



ANALISIS SOAL *INTERNATIONAL COMPETITIONS AND ASSESSMENTS FOR SCHOOLS (ICAS)* DENGAN MENGGUNAKAN METODE *ITEM RESPONSE THEORY (IRT)* DAN *CLASSICAL TEST THEORY (CTT)*

Skripsi

**disusun sebagai salah satu syarat
untuk memperoleh gelar Sarjana Pendidikan
Program Studi Pendidikan Biologi**

oleh

**Mahanani
4401408105**

**JURUSAN BIOLOGI
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS NEGERI SEMARANG**

2015

PERNYATAAN KEASLIAN SKRIPSI

Saya menyatakan dengan sebenar-benarnya bahwa skripsi saya yang berjudul "Analisis Soal *International Competitions and Assessments for Schools* (ICAS) dengan Menggunakan Metode *Item Response Theory* (IRT) dan *Classical Test Theory* (CTT)" disusun berdasarkan hasil penelitian saya dengan arahan dosen pembimbing. Sumber informasi atau kutipan yang berasal atau dikutip dari karya yang diterbitkan telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka di bagian akhir skripsi ini. Skripsi ini belum pernah diajukan untuk memperoleh gelar dalam program sejenis di perguruan tinggi manapun

Semarang, 13 Mei 2015


Mahanani
4401408104



PENGESAHAN

Skripsi yang berjudul :

Analisis Soal *International Competitions and Assessments for Schools* (ICAS) dengan Menggunakan Metode *Item Response Theory* (IRT) dan *Classical Test Theory* (CTT)

disusun oleh

Mahanani

4401408104

telah dipertahankan di hadapan sidang Panitia Ujian Skripsi FMIPA UNNES pada tanggal 11 Mei 2015.

PanitiaUjian:

Ketua,



Prof. Dr. Wiyanto, M.Si

NIP. 196310121988031001

Sekretaris,



Andin Irsadi, S.Pd, M.Si

NIP. 197403102000031001

Ketua Penguji



Sri Sukaesih, S.Pd, M.Pd

NIP. 197908292005012002

Anggota Penguji/

Pembimbing I



Dr. Saiful Ridlo, M.Si

NIP. 196604191991021002

Anggota Penguji/

Pembimbing II



Dra. Ely Rudyatmi, M.Si

NIP. 196205241987102001

MOTTO DAN PERSEMBAHAN

Allah tidak membebani seseorang melainkan sesuai dengan kesanggupannya.

(QS. Al Baqarah: 286)

Sesungguhnya sesudah kesulitan itu ada kemudahan.

(QS. Al Insyirah: 6)

PERSEMBAHAN

Untuk Bapak, Emak, Dek Irkham, dan Mas Maburr.

PRAKATA

Kami panjatkan puji syukur atas kehadiran Allah SWT yang telah melimpahkan rahmatNya sehingga penulis dapat menyelesaikan skripsi dengan judul Analisis Soal *International Competitions and Assessments for Schools* (ICAS) dengan Menggunakan Metode *Item Response Theory* (IRT) dan *Classical Test Theory* (CTT).

Penulis menyadari bahwa penyusunan skripsi ini tidak terlepas dari bantuan dan dukungan berbagai pihak, maka penulis mengucapkan terima kasih kepada :

1. Rektor Universitas Negeri Semarang atas kesempatan yang diberikan kepada penulis sehingga dapat menyelesaikan studinya.
2. Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam atas izin yang diberikan kepada penulis untuk melakukan penelitian.
3. Ketua Jurusan Biologi atas kemudahan administrasi dalam menyelesaikan skripsi ini.
4. Dr. Saiful Ridlo, M.Si sebagai dosen pembimbing I yang telah memberikan pengarahan dan bimbingan dengan penuh kesabaran.
5. Dra. Ely Rudyatmi, M.Si. sebagai dosen pembimbing II yang telah memberikan pengarahan dan bimbingan dengan penuh kesabaran.
6. Sri Sukaesih, S.Pd, M.Pd sebagai dosen penguji yang telah memberikan saran dan masukan yang sangat berguna untuk penyempurnaan skripsi ini.
7. Kepala SMP Negeri 2 Semarang, kepala SMP Negeri 5 Semarang, kepala SMP Negeri 1 Kudus, dan kepala SMP Negeri 5 Yogyakarta yang telah memberikan izin untuk melaksanakan penelitian.
8. Bapak, Ibu, dek Irkham, mas Maburur dan semua pihak yang telah memberi do'a, bantuan, dukungan serta semangat.

Akhir kata, semoga skripsi ini dapat memberi tambahan ilmu bagi para pembaca untuk meningkatkan wawasan pengetahuan.

Semarang, Mei 2015

Penulis

ABSTRAK

Mahanani. 2015. *Analisis Soal International Competitions and Assessments for Schools (ICAS) dengan Menggunakan Metode Item Response Theory (IRT) dan Classical Test Theory (CTT)*. Skripsi, Jurusan Biologi Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Negeri Semarang. Pembimbing Utama Dr. Saiful Ridlo, M.Si dan Pembimbing Pendamping Dra. Ely Rudyatmi, M.Si.

Kata kunci: analisis soal, ICAS, IRT, CTT

Penelitian ini bertujuan untuk menganalisis keakuratan dan kualitas soal ICAS IPA tahun 2010 menurut IRT dan CTT. Jenis penelitian ini adalah deskriptif kuantitatif. Data penelitian ini adalah lembar soal, kunci jawaban dan respons peserta tes yang terdiri atas seluruh sekolah, kelompok 1 (SMP N 5 Yogyakarta) dan kelompok 2 (SMP N 2 Semarang, SMP N 5 Semarang, SMP N 1 Kudus). Analisis data menggunakan metode CTT dan IRT model logistik 1 parameter dengan bantuan program TAP versi 14.7.4 dan BILOG MG 3. Hasil penelitian menurut CTT diketahui bahwa semakin kecil jumlah peserta tes dalam suatu kelompok maka semakin besar jumlah soal yang berkualitas tidak baik. Hasil penelitian menurut IRT model logistik 1 parameter diketahui bahwa semakin kecil jumlah peserta tes dalam suatu kelompok maka semakin kecil jumlah soal yang tidak baik. Koefisien SEM pada IRT lebih kecil daripada CTT. Analisis soal ICAS IPA tahun 2010 menggunakan pendekatan IRT lebih akurat daripada CTT.

Jumlah soal ICAS IPA yang baik menurut IRT lebih banyak daripada CTT. Analisis soal ICAS IPA tahun 2010 menggunakan pendekatan IRT lebih akurat daripada CTT.

DAFTAR ISI

	Halaman
PERNYATAAN KEASLIAN SKRIPSI.....	ii
PENGESAHAN.....	iii
MOTTO DAN PERSEMBAHAN.....	iv
PRAKATA.....	v
ABSTRAK.....	vi
DAFTAR ISI.....	vii
DAFTAR TABEL.....	ix
DAFTAR GAMBAR.....	x
DAFTAR LAMPIRAN.....	xi
BAB 1 PENDAHULUAN.....	1
1.1. Latar Belakang.....	1
1.2. Rumusan Masalah.....	3
1.3. Penegasan Istilah.....	3
1.3.1 Analisis soal.....	3
1.3.2 <i>International competitions and assessments for schools (ICAS)</i>	3
1.3.3 <i>Classical test theory (CTT)</i>	4
1.3.4 <i>Item response theory (IRT)</i>	4
1.4. Tujuan Penelitian.....	5
1.5. Manfaat Penelitian.....	5
1.5.1 Bagi guru.....	5
1.5.2 Bagi peneliti.....	5
1.5.3 Bagi Sekolah.....	5
1.5.4 Bagi Dinas Pendidikan.....	5
BAB 2 TINJAUAN PUSTAKA.....	6
2.1 Analisis soal.....	6
2.2 <i>International competitions and assessments for schools (ICAS)</i>	6
2.3 <i>Classical Test Theory (CTT)</i>	7
2.4 <i>Item response theory (IRT)</i>	11
2.5 Klasifikasi kualitas soal yang baik.....	16
2.6 BILOG MG 3.....	18
2.7 <i>Test analysis program (TAP)</i> versi 14.7.4.....	23
2.8 Kerangka Berpikir.....	25

BAB 3 METODE PENELITIAN	26
3.1 Lokasi dan Waktu Penelitian.....	26
3.2 Obyek Penelitian.....	26
3.3 Rancangan Penelitian.....	26
3.4 Prosedur Penelitian	26
3.5 Data dan Metode Pengumpulan Data.....	27
3.6 Metode Analisis Data.....	27
3.7 Analisis soal.....	29
3.8 Kriteria kualitas soal	32
BAB 4 HASIL PENELITIAN DAN PEMBAHASAN	34
4.1 Hasil Penelitian	34
4.2 Pembahasan	36
BAB 5 PENUTUP	42
5.1 Simpulan.....	42
5.2 Saran.....	42
DAFTAR PUSTAKA	43
LAMPIRAN	46

DAFTAR TABEL

Tabel		Halaman
2.1	Klasifikasi tingkat kesukaran soal menurut CTT.....	9
2.2	Klasifikasi koefisien daya pembeda.....	9
2.3	Klasifikasi koefisien reliabilitas.....	10
2.4	Klasifikasi kemampuan peserta tes.....	15
2.5	Klasifikasi tingkat kesukaran soal menurut IRT.....	15
2.6	Klasifikasi kualitas soal menurut teori tes klasik.....	17
3.1	Klasifikasi tingkat kesukaran soal menurut CTT.....	28
3.2	Klasifikasi tingkat kesukaran soal menurut IRT.....	28
3.3	Klasifikasi daya pembeda soal.....	28
3.4	Klasifikasi reliabilitas.....	29
3.5	Klasifikasi kemampuan peserta tes.....	29
3.6	Tingkat kesukaran soal ICAS IPA tahun 2010 menurut CTT.	30
3.7	Daya pembeda soal ICAS IPA tahun 2010.....	30
3.8	Efektivitas pengecoh soal ICAS IPA tahun 2010.....	30
3.9	Reliabilitas soal ICAS IPA tahun 2010.....	30
3.10	SEM soal ICAS IPA tahun 2010 menurut CTT.....	31
3.11	Fungsi informasi tes ICAS IPA tahun 2010.....	31
3.12	SEM soal ICAS IPA tahun 2010 menurut IRT.....	31
3.13	Kecocokan soal ICAS IPA tahun 2010 dengan model logistik 1 parameter.....	32
3.14	Tingkat kesukaran soal ICAS IPA tahun 2010 menurut IRT..	31
3.15	Klasifikasi kualitas soal ICAS IPA tahun 2010 menurut IRT	32
3.16	Klasifikasi kualitas soal ICAS IPA tahun 2010 menurut CTT	33
3.17	Klasifikasi keakuratan soal ICAS IPA tahun 2010 menurut CTT dan IRT.....	33
4.1	Kualitas soal ICAS IPA tahun 2010 menurut CTT.....	34
4.2	Kualitas soal ICAS IPA tahun 2010 menurut IRT.....	35
4.3	Keakuratan soal ICAS IPA tahun 2010 menurut CTT dan IRT.....	35
4.4	Pola jawaban peserta tes pada soal ICAS IPA tahun 2010.....	37

