas tes 0,88 maka akan menghasilkan $0,94 > 0,88$ yang berarti Reliabilitas instrumen tes yang digunakan berada pada kategori tinggi atau reliabel.



Pengolahan data validitas konstruk yang terdiri atas tingkat kesukaran, daya pembeda, dan indeks pengecoh. Analisis validitas konstruk dilakukan dengan menggunakan program Anates. Tingkat kesukaran diklasifikasikan berdasarkan kriteria tingkat kesukaran, disajikan pada Tabel 2 sebagai berikut :

Tabel 3. Klasifikasi soal Berdasarkan Tingkat Kesukaran Soal

No	Kriteria	Nomor Soal	Jumlah	Prosentase
1	Mudah (0,71-1,00)	1, 2, 3, 4, 6, 7, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 20, 21, 22, 23, 24, 26, 27, 28, 30.	25 butir	83,3%
2	Sedang (0,31-0,70)	8, 25, 29.	3 butir	10%
3	Sukar (0,00-0,30)	5, 19,	2 butir	6,7%
Total			30	100%



Gambar 2. Diagram Distribusi Tingkat Kesukaran Soal

Berdasarkan Tabel 3 dan Gambar 2, diketahui bahwa total soal tes sebanyak 30 butir soal didapatkan indeks tingkat kesukaran yang menunjukkan bahwa soal dengan kategori “Mudah” memiliki prosentase 83,3% dengan jumlah 25 butir soal. Sedangkan soal dengan kategori “Sedang” memiliki prosentase sebesar 10% dengan jumlah 3 butir soal dan soal dengan kategori “Sukar” memiliki prosentase sebesar 6,7% dengan jumlah 2 butir soal. Hasil tersebut memang belum memenuhi kriteria instrumen evaluasi yang baik seperti yang dikemukakan oleh Sudjana (Sudjana, 2017), dimana salah satu proporsi yang baik sebuah instrumen evaluasi harus memenuhi

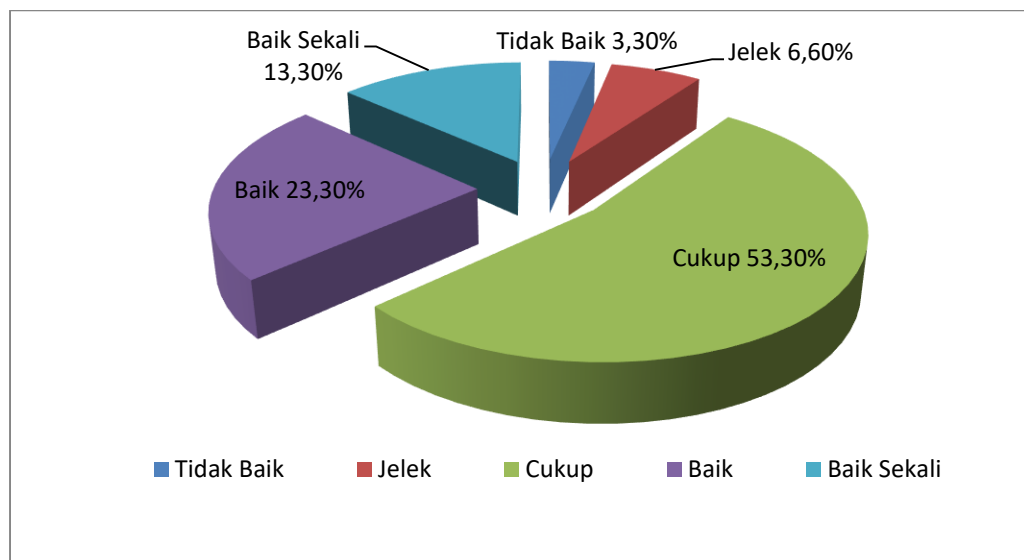


perbandingan 3-5-2 atau 30% butir soal dengan kategori mudah, 50% butir soal kategori sedang, dan 20 % butir soal dengan kategori sukar.