DAFTAR GAMBAR

Gambar		Halaman
2.1	Karakteristik butir	13
2.2	Kerangka berpikir penelitian	25

DAFTAR LAMPIRAN

Lampiran		Halaman
1	Perangkat Tes ICAS IPA Tahun 2010.....	47
2	Hasil keluaran (<i>output</i>) Program TAP Versi 14.7.4	71
3	Hasil keluaran (<i>output</i>) BILOG MG Fase II dan Fase III	72
4	Kurva Fungsi Informasi Tes dan Kesalahan Baku Pengukuran.....	73
5	Kualitas Soal ICAS IPA Tahun 2010 Menurut CTT Secara Keseluruhan.....	74
6	Kualitas Soal ICAS IPA Tahun 2010 Menurut CTT pada Kelompok 1.....	75
7	Kualitas Soal ICAS IPA Tahun 2010 Menurut CTT pada Kelompok 2.....	76
8	Kualitas Soal ICAS IPA Tahun 2010 Menurut IRT Secara Keseluruhan.....	77
9	Kualitas Soal ICAS IPA Tahun 2010 Menurut IRT pada Kelompok 1.....	78
10	Kualitas Soal ICAS IPA Tahun 2010 Menurut IRT pada Kelompok 2.....	79
11	Invariansi Parameter.....	80
12	Peluang Menjawab Benar Soal ICAS IPA tahun 2010 Secara Keseluruhan.....	82
12	Peluang Menjawab Benar Soal ICAS IPA tahun 2010 pada Kelompok 1.....	82
12	Peluang Menjawab Benar Soal ICAS IPA tahun 2010 pada Kelompok 2.....	83
13	Kurva Karakteristik Butir Soal ICAS IPA tahun 2010 Secara Keseluruhan.....	84
14	Surat Ijin Penelitian Dinas Pendidikan Kota Semarang.....	85
15	Surat Keterangan Telah Melakukan Penelitian SMP N 2 Semarang.....	86
16	Surat Keterangan Telah Melakukan Penelitian SMP N 5 Semarang.....	87
17	Surat Permohonan Ijin Observasi SMP N 1 Kudus	88
18	Surat Keterangan Telah Melakukan Penelitian SMP N 1 Kudus.....	89
19	Surat Permohonan Ijin Observasi Dinas Pendidikan Kota Yogyakarta.....	90
20	Surat Permohonan Ijin Observasi SMP N 5 Yogyakarta	91

BAB 1

PENDAHULUAN

1.1. Latar Belakang

International Competitions and Assessments for Schools (ICAS) merupakan suatu kompetisi internasional yang diselenggarakan oleh *Educational Assessment Australia (EAA)*. ICAS mencakup mata pelajaran komputer, Bahasa Inggris, matematika, IPA, cara menega dan menulis pada siswa SMP dan SMA (EAA 2013). Antusias siswa SMP di Indonesia begitu besar dalam mengikuti ICAS IPA tahun 2010. Hal ini diketahui dari hasil wawancara dengan salah satu admin EAA, Andreana Johnson, melalui surat elektronik pada tanggal 8 November 2012. 37% peserta ICAS IPA tahun 2010 berasal dari negara Indonesia. 7.340 siswa dari 180 SMP di Indonesia mengikuti ICAS IPA.

Berdasarkan hasil wawancara terhadap guru SMP N 2 Semarang, SMP N 5 Semarang, SMP N 1 Kudus, dan SMP N 5 Yogyakarta, motivasi sekolah mengikuti kompetisi ICAS IPA adalah untuk mengetahui kemampuan siswa pada mata pelajaran IPA dalam lingkup internasional. Selain menggunakan hasil *placement test* yang dilakukan secara mandiri, sekolah juga menggunakan hasil kompetisi ICAS sebagai bahan pertimbangan penempatan kelas pada siswa kelas 9. Salah satu motivasi siswa dalam mengikuti kompetisi ICAS adalah untuk mendapatkan piagam penghargaan sebagai alat pendukung untuk melanjutkan ke sekolah favorit.

ICAS IPA merupakan salah satu kegiatan mengungkap dan menaksir kemampuan peserta tes pada mata pelajaran IPA dalam lingkup internasional. Sebagai estimator kemampuan peserta tes, ICAS memerlukan perangkat tes berupa paket soal ICAS. Dapat diasumsikan bahwa estimasi kemampuan peserta tes ditentukan oleh kualitas soal ICAS yang diujikan. Soal ICAS yang berkualitas baik akan memberikan estimasi yang tepat dan akurat terhadap kemampuan peserta tes.

Upaya mengetahui kualitas soal ICAS dapat dilakukan dengan menganalisis soal ICAS. Analisis soal ICAS dapat dilakukan dengan

menggunakan dua pendekatan yaitu teori respons butir atau *item response theory* (IRT) dan teori tes klasik atau *classical test theory* (CTT). Kualitas soal ICAS ditentukan oleh parameter dan keakuratan soal ICAS.

Parameter soal yang digunakan pada analisis soal menurut IRT dan CTT pada dasarnya sama yaitu tingkat kesukaran soal, daya pembeda soal, tebakan semu, dan kemampuan. Perbedaan antara IRT dan CTT terletak pada rumus dan skala yang digunakan.

CTT merupakan teori pengukuran tes secara klasik. Keunggulan CTT adalah konsep pengukurannya mudah dipahami, analisisnya sederhana dan memerlukan ukuran sampel kecil. CTT juga memiliki berbagai kelemahan, tiga diantaranya yaitu *group dependent*, *item dependent*, dan estimasi SEM berlaku untuk semua peserta tes.

IRT dikembangkan untuk mengatasi kelemahan-kelemahan CTT. IRT memiliki keunggulan yaitu *not group dependent and not item dependent*. Sehingga kesulitan untuk memenuhi konsep tes paralel pada CTT, dapat diatasi dengan metode statistik IRT yang memungkinkan mengestimasi SEM pada peserta tes dengan tingkat kemampuan yang berbeda (Hambleton *et al.* 1991). IRT memiliki kelemahan yaitu konsep pengukurannya relatif sulit dipahami karena memerlukan pemahaman matematika yang lebih rumit dan memerlukan ukuran sampel besar. IRT memiliki 3 model pengukuran, yaitu model logistik satu parameter yang analisisnya menitikberatkan pada tingkat kesukaran soal, model logistik dua parameter yang analisisnya menitikberatkan pada tingkat kesukaran soal dan daya pembeda soal, serta model logistik tiga parameter yang analisisnya menitikberatkan pada tingkat kesukaran soal, daya pembeda soal, dan tebakan semu. IRT pada penelitian ini menggunakan model logistik satu parameter karena ingin menitikberatkan kegiatan analisis soal pada parameter tingkat kesukaran soal.

Keakuratan soal ICAS dapat diketahui dengan menghitung besarnya nilai kesalahan baku pengukuran atau *standar error measurement* (SEM) menurut IRT dan CTT. Nilai SEM pada IRT bergantung pada kemampuan peserta tes dan fungsi informasi tes. Nilai SEM pada CTT bergantung pada indeks reliabilitas.

Semakin kecil nilai SEM menunjukkan bahwa soal ICAS semakin akurat (Hidayati 2010).

Sebagaimana telah diuraikan di atas, bahwa upaya mengetahui kualitas soal ICAS dapat dilakukan dengan menganalisis parameter dan keakuratan soal ICAS menurut CTT dan IRT. Parameter soal dianalisis berdasarkan masing-masing pendekatan. IRT pada penelitian ini menitikberatkan pada tingkat kesukaran soal sehingga menggunakan IRT model logistik satu parameter. Keakuratan soal ICAS dapat diketahui dari besarnya nilai SEM. Laporan hasil analisis soal ICAS IPA tahun 2010 yang diserahkan penyelenggara kepada sekolah maupun kepada peserta tes belum memuat kualitas soal ICAS secara lengkap. Oleh karena itu, kegiatan menganalisis soal ICAS IPA menggunakan metode IRT model logistik satu parameter dan CTT perlu dilakukan untuk mengetahui kualitas soal ICAS menurut IRT dan CTT.

1.2. Rumusan Masalah

Berdasarkan latar belakang di atas, maka rumusan masalah penelitian ini adalah:

- a. Bagaimana kualitas soal ICAS IPA tahun 2010 menurut IRT dan CTT?
- b. Bagaimana keakuratan soal ICAS IPA tahun 2010 menurut IRT dan CTT?

1.3. Penegasan Istilah

Dalam penelitian ini digunakan beberapa istilah sebagai berikut:

1.3.1 Analisis soal

Analisis soal merupakan prosedur sistematis yang memberikan informasi khusus terhadap tes yang disusun (Arikunto 2006). Analisis soal dalam penelitian ini merupakan prosedur sistematis untuk mengkaji soal ICAS agar diperoleh informasi khusus tentang soal ICAS, sehingga dapat menentukan kesimpulan yang didukung oleh data yang ada.

1.3.2 *International competitions and assessments for schools (ICAS)*

International Competitions and Assessments for Schools (ICAS) merupakan suatu kompetisi internasional yang diselenggarakan oleh *Educational Assessment Australia (EAA)*. ICAS mencakup mata pelajaran komputer, Bahasa

Inggris, matematika, IPA, cara mengeja dan menulis pada siswa SMP dan SMA. ICAS IPA tahun 2010 memiliki 10 tipe perangkat tes, yaitu A, B, C, D, E, F, G, H, I, dan J. Setiap tipe perangkat tes memuat 45 soal yang berisi tentang mata pelajaran biologi, fisika, dan kimia (EAA 2013). ICAS dalam penelitian ini adalah kompetisi ICAS IPA yang diselenggarakan pada tahun 2010 di SMP N 2 Semarang, di SMP N 5 Semarang, di SMP N 1 Kudus, dan di SMP N 5 Yogyakarta.

1.3.3 Classical test theory (CTT)

Classical test theory (CTT) atau teori tes klasik adalah teori tentang skor tes yang memperkenalkan tiga konsep yaitu skor amatan, skor murni, dan skor kesalahan (Hambleton *et al.* 1991). *Test analysis program* (TAP) versi 14.7.4 adalah program windows yang digunakan untuk menganalisis soal berdasarkan teori tes klasik. Hasil analisis (TAP) versi 14.7.4 adalah skor total peserta tes, tingkat kesukaran soal, daya pembeda soal, korelasi biserial soal, korelasi pengecoh, kesalahan baku pengukuran (SEM), reliabilitas, dan lain-lain (Brooks & Johanson 2003).

CTT dalam penelitian ini adalah suatu pengkajian tentang teori skor tes yang terdiri atas tiga konsep yaitu skor amatan, skor murni, dan skor kesalahan. Analisis soal menurut CTT dilakukan dengan bantuan program TAP versi 14.7.4.

1.3.4 Item response theory (IRT)

Item response theory (IRT) atau teori respons butir merupakan suatu kajian soal yang didasarkan pada parameter soal dan kemampuan peserta tes (Rudner 2001). IRT memiliki 3 model pengukuran, yaitu: model logistik satu parameter, model logistik dua parameter, dan model logistik tiga parameter (Sudaryono 2013). Analisis soal menurut IRT dilakukan dengan bantuan program BILOG MG 3 pada fase II dan III. Hasil analisis fase II adalah parameter soal menurut IRT dan peluang menjawab benar. Hasil analisis fase III adalah parameter kemampuan peserta tes, informasi tes, dan SEM menurut IRT (Zimowski *et al* 2003).

IRT dalam penelitian ini adalah suatu kajian soal yang didasarkan pada parameter soal dan kemampuan peserta tes dengan menggunakan model logistik satu parameter yang analisisnya menitikberatkan pada tingkat kesukaran soal.

Analisis soal menurut IRT dilakukan dengan bantuan program BILOG MG 3 pada fase II dan III.

1.4. Tujuan Penelitian

Berdasarkan rumusan masalah di atas, maka tujuan penelitian ini adalah:

- a. Menganalisis kualitas soal ICAS IPA tahun 2010 menurut IRT dan CTT.
- b. Menganalisis keakuratan soal ICAS IPA tahun 2010 menurut IRT dan CTT.

1.5. Manfaat Penelitian

Hasil penelitian ini diharapkan dapat digunakan sebagai bahan pertimbangan bagi guru, bagi peneliti, bagi sekolah, dan bagi dinas pendidikan. Antara lain:

1.5.1 Bagi guru

Memberi informasi kepada guru mengenai kemampuan peserta tes ICAS IPA tahun 2010 menurut IRT dan CTT secara akurat.

1.5.2 Bagi peneliti

Menambah wawasan mengenai kualitas dan keakuratan soal ICAS IPA tahun 2010 menurut IRT dengan CTT.

1.5.3 Bagi Sekolah

Sebagai bahan pertimbangan untuk memilih kompetisi internasional yang akan diikuti oleh sekolah.

1.5.4 Bagi Dinas Pendidikan

Menjadi masukan dalam hal evaluasi belajar bagi dinas pendidikan sebagai bahan untuk menentukan kebijakan dan langkah-langkah yang dipandang efektif di bidang pendidikan.

BAB 2

TINJAUAN PUSTAKA

2.1 Analisis soal

Analisis soal merupakan proses pengumpulan, peringkasan, dan penggunaan informasi dari jawaban siswa untuk membuat keputusan tentang setiap penilaian (Nitko 2001). Analisis soal merupakan proses untuk mengetahui informasi yang telah dikumpulkan, termasuk mengolah data untuk menentukan kesimpulan yang didukung data (Tayipnapi 2000).

Kustriyono (2004) menyebutkan bahwa ada dua cara yang dapat digunakan untuk mengetahui kualitas tes yaitu analisis kualitatif dan kuantitatif. Surapranata (2006) menjelaskan bahwa analisis kualitatif disebut validitas logis (*logical validity*) dan analisis kuantitatif disebut validitas empiris (*empirical validity*). Menurut Kustriyono (2004) pendekatan kualitatif dilakukan dengan cara menelaah soal sebelum perangkat tes diujikan. Hal yang ditekankan adalah penilaian dari aspek materi, konstruksi, dan bahasa. Sedangkan pendekatan kuantitatif merupakan metode penelaahan soal yang didasarkan pada data empiris yang diperoleh melalui respons peserta tes. Pendekatan kuantitatif dilakukan setelah tes diberikan kepada peserta tes. Melalui validitas empiris akan diperoleh informasi butir soal yang meliputi tingkat kesukaran, daya pembeda dan efektivitas pengecoh.

2.2 *International competitions and assessments for schools (ICAS)*

International competitions and assessments for schools (ICAS) merupakan suatu kompetisi internasional sekaligus suatu kegiatan evaluasi belajar yang independen untuk sekolah-sekolah di dunia dengan peserta lebih dari 1 juta setiap tahunnya berasal dari 6300 sekolah di Australia, Selandia Baru dan berasal dari 20 negara lain termasuk Hong Kong, India, Malaysia, Singapura, Afrika Selatan and Amerika. ICAS diselenggarakan oleh *Educational Assessment Australia (EAA)*. ICAS mencakup mata pelajaran komputer, Bahasa Inggris, matematika, IPA, cara mengeja dan menulis pada siswa SMP dan SMA (EAA 2013).

ICAS IPA tahun 2010 adalah kompetisi ICAS yang diselenggarakan pada tahun 2010. ICAS IPA tahun 2010 memiliki 10 tipe perangkat tes, yaitu A, B, C, D, E, F, G, H, I, dan J. Setiap tipe perangkat tes memuat 45 soal yang berisi tentang mata pelajaran biologi, fisika, dan kimia. Pokok soal ICAS IPA tahun 2010 berupa pernyataan diikuti dengan pertanyaan. Soal ICAS IPA tahun 2010 berupa pilihan ganda yang memiliki 4 pilihan jawaban (1 jawaban benar dan 3 jawaban pengecoh). Ranah kognitif yang termuat dalam pertanyaan adalah menginterpretasi data (mengamati, mengukur, dan menginterpretasikan data, diagram, dan grafik), mengaplikasikan data (memprediksi, menginferensikan, dan menyimpulkan) dan kecakapan tinggi (mengevaluasi data, menginvestigasi, dan menyelesaikan masalah) (EAA 2011).

2.3 Classical Test Theory (CTT)

Menurut Azwar (2007) *classical test theory* atau teori tes klasik terdiri atas skor amatan (X), skor murni (T) dan skor kesalahan pengukuran (E). Jika dituliskan dengan pernyataan matematis, maka kalimat tersebut menjadi

$$X = T + E \dots\dots\dots (1)$$

Keterangan:

X : skor amatan

T : skor murni

E : skor kesalahan pengukuran

Kesalahan pengukuran yang dimaksud dalam teori ini merupakan kesalahan yang tidak sistematis atau acak. Kesalahan ini merupakan penyimpangan secara teoritis dari skor amatan yang diperoleh dengan skor amatan yang diharapkan. Kesalahan pengukuran yang sistematis dianggap bukan merupakan kesalahan pengukuran (Retnawati 2003).

Allen & Yen (2002) menyatakan bahwa tujuh asumsi-asumsi CTT sebagai berikut. Pertama, skor tampak (X) merupakan hasil penjumlahan dari skor murni (T) dan skor kesalahan pengukuran (E). Kedua, skor murni merupakan nilai rata-rata skor perolehan teoretis dari pengujian independen yang dilakukan berulang-ulang (sampai tak terhingga) terhadap peserta tes dengan menggunakan alat ukur. Ketiga, tidak terdapat korelasi antara skor murni dan skor kesalahan pengukuran ($\rho_{eT} = 0$). Keempat, tidak terdapat korelasi antara kesalahan pada pengukuran pertama dan kesalahan pada pengukuran kedua ($\rho_{e1e2} = 0$). Kelima, skor

kesalahan pada tes pertama tidak berkorelasi dengan skor murni pada tes kedua ($\rho_{e_1t_2} = 0$). Keenam, dua tes dikatakan paralel apabila skor murni sama ($T_1=T_2$) dan varian skor-skor kesalahannya sama ($\sigma_e^1 = \sigma_e^2$). Ketujuh, dua tes dikatakan setara (*equivalent test*) apabila memiliki skor tampak X_1 dan X_2 yang memenuhi asumsi 1 sampai 5, dan setiap populasi peserta tes $T_1=T_2+C$, dengan C sebagai suatu bilangan konstan.

Menurut Retnawati (2003) asumsi-asumsi CTT dijadikan dasar untuk mengembangkan formula-formula dalam menentukan validitas dan reliabilitas tes. Validitas dan reliabilitas pada perangkat tes digunakan untuk menentukan kualitas tes. Kriteria lain untuk menentukan kualitas tes adalah tingkat kesukaran soal dan daya pembeda soal. Menurut Depdikbud (1997) kualitas soal ditentukan oleh parameter soal yaitu tingkat kesukaran soal, daya pembeda soal dan efektivitas pengecoh. Penjelasan mengenai validitas, reliabilitas, dan parameter soal menurut CTT disajikan sebagai berikut.

Menurut Sudjana (2001) validitas adalah ketepatan alat penilaian terhadap konsep yang dinilai sehingga betul-betul menilai apa yang seharusnya dinilai. Menurut Surapranata (2006) validitas terdiri atas validitas logis dan validitas empiris. Validitas logis terdiri atas validitas isi dan validitas konstruksi. Validitas empiris terdiri atas tingkat kesukaran soal, daya pembeda soal, dan efektivitas pengecoh.

Menurut Azwar (2003) parameter soal CTT yang pertama adalah tingkat kesukaran soal. Tingkat kesukaran soal adalah rasio antara peserta tes yang menjawab soal secara benar dengan banyaknya peserta tes yang menjawab soal. Tingkat kesukaran soal berkisar antara 0 sampai 1. Rumus untuk menentukan besarnya tingkat kesukaran secara matematis sebagai berikut.

$$P = \frac{B}{JS} \dots\dots\dots (1)$$

Keterangan:

p : tingkat kesukaran

B : banyaknya peserta tes yang menjawab soal dengan benar

JS : banyaknya peserta tes

Tingkat kesukaran soal disimbolkan dengan huruf p. Menurut Surapranata (2006) klasifikasi tingkat kesukaran soal menurut CTT dapat dilihat dalam Tabel 2.1 berikut ini.

Tabel 2.1 Klasifikasi tingkat kesukaran soal menurut CTT

No	P	Kriteria tingkat kesukaran
1	$p < 0,30$	Sukar
2	$0,30 \leq p \leq 0,70$	Sedang
3	$p > 0,70$	Mudah

*Diadaptasi dari Surapranata (2006)

Menurut Arikunto (2006) tingkat kesukaran soal sebaiknya terletak pada interval 0,30 – 0,70. Pada interval 0,30 – 0,70 akan diperoleh informasi tentang kemampuan peserta tes secara maksimal.