Analisis daya pembeda dilakukan dengan membagi data ke dalam dua kelompok yaitu kelompok atas dan kelompok bawah. Daya pembeda soal diklasifikasikan pada kriteria daya pembeda soal dengan hasil rekapitulasi yang disajikan pada Tabel 3 sebagai berikut:

Tabel 4. Klasifikasi soal Berdasarkan Daya Pembeda

No	Kriteria	Nomor Soal	Jumlah	Prosentase
1	Tidak Baik ($DP < 0$)	5	1	3,3%
2	Jelek (0,00 – 0,20)	8, 14.	2	6,6%
3	Cukup (0,21 – 0,40)	2, 3, 4, 6, 7, 9, 10, 11, 12, 13, 15, 16, 23, 24, 26, 30.	16	53,3%
4	Baik (0,41 – 0,70)	1, 17, 18, 20, 22, 28, 28.	7	23,3%
5	Baik Sekali (0,71 – 1,00)	19, 21, 25, 27.	4	13,3%
Total			30	100%



Gambar 3. Diagram distribusi soal berdasarkan Daya Pembeda

Berdasarkan Tabel 4 dan Gambar 3, diketahui bahwa dari keseluruhan soal tes sebanyak 30 butir soal didapatkan bahwa prosentase tertinggi didapatkan pada soal dengan Daya Pembeda pada kategori “Cukup” memiliki prosentase sebesar 53,3% dengan jumlah sebanyak 16 butir soal. Soal dengan kategori “Baik” memiliki prosentase sebesar 23,3% dengan jumlah sebanyak 7 butir soal. Soal dengan

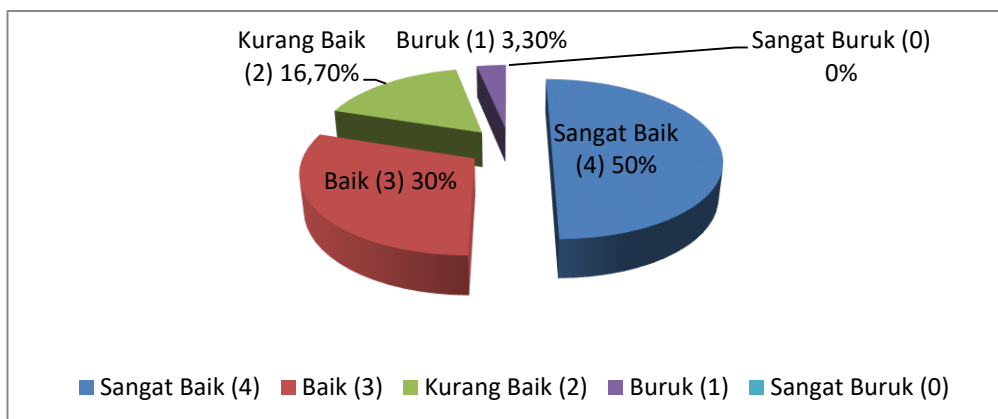


kategori “Baik Sekali” memiliki prosentase se.besar 13,3% dengan jumlah sebanyak 4 butir soal. Soal dapat dikatakan tidak memenuhi kriteria apabila terletak pada kategori tidak baik atau jelek. Soal yang tidak memenuhi kriteria berarti soal tersebut belum mampu membedakan siswa yang memiliki

Analisis indeks pengecoh dilakukan menggunakan program Anates. Indeks pengecoh soal dikatakan berfungsi apabila lebih dari 5% peserta tes menjawab/memilih pada setiap alternatif jawaban, dan kurang dari 10% peserta tes tidak menjawab/omit. Indeks pengecoh soal dapat diklasifikasikan pada kriteria berfungsi dan tidak berfungsi (Arikunto, 2012). Rekapitulasi Indeks Pengecoh soal disajikan pada Tabel 4 berikut ini:

Tabel 5. Klasifikasi soal Berdasarkan Indeks Pengecoh

No	Kriteria	Nomor Soal	Jumlah	Prosentase
1	4 (Sangat Baik)	2, 4, 6, 7, 12, 15, 17, 18, 20, 21, 24, 26, 28, 29, 30.	15	50%
2	3 (Baik)	5, 9, 11, 13, 14, 16, 23,25, 27.	9	30%
3	2 (Kurang Baik)	1, 3, 10, 19, 22.	5	16,7%
4	1 (Buruk)	8.	1	3,3%
5	0 (Sangat Buruk)	-	-	-
Total			30	100%



Gambar 4. Diagram distribusi soal berdasarkan Indeks Pengecoh

Berdasarkan Tabel 5 dan Gambar 4, dapat diketahui masing-masing prosentase efektivitas pengecoh pada setiap kategori. Hasil diatas menunjukkan bahwa dari 120 pengecoh yang ada pada 30 butir soal, sejumlah 15 butir soal atau 50% pengecoh masuk dalam kategori “Sangat Baik”, 30%

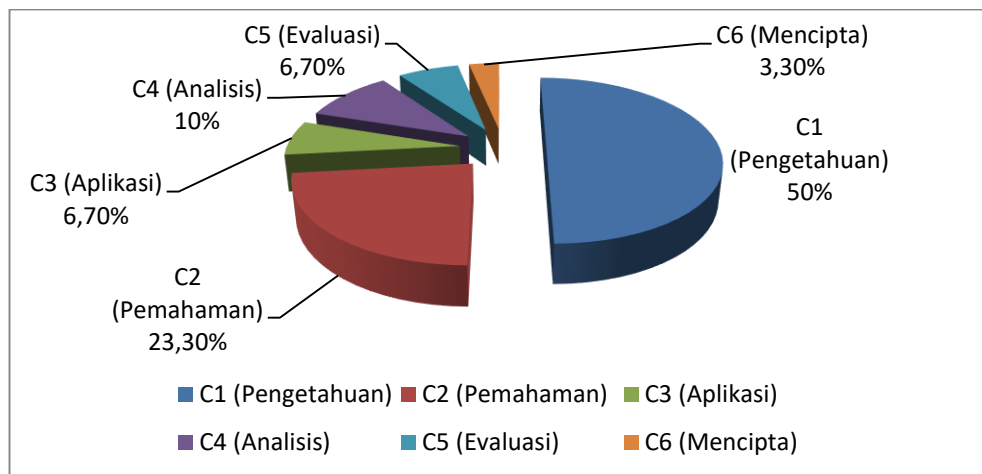


pengecoh masuk kategori “Baik”, 16,7% pengecoh masuk kategori Kurang Baik, 3,3% pengecoh masuk kategori “Buruk” dan tidak ada pengecoh yang berkategori “Sangat Buruk”.

Pengklasifikasian soal tes berdasarkan tingkatan kognitif dilakukan dengan menelaah soal berdasarkan indikator ketercapaian dari kisi-kisi soal yang dibuat. Rekapitulasi sebaran tingkatan kognitif soal tes disajikan pada Tabel 5 berikut ini :

Tabel 6. Klasifikasi soal Berdasarkan Tingkatan Kognitif Taksonomi Bloom

No	Kriteria	Nomor Soal	Jumlah	Prosentase
1	C1 (Pengetahuan)	1, 2, 3, 12, 13, 14, 15, 16,17, 20, 25, 26, 27, 28, 30.	15	50%
2	C2 (Pemahaman)	4, 11, 28, 19, 21, 22, 23,	7	23,3%
3	C3 (Aplikasi)	5, 24.	2	6,7%
4	C4 (Analisis)	6, 7, 10,	3	10%
5	C5 (Evaluasi)	8, 9,	2	6,7%
6	C6 (Mencipta)	29.	1	3,3%
Total			30	100%



Gambar 5. Diagram distribusi soal berdasarkan Tingkatan Kognitif Taksonomi Bloom

Berdasarkan Tabel 6 dan Gambar 5 dapat diketahui bahwa prosentase tingkatan kognitif yang paling tinggi yaitu sebesar 50% setara dengan 15 butir soal berada pada tingkatan C1 (Pengetahuan), 23,3% setara dengan 7 butir soal untuk tingkatan C2 (Pemahaman), 10% setara dengan 3 butir soal untuk C4 (Analisis), masing-masing 6,7% setara dengan 2 butir soal untuk tingkatan C3 (Aplikasi) dan C5 (Evaluasi) dan 3,3% setara dengan 1 butir soal untuk tingkatan C6 (Mencipta).