Parameter soal CTT yang kedua adalah daya pembeda soal. Menurut Lababa (2008) daya pembeda soal adalah kemampuan soal untuk membedakan antara peserta tes yang berkemampuan tinggi dan berkemampuan rendah. Menurut Crocker & Algina (2006) daya pembeda soal dapat ditentukan dengan berbagai cara. Salah satu cara untuk menentukan daya pembeda soal adalah menghitung koefisien korelasi biserial soal. Koefisien korelasi biserial ditentukan dengan rumus di bawah ini.

$$\rho_{pbis} = \frac{\mu_+ - \mu_\tau}{\sigma_\tau} \sqrt{\frac{p}{q}} \dots\dots\dots (2)$$

Keterangan:

- ρ_{pbis} : korelasi biserial soal
- μ_+ : rata-rata skor peserta tes yang menjawab benar soal
- μ_τ : rata-rata skor total
- σ_τ : simpangan baku skor total
- p : proporsi banyaknya peserta yang menjawab benar
- q : 1- p

Menurut Azwar (2003) besarnya daya pembeda berkisar antara -1 sampai dengan +1. Menurut Ebel & Frisbie (1991) klasifikasi daya pembeda terdiri atas empat kriteria yang dapat dilihat pada Tabel 2.2 di bawah ini.

Tabel 2.2 Klasifikasi koefisien daya pembeda

Daya pembeda	Kriteria Daya pembeda
$\geq 0,40$	Sangat baik
0,30 - 0,39	Lumayan baik
0,20 - 0,29	Perlu direvisi
$< 0,20$	Jelek dan perlu dibuang

* Diadaptasi dari Ebel & Frisbie (1991)

Analisis pada soal pilihan ganda terdapat alternatif jawaban/option yang merupakan pengecoh (distraktor). Soal yang baik, pengecohnya akan dipilih secara merata oleh peserta tes. Sebaliknya soal yang kurang baik, pengecohnya

akan dipilih secara tidak merata (Arikunto 2006). Pengecoh dianggap baik apabila jumlah peserta tes yang memilih pengecoh itu sama atau mendekati jumlah ideal. Suatu pengecoh dapat dikatakan berfungsi baik jika paling sedikit dipilih oleh 5% peserta tes (Surapranata 2006).

Purwanto (2009) menjelaskan bahwa reliabilitas adalah keterpercayaan, ketetapan dan konsistensi dari alat ukur. Menurut Azwar (2007) tinggi rendahnya reliabilitas ditunjukkan oleh suatu angka yang disebut koefisien reliabilitas. Koefisien reliabilitas berkisar 0 sampai dengan 1. Menurut Arikunto (2006) nilai dari koefisien reliabilitas diberi lambang r yang dapat dicari dengan menggunakan rumus *Alpha Croncbanch*. Rumus *Alpha Croncbanch* sebagai berikut.

$$r_{11} = \left[\frac{n}{n-1} \right] \left[\frac{s^2 + \sum pq}{s^2} \right] \dots\dots\dots (3)$$

Keterangan:

- r₁₁ : reliabilitas instrumen
- p : proporsi siswa yang menjawab benar
- q : proporsi siswa yang menjawab salah (q=1-p)
- Σpq : jumlah perkalian antara p dan q
- N : banyaknya item
- s : standar deviasi dari tes

Koefisien r yang diperoleh, dikonsultasikan dengan r tabel *product moment* dengan taraf signifikan 5%. Jika harga r hitung > r tabel, maka soal tersebut reliabel. Klasifikasi koefisien reliabilitas terdiri atas lima kriteria yang dapat dilihat pada Tabel 2.3 di bawah ini.

Tabel 2.3 Klasifikasi koefisien reliabilitas

Koefisien reliabilitas	Kriteria reliabilitas
0,800 < r ≤ 1,000	Sangat tinggi
0,600 < r ≤ 0,800	Tinggi
0,400 < r ≤ 0,600	Cukup
0,200 < r ≤ 0,400	Rendah
0,000 < r ≤ 0,200	Sangat rendah

* Diadaptasi dari Arikunto (2006)

Menurut Purwanto (2009) kesalahan baku pengukuran adalah ukuran yang mencerminkan tidak akuratnya skor dari tes hasil belajar yang digunakan untuk mengukur. Pada metode pengukuran CTT, semakin besar nilai kesalahan baku pengukuran maka koefisien reliabilitasnya makin kecil dan sebaliknya. Rumus kesalahan baku pengukuran juga dapat dihitung dengan rumus berikut.

$$SEM = SD\sqrt{1 - r_{tt}} \dots\dots\dots (4)$$

Keterangan:

SEM : kesalahan baku pengukuran

SD : standar deviasi

r_{tt} : koefisien reliabilitas

Menurut Hidayati (2010) semakin kecil nilai SEM dan semakin besar indeks reliabilitas maka akan semakin akurat perangkat tes. Pada metode CTT, nilai SEM kecil dan indeks reliabilitas besar dapat dicapai apabila analisis soal menggunakan ukuran sampel kecil.

Menurut Hambleton *et al.* (1985) parameter soal pada CTT berupa tingkat kesukaran soal dan daya pembeda soal bergantung pada karakteristik peserta tes. Hal ini menjadi kelemahan CTT. Jika kemampuan peserta tes tinggi maka indeks kesukaran soal akan tinggi dan sebaliknya (*group dependent*). Peserta tes juga bergantung pada parameter soal. Hasil pengukuran kemampuan peserta tes bergantung dari soal yang diujikan. Jika soal yang diujikan memiliki tingkat kesulitan yang tinggi, estimasi kemampuan peserta akan rendah dan sebaliknya (*item dependent*). Besarnya daya pembeda soal juga bergantung pada homogenitas kelompok peserta tes. Semakin tidak homogen kemampuan peserta tes maka semakin tinggi koefisien daya pembeda soalnya. Menurut Sulistiyono (2011) Besarnya indeks kesukaran soal dan daya pembeda soal akan mempengaruhi reliabilitas soal. Tes yang sangat sukar atau sangat mudah cenderung memiliki daya pembeda soal rendah dan reliabilitas rendah.

2.4 Item response theory (IRT)

Item response theory (IRT) atau teori respons butir adalah penelaahan soal dengan menggunakan jawaban soal. IRT menjelaskan hubungan peluang menjawab benar suatu soal dengan kemampuan peserta tes yang mendasarinya dalam fungsi matematika (Diknas 2008). Konsep teori respons butir didasarkan pada dua postulat, yaitu (a) prestasi peserta tes pada suatu soal dapat diprediksikan oleh sekelompok faktor yang disebut *latent trait* atau kemampuan, dan (b) hubungan antara prestasi peserta tes pada suatu soal dengan kemampuan yang mendasarinya digambarkan oleh suatu fungsi naik secara monoton yang disebut kurva karakteristik butir atau *Item Characteristic Function/ICC* (Retnawati 2003).

Tujuan IRT adalah membentuk parameter soal dan parameter peserta tes yang bersifat invarian sehingga dapat mengatasi kelemahan-kelemahan CTT yaitu

group dependent dan *item dependent*. *Group dependent* artinya hasil pengukuran bergantung pada peserta tes yang mengerjakan tes. Jika tes diujikan kepada kelompok peserta tes dengan kemampuan tinggi, maka tingkat kesukaran soal menjadi rendah, begitu pula sebaliknya. *item dependent* artinya hasil pengukuran bergantung pada soal yang diujikan. Jika soal yang diujikan memiliki tingkat kesukaran tinggi, maka estimasi kemampuan peserta tes rendah, begitu pula sebaliknya (Hambleton & Swaminathan 1985). Tujuan IRT terwujud jika ada kecocokan antara data tes dengan model yang digunakan (Retnawati 2003). Pemilihan model yang cocok dengan data bergantung pada banyaknya data yang diperoleh untuk mengestimasi parameter-parameter yang digunakan. Semakin banyak parameter yang diestimasi maka semakin banyak data respons peserta tes yang dibutuhkan (Stocking 1999). Sampel atau data respons peserta tes yang dibutuhkan paling tidak 100 orang untuk model Rasch. Selain itu, syarat suatu model agar cocok dengan data tes yaitu harus memenuhi asumsi-asumsi dalam IRT (Mardapi 1991). Ukuran sampel untuk model logistik 1 parameter atau model Rasch antara 30 sampai 300 (Bond & Fox 2007).

IRT memiliki tiga asumsi. Asumsi pertama adalah unidimensi. Unidimensi adalah asumsi yang menyatakan bahwa setiap soal hanya mengukur satu kemampuan peserta tes (Sudaryono 2013). Suatu soal dapat dinyatakan unidimensi apabila telah dibuktikan dengan analisis faktor. Apabila hasil pengukuran menunjukkan bahwa dimensi yang dominan adalah satu, maka asumsi unidimensi sudah terpenuhi. Penggunaan analisis faktor bertujuan untuk memperlihatkan pada kelompok faktor mana butir itu berada. Setiap faktor hanya menunjukkan suatu dimensi kemampuan peserta tes (Syarifah 2007).

Asumsi kedua adalah independensi lokal. Independensi lokal dibagi menjadi dua yaitu independensi lokal terhadap respons peserta tes dan independensi lokal terhadap soal. Independensi lokal terhadap respons peserta tes memiliki arti bahwa benar salahnya peserta tes menjawab sebuah soal tidak dipengaruhi oleh benar salahnya peserta tes lain dalam menjawab benar soal. Sedangkan independensi lokal terhadap soal memiliki arti bahwa benar salahnya seorang peserta tes menjawab sebuah soal tidak dipengaruhi oleh benar salahnya peserta tes dalam menjawab benar soal yang lain (Allen & Yen 2002).

Independensi lokal dapat diuji dengan dua cara, yaitu: secara eksak melalui rumus peluang, dan secara statistika melalui *chi square*. Pengujian independensi lokal menggunakan uji *chi square* atau khi-kuadrat dilakukan dengan taraf signifikansi tertentu melalui hipotesis. H_0 menunjukkan ada independensi lokal. Sedangkan H_2 menunjukkan tidak ada independensi lokal. Rumus khi-khuadrat adalah sebagai berikut.

$$x^2 = \left[\frac{\sum (f_o - f_e)^2}{f_e} \right] \dots \dots \dots (5)$$

Keterangan:

x^2 : khi-kuadrat

f_0 : frekuensi amatan

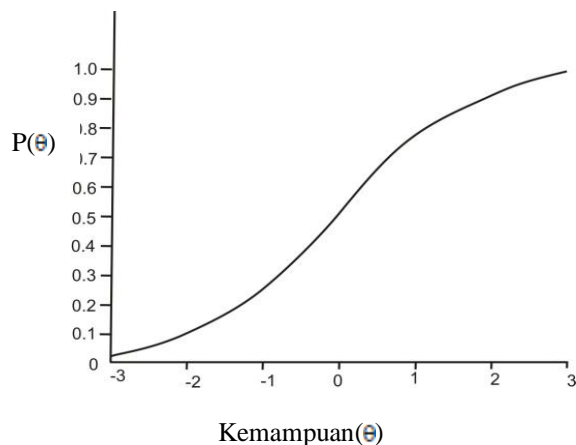
f_e : frekuensi harapan

α : 0.05

V : (baris tabel-1)(kolom tabel-1)

Kriteria pengujian adalah H_0 ditolak jika x^2 hitung $>$ x^2 tabel (α)(v) atau jika Sign. $x^2 < \alpha$, H_0 diterima jika x^2 hitung $<$ x^2 tabel (α)(v) atau Sign. $x^2 > \alpha$.

Asumsi ketiga adalah kurva karakteristik butir. Menurut Baker (2001) IRT memiliki kurva karakteristik butir sebagai fungsi matematika yang menyatakan hubungan antara peluang menjawab benar peserta tes dengan kemampuan peserta tes. Kurva karakteristik butir menjelaskan bahwa peserta tes berkemampuan tinggi memiliki peluang tinggi untuk menjawab soal dengan benar dan sebaliknya. Peluang menjawab benar disimbolkan dengan $P(\Theta)$ dan kemampuan peserta tes disimbolkan dengan Θ . Kurva karakteristik butir digambarkan pada Gambar 1 sebagai berikut.



Gambar 2.1 Karakteristik butir

Menurut Sudaryono (2013) pada IRT, peserta tes memiliki kemampuan (θ) yang biasanya berbeda diantara peserta tes yang lain. Soal memiliki tingkat kesukaran soal (b) yang biasanya berbeda di antara soal. Pada pengukuran dengan metode pengukuran IRT, terjadi pertemuan antara kemampuan peserta tes dengan tingkat kesukaran soal. Respons peserta tes terhadap soal menghasilkan hasil ukur. Hasil ukur dinyatakan dalam bentuk peluang menjawab benar. Peluang menjawab benar ditentukan oleh padanan di antara kemampuan peserta tes dengan tingkat kesukaran soal. Peluang menjawab benar adalah peluang menjawab benar peserta tes ke- g pada soal ke i . soal yang terlalu sukar atau terlalu sulit tidak dapat menunjukkan kemampuan peserta tes. Sehingga akurasi pengukuran menjadi rendah. Kecocokan antara kemampuan dengan tingkat kesukaran soal menghasilkan akurasi pengukuran yang tinggi dengan ketentuan $P(\theta) = P_{\min} + 0,5(P_{\max} - P_{\min})$. Kecocokan antara kemampuan dengan tingkat kesukaran soal dapat dicapai jika keduanya independen. Jika tingkat kesukaran soal independen dari kemampuan maka dapat dicari nilai kesukaran soal yang cocok dengan kemampuan.

Menurut Hambleton *et al.* (1991) IRT memiliki 3 model logistik yaitu model logistik satu parameter, model logistik dua parameter, dan model logistik tiga parameter. Model logistik satu parameter atau model Rasch, merupakan model logistik untuk menganalisis soal yang hanya menitikberatkan pada parameter tingkat kesukaran soal (b_i). Model ini dipublikasikan oleh George Rasch pada tahun 1960an. Model Rasch hanya menggunakan satu parameter soal, yaitu parameter tingkat kesulitan soal, sedang parameter lainnya, seperti daya pembeda soal dianggap sama untuk semua butir, dan tebakan semu sama dengan nol. Rumus model logistik satu parameter sebagai berikut.

$$P_i(\theta) = \frac{e^{(\theta - b_i)}}{1 + e^{(\theta - b_i)}} \dots \dots \dots (6)$$

Keterangan:

$P_i(\theta)$: peluang peserta tes menjawab benar soal ke- i

θ : tingkat kemampuan peserta tes

e : angka transcendent yang nilainya = 2,718

b_i : parameter tingkat kesukaran soal pada soal ke- i

Menurut Hambleton *et al.* (1991) parameter b_i didefinisikan sebagai nilai skala kemampuan peserta tes yang memiliki peluang menjawab benar sebesar 50% pada soal tertentu. Soal yang memiliki parameter b_i sebesar 0,3 berarti diperlukan kemampuan minimal 0,3 pada skala untuk dapat menjawab benar dengan peluang 50%. Semakin besar nilai parameter b_i , maka semakin besar kemampuan yang diperlukan untuk menjawab benar dengan peluang 50%. Semakin besar nilai parameter b_i , maka makin sulit soal. Hubungan peluang menjawab benar $P_i(\theta)$ dengan tingkat kemampuan peserta (θ) dapat digambarkan sebagai kurva karakteristik butir.

Menurut Hambleton & Swaminathan (1985) secara teoritis, nilai kemampuan dan tingkat kesukaran (b_i) terletak di antara $-\infty$ dan $+\infty$. Sesuai dengan daerah asal distribusi normal, nilai kemampuan peserta (θ) terletak di antara -3 dan 3. Tingkat kesukaran soal (b_i) yang baik terletak di antara -2 dan 2. Menurut Hidayatulloh (2013) klasifikasi kemampuan peserta tes disajikan pada Tabel 2.4 dan klasifikasi tingkat kesukaran soal menurut IRT disajikan pada Tabel 2.5 sebagai berikut.

Tabel 2.4 Klasifikasi kemampuan peserta tes

θ	Kriteria kemampuan peserta tes
$\theta > 1$	Kemampuan tinggi
$-1 \leq \theta \leq 1$	Kemampuan sedang
$\theta < -1$	Kemampuan rendah

* Diadaptasi dari Hidayatulloh (2013)

Tabel 2.5 Klasifikasi tingkat kesukaran soal menurut IRT

b	Kriteria tingkat kesukaran
$b > 2$	Sangat sukar
$1 < b \leq 2$	Sukar
$-1 < b \leq 1$	Sedang
$-1 > b \geq -2$	Mudah
$b < -2$	Sangat mudah

* Diadaptasi dari Hidayatulloh (2013)

Menurut Hambleton & Swaminathan (1985) IRT menggunakan istilah informasi butir untuk menyatakan reliabilitas soal dan informasi tes untuk menyatakan reliabilitas tes. Fungsi informasi butir menyatakan kekuatan atau sumbangan soal dalam mengungkap *latent trait* yang diukur dengan tes. Dengan fungsi informasi butir diketahui soal mana yang cocok dengan model sehingga

membantu dalam seleksi soal tes. Secara matematis, fungsi informasi butir memenuhi persamaan sebagai berikut.

$$I_i(\theta) = \frac{[P_i(\theta)]^2}{P_i(\theta)[1-P_i(\theta)]} \dots\dots\dots (7)$$

Keterangan:

$I_i(\theta)$: fungsi informasi butir ke-i

i : 1, 2, 3, ..., n

$P_i(\theta)$: peluang peserta dengan kemampuan θ menjawab benar soal ke-i

$P'_i(\theta)$: turunan fungsi $P_i(\theta)$ terhadap (θ)

$Q_i(\theta)$: peluang peserta dengan kemampuan θ menjawab benar soal ke-i

Fungsi informasi tes merupakan jumlah dari fungsi informasi butir penyusun tes. Fungsi informasi tes akan tinggi jika soal memiliki fungsi informasi yang tinggi pula. Fungsi informasi perangkat tes secara matematis dapat dituliskan sebagai berikut.

$$I(\theta) = \sum_{i=1}^n I_i(\theta) \dots\dots\dots (8)$$

Keterangan:

$I(\theta)$: fungsi informasi tes

Menurut Retnawati (2003) kesalahan baku pengukuran dalam IRT menunjukkan ketidakakuratan hasil analisis dan berkaitan erat dengan fungsi informasi tes. Fungsi informasi tes dengan SEM memiliki hubungan berbanding terbalik kuadratik. Semakin besar fungsi informasi tes maka SEM semakin kecil atau sebaliknya. Hubungan matematis antara fungsi informasi tes dengan SEM sebagai berikut.

$$SEM(\theta) = \frac{1}{\sqrt{I(\theta)}} \dots\dots\dots (9)$$

Keterangan:

$SEM(\theta)$: kesalahan baku pengukuran

2.5 Klasifikasi kualitas soal yang baik

Menurut Hidayatulloh (2013) kualitas soal menurut teori tes klasik ditentukan oleh parameter soal penyusunnya yaitu tingkat kesukaran, daya pembeda, dan efektivitas pengecoh. Kualitas soal menurut teori tes klasik disajikan pada Tabel 2.6 sebagai berikut.

Tabel 2.6 Klasifikasi kualitas soal menurut teori tes klasik

Kriteria	Tingkat kesukaran soal (p)	Korelasi biserial	Korelasi Biserial
Baik	$30\% \leq p \leq 70\%$	$\rho_{bis} \geq 0,3$	Negatif kecuali kunci jawaban
Cukup Baik	$p \geq 70\%$ atau $p \leq 30\%$	$\rho_{bis} \geq 0,3$	Negatif kecuali kunci jawaban
	$30\% \leq p \leq 70\%$	$\rho_{bis} \geq 0,3$	Ada korelasi biserial positif selain kunci jawaban
	$30\% \leq p \leq 70\%$	$0,2 \leq \rho_{bis} < 0,3$	Negatif kecuali kunci jawaban
Tidak baik	$p \geq 70\%$ atau $p \leq 30\%$	$\rho_{bis} < 0,20$	Ada korelasi biserial positif selain kunci jawaban
		$\rho_{bis} < 0,20$	Ada korelasi biserial positif selain kunci jawaban

* Diadaptasi dari Hidayatulloh (2013)

Menurut Hambleton *et al.* (1991) kualitas soal yang baik menurut IRT adalah soal yang memiliki parameter tingkat kesukaran soal $-2 \leq b \leq 2$ dalam skala logit pada kurva normal. Soal yang ditolak/dikriteriakan tidak baik adalah soal yang memiliki parameter tingkat kesukaran soal (b) > 2 atau < -2 dalam skala logit. Menurut Hidayatulloh (2013) kualitas soal yang baik menurut IRT adalah soal yang cocok dengan model IRT dan memiliki tingkat kesukaran soal $-2 \leq b \leq 2$ dalam skala logit. Menurut Hayat (1994) soal yang tidak cocok dengan model, secara empiris merupakan soal yang tidak berfungsi sebagaimana mestinya atau soal yang bias. Soal yang tidak cocok terhadap model disebabkan oleh materi, konstruksi, atau bahasa.

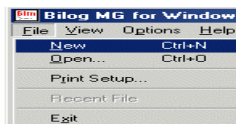
Hidayati (2010) menjelaskan bahwa klasifikasi kualitas soal yang baik menurut teori tes klasik adalah memiliki koefisien reliabilitas besar dan nilai SEM kecil pada ukuran sampel kecil (≤ 500). Sedangkan klasifikasi kualitas soal yang baik menurut teori respons butir adalah memiliki fungsi informasi tes tinggi dan nilai SEM kecil pada ukuran sampel besar (≥ 500). Secara keseluruhan SEM yang paling kecil diperoleh jika melakukan analisis menggunakan teori respons butir. Analisis soal menurut teori tes klasik akan akurat apabila menggunakan ukuran

sampel kecil dan menurut teori respons butir akan akurat apabila menggunakan ukuran sampel besar.