Tingkat kesulitan soal berdasarkan Taksonomi Bloom (Kemendikbud, 2013) dijelaskan bahwa ada 6 tingkatan Taksonomi Bloom diantaranya ingatan (C1), pemahaman (C2), aplikasi (C3), analisis



(C4), evaluasi (C5), dan mencipta (C6). Sedangkan tingkatan sintesis tidak terdapat dalam diagram Taksonomi Bloom tersebut. Sebaran tingkatan Taksonomi Bloom untuk tiap tingkatan sekolah SD terdapat 2 tingkatan yaitu ingatan (C1), dan pemahaman (C2), tingkat SMP terdapat 3 tingkatan yaitu ingatan (C1), pemahaman (C2), dan aplikasi (C3), tingkatan SMA terdapat 4 tingkatan yaitu ingatan (C1), pemahaman (C2), aplikasi (C3), dan analisis (C4), sedangkan pada tingkatan PT (Perguruan Tinggi), terdapat 6 tingkatan yaitu ingatan (C1), pemahaman (C2), aplikasi (C3), analisis (C4), evaluasi (C5) dan mencipta (C6). Distribusi soal tengah semester mata kuliah Algoritma pemrograman dan struktur data belum sesuai dengan distribusi tingkat kesulitan soal yang baik, karena jumlah prosentase distribusi tingkat kesulitan soal yang paling banyak adalah jenjang pengetahuan (C1) dan pemahaman (C2) sebanyak 73,3%. Sesuai dengan tingkatan Perguruan Tinggi Ranah Taksonomi Bloom yang paling dieksplor adalah analisis (C4), evaluasi (C5) dan mencipta (C6) akan tetapi pada tes ujian tengah semester ini hanya memiliki prosentase 20% dari keseluruhan soal tes.

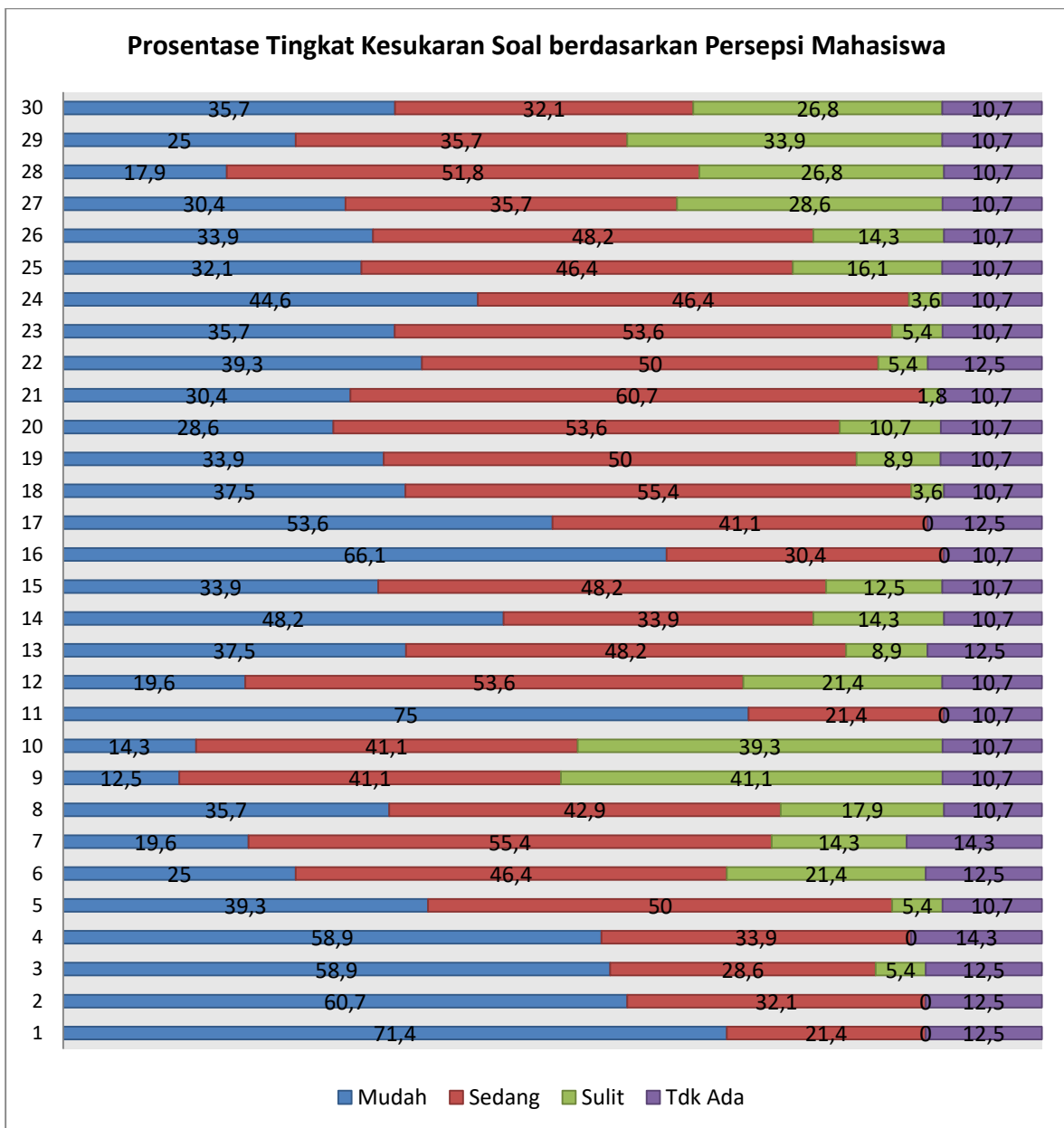
Penilaian mahasiswa berdasarkan persepsinya tentang kesulitan soal ditinjau dari kemampuan masing-masing dilakukan dengan angket persepsi soal yang diisi oleh mahasiswa saat mengerjakan soal tes. Rekapitulasi sebaran persepsi mahasiswa tentang kesulitan soal dirangkum dalam grafik pada Gambar 6. Berdasarkan Gambar 6 didapat bahwa Persepsi mahasiswa tentang kesulitan soal tes bahwa hampir seluruh butir soal mendapatkan persepsi yang berbeda oleh seluruh mahasiswa. Setiap soal memiliki variasi tingkat kesulitan dari mudah, sedang dan sulit bahkan ada sebagian kecil mahasiswa tidak bisa menentukan tingkatan kesulitan berdasarkan kemampuan dirinya sendiri.

Berpatokan pada hasil Uji Tingkat Kesukaran soal jika dibandingkan dengan persepsi mahasiswa tentang kesukaran soal dimana untuk butir soal 8, 25 dan 29 merupakan soal dengan tingkat kesukaran “Sedang” akan tetapi masih terdapat sebagian mahasiswa dengan rata-rata 22,63% mahasiswa menganggap bahwa soal tersebut dirasa sulit, rata-rata 30,93% mahasiswa menganggap soal-soal tersebut mudah dikerjakan, serta rata-rata 41,67% mahasiswa menganggap pada tingkatan kesukaran yang sebenarnya yaitu sedang dan sisanya 10,7% mahasiswa tidak mampu menentukan tingkatan kesulitan soal berdasarkan kemampuannya.