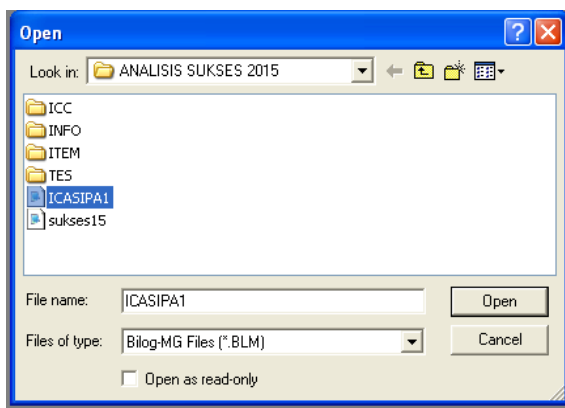
2.6 BILOG MG 3

Menurut Zimowski *et al* (2003) BILOG MG 3 adalah sebuah ekstensi dari program BILOG yang didesain untuk analisis soal berbentuk pilihan ganda menurut teori respons butir. Hasil analisis BILOG MG 3 ada tiga fase. Hasil analisis fase I adalah tingkat kesukaran soal dan daya pembeda soal menurut CTT. Hasil analisis fase II adalah parameter soal menurut IRT dan peluang menjawab benar. Hasil analisis fase III adalah parameter kemampuan peserta tes, informasi tes, SEM menurut IRT, dan reliabilitas. Penggunaan program BILOG MG 3 untuk menganalisis soal dengan metode pengukuran IRT dan CTT terdiri atas 16 langkah. Langkah-langkah penggunaan program BILOG MG 3 disajikan sebagai berikut.

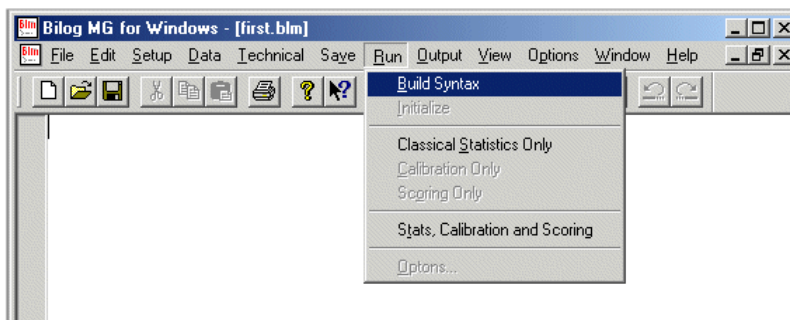
Langkah 1: membuat file baru dengan cara menekan menu *File, New*.



Langkah 2: membuat *file* dalam bentuk *.blm*.



Langkah 3: membuat *syntax* dengan cara menekan menu *Run-Build Syntax*



Langkah 4: menetik *syntax*

```
>GLOBAL DFName = 'F:\examples\examp101.dat',
  NPArm = 3;
>LENGTH NITems = (4);
>INPUT NTOtal = 4,
  NIDchar = 3;
>ITEMS ;
>TEST1 TName = 'TEST0001',
  INumber = (1(1)4);
(3A1, 1X, 4A1)
>CALIB ACCel = 1.0000;
>SCORE ;
```

Langkah 5: menyimpan *syntax* dengan menekan menu *Save*

```
'F:\examples\examp101.dat',
3;
(1);
4,
3;
TEST0001'
```

Langkah 6: membuat *file* dalam bentuk.DAT (jawaban peserta tes).

```
1 DCCDACDACCACDCAAC
2 BCCDDCDACCACDDBB
3 DCCDDCDACCACDDBAD
4 BCCAACDCCDDDDOAD
5 DCCBCDACCACDCAAD
```

Langkah 7: membuat *file* dalam bentuk.KEY (jawaban benar peserta tes).

```
KEY DCCDACDACCACDDBAD
```

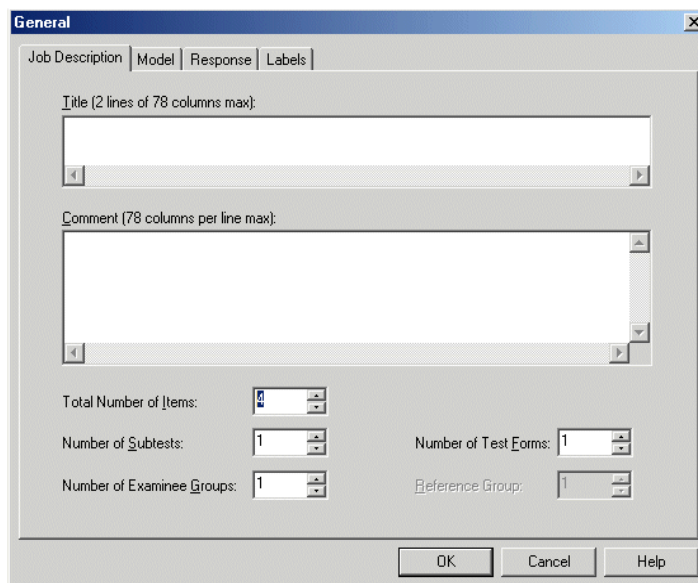
Langkah 8: membuat *file* dalam bentuk.OMT (jawaban salah peserta tes).

```
KEY 0000000000000000
```

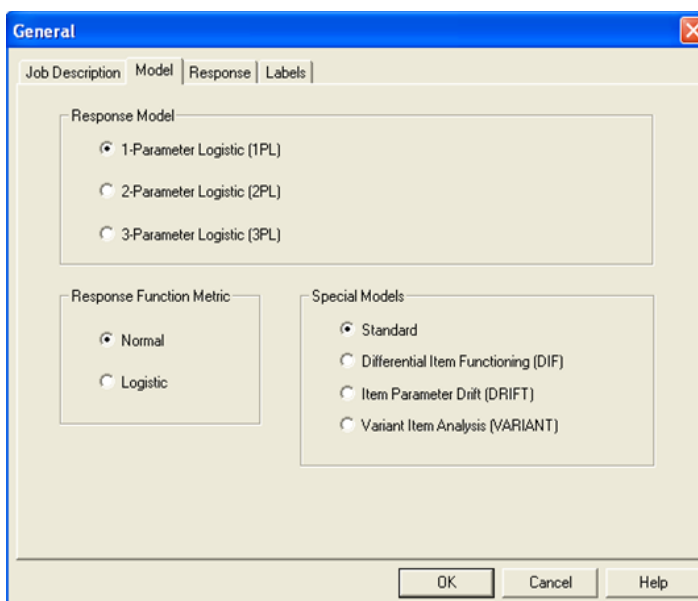
Langkah 9: menekan menu *Setup, General*



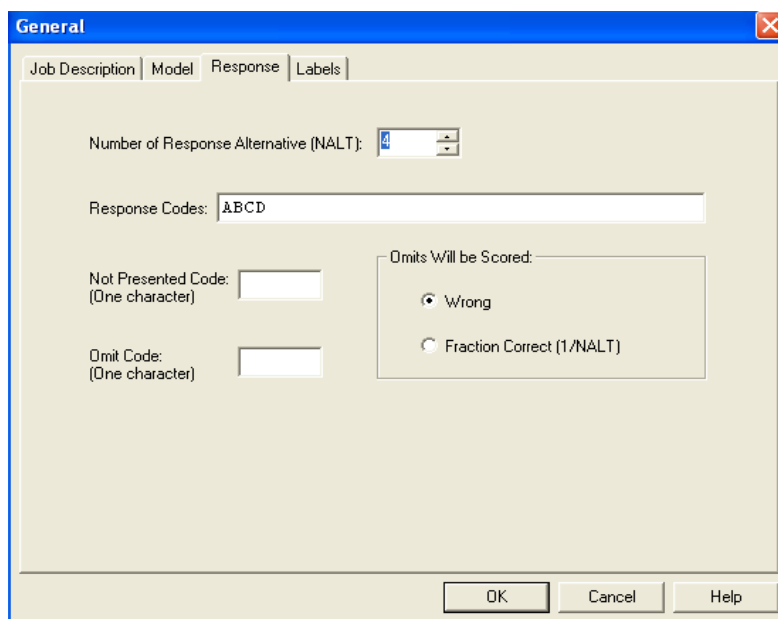
Langkah 10: menekan menu *General, Job Description*, mengisi *Total Number of Items*, *Number of Subtests*, *Number of Examinee Groups*, *Number of Test Form*, *OK*.



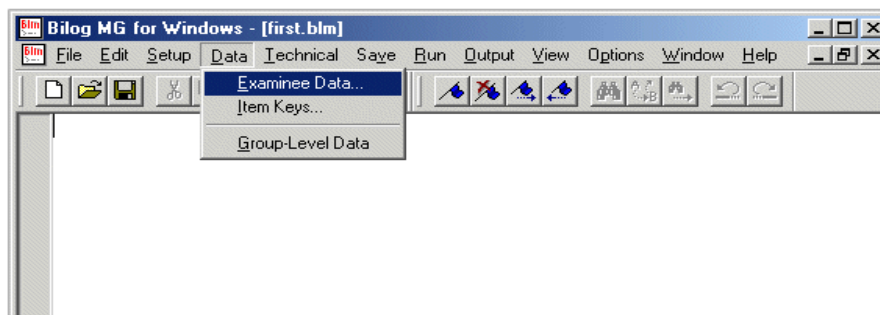
Langkah 11: menekan menu *Model, Response Model (1-Parameter)*, *Response Function Metric (Normal)*, *Special Models (Standar)*, *OK*



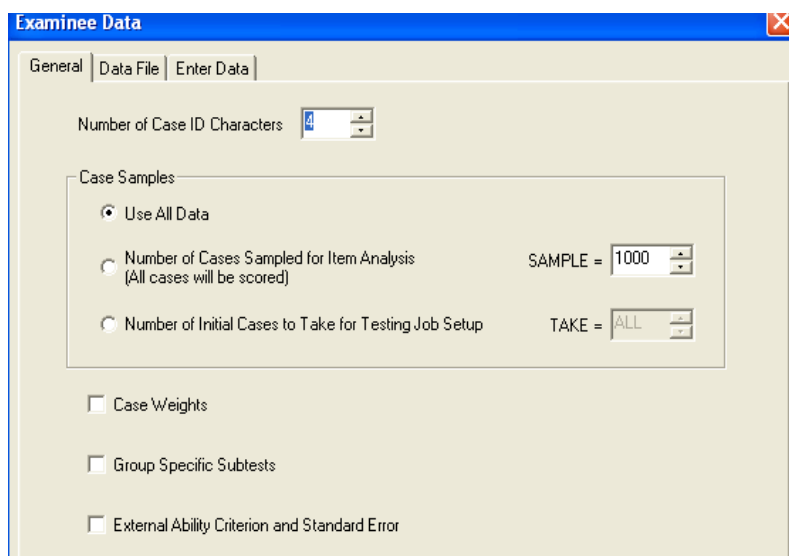
Langkah 12: menekan menu *Response, NALT, Response Codes (ABCD), OK*



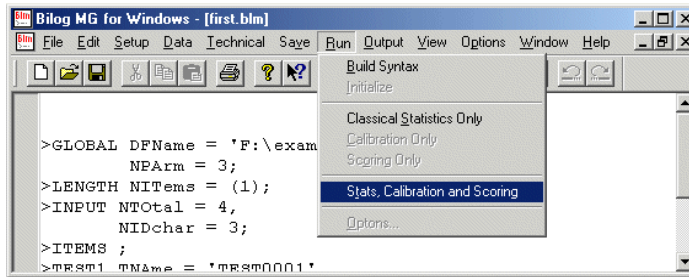
Langkah 13: menekan menu *Data, Examinee Data*



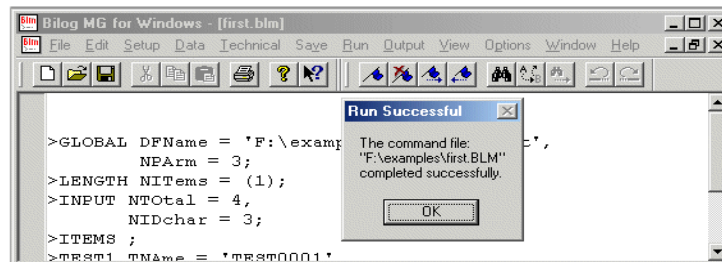
Langkah 14: menekan menu *Examinee Data, Number of Case ID characters, Number of Case Sample of Item Analysis, OK*



Langkah 15: menekan menu *Run, Stats, calibration, and scoring*.

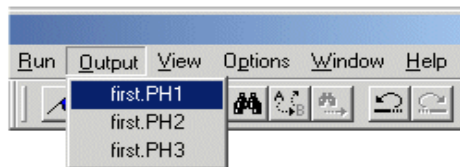


Jika langkah 15 berhasil, maka akan ditemukan tulisan *Run Successful*.



Langkah 16: mendapatkan hasil atau *output* berupa *PHASE1, PHASE2, PHASE 3*.

Dari hasil tersebut kita bisa menganalisis butir soal dengan teori tes klasik dan teori tes modern.



Hasil analisis *PHASE 1*

STATISTICS FOR SUBTEST TEST0001						
NAME	#TRIED	#RIGHT	PCT	LOGIT/1.7	PEARSON	BISERIAL
ITEM0001	660.0	520.0	78.8	-0.77	0.236	0.333
ITEM0002	660.0	411.0	62.3	-0.29	0.319	0.407
ITEM0003	660.0	201.0	30.5	0.49	0.288	0.379
ITEM0004	660.0	351.0	53.2	-0.07	0.322	0.405

Hasil analisis *PHASE 2*

ITEM	INTERCEPT	SLOPE	THRESHOLD	LOADING	ASYMPTOTE
ITEM0001	0.757 0.126*	0.661 0.124*	-1.146 0.265*	0.551 0.103*	0.194 0.087*
ITEM0002	0.145 0.170*	1.032 0.251*	-0.141 0.181*	0.718 0.175*	0.181 0.078*
ITEM0003	-1.261 0.454*	1.233 0.459*	1.022 0.135*	0.777 0.289*	0.114 0.039*
ITEM0004	-0.225 0.210*	1.033 0.275*	0.218 0.170*	0.719 0.191*	0.168 0.069*

Langkah 3: Setelah menekan menu *close & analyze* maka akan muncul hasil sebagai berikut.

TAP: Test Analysis Program (version 14.7.4) N=10K version

File Analysis Options Examples Help

INPUT

- Enter New Data
- Open TAP file
- Insert a TEXT file (see Help)
- Generate Sample Data
- Edit Data in Data Editor

Analyze (F9)

- View Grade Reports (if n<200)
- View Examinee Output (F5)
- View Options Analysis (F7)
- View Item Analysis (F6)
- View Full Test Results (F2)

OUTPUT

When Saving FULL Results, also...

- Save Quick Examinee Results (*.EXM)
- Save Quick Item Analysis (*.ITM)
- Save Quick Options Analysis (*.OPT)
- Save Individual Grade Reports (*.RPT)

Save Results (F3)

Title: ANALISIS SEM ICAS SCIENCE TAHUN 2010

QUICK EXAMINEE RESULTS

```

Number of Examinees = 252
Total Possible Score= 45
Minimum Score      = 10.000 = 22.2%
Maximum Score      = 38.000 = 84.4%
Median Score       = 25.000 = 55.6%
Mean Score         = 24.841 = 55.2%
Standard Deviation = 5.114
Variance           = 26.157
Skewness           = -0.123
Kurtosis           = -0.113

```

Click the "View Examinee Output" button above to VIEW or SAVE or PRINT the detailed Examinee results or hit F5 on your keyboard (similarly for detailed Items and Options output, but FULL TEST does not include Options or Examinee).

QUICK TEST/ITEM RESULTS

```

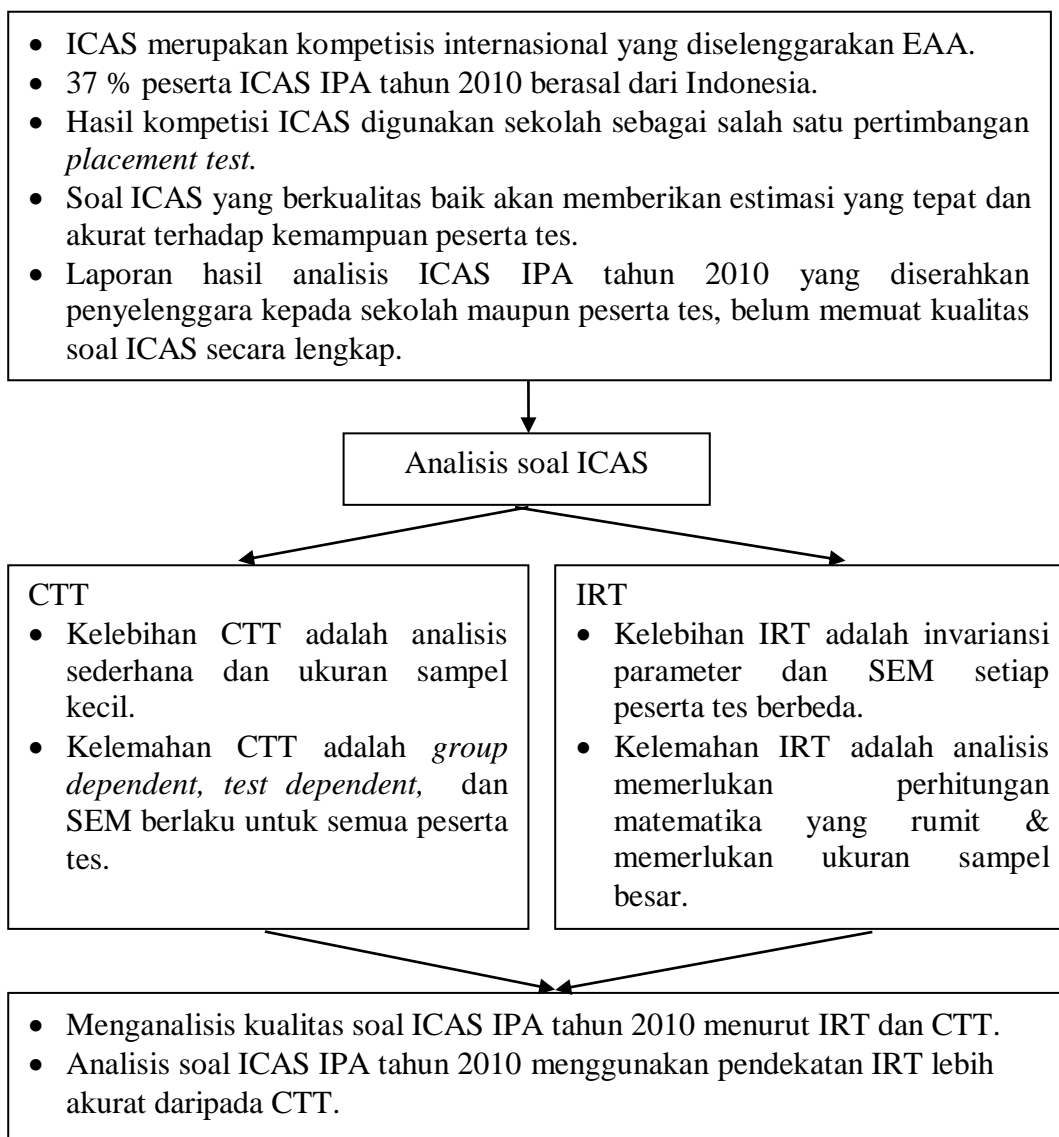
Number of Items Excluded = 0
Number of Items Analyzed = 45
Mean Item Difficulty     = 0.552
Mean Discrimination Index = 0.261
Mean Point Biserial      = 0.256
Mean Adj. Point Biserial = 0.175
KR20 (Alpha)             = 0.683
KR21                     = 0.588
SEM (from KR20)          = 2.881
High Grp Min Score (n=74) = 28.000
Low Grp Max Score (n=81)  = 22.000

```

Potential Problem Items = 8
defined as: difficulty <= 0.20(4)
or: difficulty >= 0.95(1)
or: D index <= 0.00(0)
or: AdjPtBiserial <= 0.00(3)

2.8 Kerangka Berpikir

Berdasarkan uraian di atas, maka kerangka berpikir dalam penelitian ini disajikan pada Gambar 2.2.



Gambar 2.2 Kerangka berpikir penelitian

BAB 3

METODE PENELITIAN

3.1 Lokasi dan Waktu Penelitian

Lokasi penelitian ini adalah 3 SMP di Jawa Tengah dan 1 SMP di Daerah Istimewa Yogyakarta. Keempat SMP adalah SMP N 2 Semarang, SMP N 1 Kudus, SMP N 5 Semarang, dan SMP 5 Yogyakarta. Penelitian ini dilakukan mulai bulan Juli 2012 sampai dengan bulan Desember 2014.

3.2 Obyek Penelitian

Obyek penelitian ini adalah soal, kunci jawaban dan lembar jawab peserta tes ICAS IPA tahun 2010 dari SMP N 2 Semarang (89 lembar jawaban ICAS IPA), SMP N 5 Semarang (19 lembar jawaban ICAS IPA), SMP N 1 Kudus (4 lembar jawaban ICAS IPA), dan SMP N 5 Yogyakarta (140 lembar jawaban ICAS IPA). Dalam penelitian ini tidak dilakukan sampling karena populasinya tersebar dalam daerah yang luas dan sulit dalam hal perijinan.

3.3 Rancangan Penelitian

Rancangan penelitian ini adalah deskriptif kuantitatif. Rancangan penelitian ini mendeskripsikan tingkat kesukaran soal, daya pembeda soal, efektivitas pengecoh, reliabilitas, fungsi informasi tes, kemampuan peserta tes, kecocokan soal dengan model logistik 1 parameter, SEM menurut CTT dan IRT.