Pada butir soal dengan hasil uji Tingkat Kesukaran Soal berada pada tingkat “Sukar” yaitu nomor butir 5 dan 19 diperoleh data sebanyak rata-rata 36,6% dari mahasiswa menganggap bahwa soal tersebut merupakan soal mudah, rata-rata 50% dari mahasiswa menganggap sebagai soal sedang, rata-rata 7,15% mahasiswa menganggap soal sukar dan rata-rata 10,7% mahasiswa tidak mampu menentukan tingkat kesukaran soal berdasarkan kemampuan diri.



Pada butir soal dengan hasil uji Tingkat Kesukaran Soal berada pada tingkat “Mudah” yaitu nomor 1, 2, 3, 4, 6, 7, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 20, 21, 22, 23, 24, 26, 27, 28, dan 30 diperoleh data sebesar rata-rata 39,6% dari mahasiswa menganggap soal tersebut mudah, sebesar rata-rata 42,6% mahasiswa menganggap soal sedang, sebesar rata-rata 12,2% mahasiswa menganggap soal sulit, dan sebesar rata-rata 11,5% mahasiswa tidak mampu menentukan tingkat kesukaran soal berdasarkan kemampuan diri.



Gambar 6. Persepsi Mahasiswa berdasarkan Tingkat Kesukaran Soal



Beberapa hal yang menjadi sebab persepsi mahasiswa tentang kesulitan soal kemungkinan disebabkan oleh rasa kurang percaya diri dan pemahaman yang masih belum cukup. Selain itu mata kuliah Algoritma Pemrograman dan Struktur Data adalah salah satu mata kuliah yang menuntut lebih banyak tentang pemahaman analitis matematis mahasiswa. Untuk memiliki pemahaman analitis matematis terlebih dahulu memiliki kemampuan koneksi matematis yang baik. Untuk dapat melakukan koneksi terlebih dahulu harus mengerti dengan permasalahannya dan untuk dapat mengerti permasalahan harus mampu membuat koneksi dengan topik- topik yang terkait.

Kegiatan yang dapat menunjukkan atau tergolong kemampuan koneksi matematis ada 4 yaitu :

(a) Mengaplikasikan dan menghubungkan antar konsep dalam suatu materi matematika, (b) Mengaplikasikan dan menghubungkan konsep antar materi matematika, (c) Mengaplikasikan konsep matematika dalam menyelesaikan masalah matematika yang berkaitan dengan kehidupan sehari- hari, (d) Mengaplikasikan dan menghubungkan konsep pada materi matematika dengan konsep pada materi ilmu selain matematika (Pijarno, 2014). Kurangnya kemampuan untuk mengenali kemampuan koneksi matematis dalam diri akan mempengaruhi persepsi individu terhadap kemampuan belajarnya.

SIMPULAN

Berdasarkan hasil pengujian dan analisa dalam pembahasan, maka dapat disimpulkan Hasil analisis data menunjukkan bahwa, soal tes tengah semester Mata Kuliah Algoritma Pemrograman dan Struktur Data layak digunakan sebagai alat penilai kompetensi. Hal ini dibuktikan dari 30 soal pilihan ganda terdapat 90% soal yang masuk pada kategori valid, sedangkan 10% soal masuk kategori tidak valid. hasil uji coba validitas menyatakan bahwa dari 30 soal (90%) terdapat 27 soal yang dinyatakan valid dan 3 (10) soal tidak valid. Kemudian untuk Uji Reliabilitas soal tes diperoleh nilai sebesar 0,94 jika dibandingkan dengan batas ambang Reliabilitas tes 0,88 maka akan menghasilkan $0,94 > 0,88$ yang berarti Reliabilitas instrumen tes yang digunakan berada pada kategori tinggi atau reliabel. Kemudian hasil uji Tingkat Kesukaran Soal terdapat 83,3% soal masuk kategori Mudah, 10% soal masuk kategori Sedang, dan 6,7% soal masuk kategori Sukar. Hasil tersebut memang belum memenuhi kriteria instrumen evaluasi yang baik jika ditinjau dari tingkat kesukaran soal . selanjutnya Daya Pembeda Soal didapat kategori “Cukup” memiliki prosentase sebesar 53,3% dengan jumlah sebanyak 16 butir soal. Soal dengan kategori “Baik” memiliki prosentase sebesar 23,3% dengan jumlah sebanyak 7 butir soal. Soal dengan kategori “Baik Sekali” memiliki prosentase se.besar 13,3% dengan jumlah sebanyak 4 butir soal. Kemudian untuk efektivitas pengecoh terdapat 120 pengecoh dari 30 butir soal didapatkan 50% pengecoh masuk dalam kategori “Sangat Baik”, 30% pengecoh



masuk kategori “Baik”, 16,7% pengecoh masuk kategori Kurang Baik, 3,3% pengecoh masuk kategori “Buruk” dan tidak ada pengecoh yang berkategori “Sangat Buruk”. Indeks Pengecoh sebagian besar telah berfungsi dengan baik. Sebaran tingkatan kognitif sebagian besar berada pada C1 (50%), C2 (23,3%), C3 (6,7%), C4 (10%), C5 (6,7%), dan C6 (3,35). Hasil ini belum proporsional untuk tingkatan Perguruan Tinggi.

Berdasarkan hasil penelitian dan simpulan yang disampaikan beberapa saran yang diajukan antara lain: (1) Penyusunan instrumen tes harus lebih memperhatikan pada indikator ketercapaian materi sehingga tingkatan kognitif yang diinginkan dapat disesuaikan dengan kebutuhan pembelajaran. (2) Dilakukan analisis butir soal terlebih agar instrumen tes dapat diketahui kevalidannya sebelum memberikan tes atau ujian kepada peserta didik. (3) Perlu dilakukan peningkatan pemahaman dan kemampuan dalam menyusun instrumen tes agar butir-butir soal yang dihasilkan menjadi lebih berkualitas. (4) Pelaksanaan tes harus dilakukan secara adil terhadap seluruh peserta didik dengan proporsi soal seimbang sesuai kebutuhan yang dipelajari serta tingkatan sekolah.

DAFTAR RUJUKAN

- Arikunto. (2006). *Dasar-Dasar Evaluasi Pendidikan*. Jakarta: Bumi Aksara.
- Karkal, Y. R., & Kundapur, G. S. (2016). Item analysis of multiple choice questions of undergraduate pharmacology examination in an International Medical School in India. *Journal of Dr. NTR University of Health Sciences* 5(3) , 183-186.
- Kemendikbud. (2013). *Implementasi Kurikulum 2013*. Jakarta: Kemendikbud.
- Odukoya, J. A., Adekeye, O., & Igbino, A. O. (May 2018). Item analysis of university-wide multiple choice objective examinations: the experience of a Nigerian private university. *Qual Quant* vol. 52, no. 3, , pp. 983-997.
- Pijarno, R. (2014). Analisis Kemampuan Koneksi Matematis Pada Materi Bangun. *Jurnal Pendidikan Matematika (online)*, <http://fkip.ump.ac.id/index.php/artikelilmiah/summary/8-pendidikan-matematika/382-analisis-kemampuankoneksi-matematis-pada-materi-bangun-ruang-sisi-lengkung-kelas-ix-smpnegeri-1-sampun> .
- Purwanto. (2011). *Evaluasi hasil belajar. Cetakan ke-2*. Yogyakarta: Pustaka Belajar.
- Ratnawulan, E., & Rusdiana, H. A. (2014). *Evaluasi Pembelajaran*. Bandung: Pustaka Setia.
- Sudjana, N. (2017). *Penilaian hasil proses belajar mengajar, 20th ed*. Bandung: Remaja Rosdakarya.



Sukardi. (2008). *Penilaian dan Pengukuran*. Jakarta: Bumi Aksara.

Sumiati, A., Widiastuti, U., & Suhud, U. (Juli 2018). Workshop Teknik Menganalisis Butir Soal dalam Meningkatkan Kompetensi Guru di SMK Cileungsi Bogor. *Jurnal Pemberdayaan Masyarakat Madani (JPMM)*, E-ISSN: 2580-4332 Vol 2 No 1 , 136-153.