3.4 Prosedur Penelitian

Prosedur penelitian terdiri atas persiapan penelitian dan pelaksanaan penelitian. Persiapan penelitian meliputi 3 langkah, yaitu (a) melakukan observasi awal untuk mengetahui data sekolah yang mengikuti olimpiade ICAS IPA pada tahun 2010 di wilayah Jawa Tengah dan Daerah Istimewa Yogyakarta, (b) meminta ijin penelitian di Dinas Pendidikan kota Semarang pada bulan Mei 2012 dan Dinas Pendidikan Provinsi Yogyakarta pada bulan Maret 2012, (c) mengajukan permohonan penelitian di SMP N 2 Semarang, SMP N 5

Semarang, SMP N 1 Kudus, dan SMP N 5 Yogyakarta pada bulan Mei 2012. Pelaksanaan penelitian yaitu mengambil dokumen soal dan jawaban siswa di sekolah yang bersangkutan serta melakukan analisis soal.

3.5 Data dan Metode Pengumpulan Data

Data awal dalam penelitian ini adalah lembar soal, kunci jawaban, dan respons peserta tes yang ada pada lembar jawab perangkat tes ICAS IPA tahun 2010. Pengumpulan data dilakukan dengan menggunakan metode dokumentasi. Metode dokumentasi yang dimaksud adalah dengan meminta lembar soal dan lembar jawab perangkat tes ICAS IPA tahun 2010.

3.6 Metode Analisis Data

Analisis soal ICAS IPA tahun 2010 menurut CTT dan IRT dilakukan meliputi dua langkah yaitu analisis secara keseluruhan dan analisis secara kelompok. Analisis secara keseluruhan adalah menganalisis seluruh respons peserta tes ICAS IPA tahun 2010 menurut CTT dan IRT, sedangkan analisis secara kelompok adalah menganalisis respons peserta tes ICAS IPA tahun 2010 menurut CTT dan IRT pada tiap kelompok. Analisis secara kelompok dilakukan dengan membagi respons peserta tes menjadi 2 kelompok berdasarkan asal sekolah dan jumlah peserta tes sehingga layak dianalisis menurut CTT dan IRT. Sebab analisis soal menurut IRT model Rasch membutuhkan data respons peserta tes paling tidak 100 (Mardapi 1991). Kelompok I terdiri atas 140 data respons peserta tes yang berasal dari SMP N 5 Yogyakarta. Kelompok II terdiri atas 112 data respons peserta tes yang berasal dari SMP N 2 Semarang, SMP N 5 Semarang, dan SMP N 1 Kudus.

Program TAP versi 14.7.4 adalah program windows yang digunakan untuk menganalisis soal berdasarkan teori tes klasik. Hasil analisis TAP versi 14.7.4 adalah tingkat kesukaran soal, korelasi biserial soal, korelasi pengecoh, kesalahan baku pengukuran (SEM), reliabilitas, dan lain-lain. BILOG MG 3 adalah program untuk menganalisis soal berbentuk pilihan ganda menurut teori respons butir pada fase II dan fase III. Hasil analisis BILOG MG 3 fase II adalah parameter soal

menurut IRT dan peluang menjawab benar, sedangkan fase III adalah parameter kemampuan peserta tes, informasi tes, SEM menurut IRT, dan reliabilitas.

Data analisis soal menurut masing-masing pendekatan baik secara keseluruhan maupun secara kelompok, dianalisis secara deskriptif yaitu menginterpretasikan kecenderungan-kecenderungan data dengan teori dan klasifikasi kriteria yang relevan. Teori dan klasifikasi analisis soal meliputi klasifikasi tingkat kesukaran soal, klasifikasi daya pembeda soal, klasifikasi koefisien reliabilitas, teori kesalahan baku pengukuran, teori fungsi informasi tes, dan teori kemampuan. Klasifikasi analisis soal disajikan pada Tabel 3.1, Tabel 3.2, dan Tabel 3.3.

a. Klasifikasi tingkat kesukaran soal menurut CTT disajikan pada Tabel 3.1 dan klasifikasi tingkat kesukaran soal menurut IRT disajikan pada Tabel 3.2 sebagai berikut.

Tabel 3.1 Klasifikasi tingkat kesukaran soal menurut CTT

P	Kriteria
$p < 0,30$	Sukar
$0,30 \leq p \leq 0,70$	Sedang
$p > 0,70$	Mudah

* Diadaptasi dari Surapranata (2006)

Tabel 3.2 Klasifikasi tingkat kesukaran soal menurut IRT

b	Kriteria tingkat kesukaran
$b > 2$	Sangat sukar
$1 < b \leq 2$	Sukar
$-1 < b \leq 1$	Sedang
$-1 > b \geq -2$	Mudah
$b < -2$	Sangat mudah

* Diadaptasi dari Hidayatulloh (2013)

b. Klasifikasi daya pembeda soal disajikan pada Tabel 3.3 sebagai berikut.

Tabel 3.3 Klasifikasi daya pembeda soal

Indeks daya beda	Kriteria Daya Beda
0,40	Sangat baik
0,30 - 0,39	Lumayan baik
0,20 - 0,29	Perlu direvisi
<0, 20	Jelek dan perlu dibuang

* Diadaptasi dari Ebel & Frisbie (1991)

c. Nilai reliabilitas yang diperoleh (r hitung) dikonsultasikan dengan r tabel *product moment* dengan taraf signifikan 5%. Jika harga r hitung $>$ r tabel, maka

soal dinyatakan reliabel. Klasifikasi reliabilitas disajikan pada Tabel 3.4 di bawah ini.

Tabel 3.4 Klasifikasi reliabilitas

Koefisien reliabilitas	Kriteria
$0,800 < r \leq 1,000$	Sangat tinggi
$0,600 < r \leq 0,800$	Tinggi
$0,400 < r \leq 0,600$	Cukup
$0,200 < r \leq 0,400$	Rendah
$0,000 < r \leq 0,200$	Sangat rendah

* Diadaptasi dari Arikunto (2006)

d. Fungsi informasi tes nilainya semakin besar menunjukkan soal-soal ICAS semakin baik.

e. Kesalahan baku pengukuran bernilai semakin kecil menunjukkan bahwa soal ICAS semakin akurat.

f. Kemampuan peserta tes (θ) terletak di antara $-\infty$ dan $+\infty$. Klasifikasi kemampuan peserta tes terdiri atas tiga kriteria yang disajikan pada Tabel 3.5 sebagai berikut

Tabel 3.5 Klasifikasi kemampuan peserta tes

θ	Kriteria kemampuan peserta tes
$\theta > 1$	Kemampuan tinggi
$-1 \leq \theta \leq 1$	Kemampuan sedang
$\theta < -1$	Kemampuan rendah

* Diadaptasi dari Hidayatulloh (2013)

3.7 Analisis soal

Soal ICAS IPA tahun 2010 dianalisis untuk menentukan tingkat kesukaran soal, daya pembeda soal, efektivitas pengecoh, reliabilitas, fungsi informasi tes, kemampuan peserta tes, peluang menjawab benar, kecocokan soal dengan model logistik 1 parameter, SEM menurut CTT dan IRT.

a. Tingkat kesukaran soal ICAS IPA tahun 2010 menurut CTT dihitung menggunakan program TAP versi 14.7.4 dan diperoleh tingkat kesukaran soal (Lampiran 4, 5, 6). Tingkat kesukaran soal disajikan dalam Tabel 3.6 sebagai berikut.

Tabel 3.6 Tingkat kesukaran soal ICAS IPA tahun 2010 menurut CTT

Analisis tingkat kesukaran soal	Sukar (%)	Sedang (%)	Mudah (%)
Seluruh sekolah	20%	46,67%	33,33%
Kelompok 1	15,55%	42,22%	42,22%
Kelompok 2	22,22%	46,67%	31,11%

b. Daya pembeda soal ICAS IPA tahun 2010 menurut CTT dihitung menggunakan program TAP versi 14.7.4 dan diperoleh daya pembeda soal (Lampiran 4, 5, 6). Daya pembeda soal disajikan dalam Tabel 3.7 sebagai berikut.

Tabel 3.7 Daya pembeda soal ICAS IPA tahun 2010

Analisis daya pembeda	Sangat baik (%)	Lumayan baik (%)	Perlu revisi (%)	Jelek & perlu dibuang (%)
Seluruh sekolah	4,44%	37,78%	31,11%	26,67%
Kelompok 1	6,67%	31,11%	35,56%	26,67%
Kelompok 2	6,67%	26,67%	35,56%	31,11%

c. Efektivitas pengecoh ICAS IPA tahun 2010 dihitung menggunakan program TAP versi 14.7.4 dan diperoleh soal yang berfungsi serta soal yang tidak berfungsi (Lampiran 4, 5, 6). Efektivitas pengecoh disajikan dalam Tabel 3.8 sebagai berikut.

Tabel 3.8 Efektivitas pengecoh soal ICAS IPA tahun 2010

Analisis pengecoh	Berfungsi (%)	Tidak berfungsi (%)
Seluruh sekolah	82,22%	17,78%
Kelompok 1	80%	20%
Kelompok 2	88,89%	11,11%

d. Reliabilitas soal ICAS IPA tahun 2010 dihitung menggunakan program TAP versi 14.7.4. dan diperoleh koefisien reliabilitas (r). r hitung dibandingkan dengan taraf signifikan 5% dengan $n_{penuh}=252$, $n_1=140$ dan $n_2=112$. $r_{penuh}= 0,124$, r_1 tabel = 0,166 dan r_2 tabel = 0,186. Berikut ini disajikan reliabilitas soal ICAS IPA tahun 2010 pada Tabel 3.9.

Tabel 3.9 Reliabilitas soal ICAS IPA tahun 2010

Analisis reliabilitas	Koefisien reliabilitas (r)	Kriteria
Seluruh sekolah	0,683	tinggi
Kelompok 1	0,651	tinggi
Kelompok 2	0,648	tinggi

e. SEM soal ICAS IPA tahun 2010 menurut CTT dihitung menggunakan program TAP versi 14.7.4 dan diperoleh nilai SEM yang disajikan pada Tabel 3.10 sebagai berikut.

Tabel 3.10 SEM soal ICAS IPA tahun 2010 menurut CTT

Analisis SEM	Koefisien SEM
Seluruh sekolah	2,881
Kelompok 1	2,812
Kelompok 2	2,929

f. Fungsi informasi tes ICAS IPA tahun 2010 dihitung menggunakan program BILOG MG pada fase III dan diperoleh nilai fungsi informasi tes atau $I(\theta)$. Fungsi informasi tes dapat dicapai apabila peserta tes memiliki kemampuan tertentu (θ). Fungsi informasi tes dan kemampuan peserta tes soal ICAS IPA tahun 2010 disajikan pada Tabel 3.11 sebagai berikut.

Tabel 3.11 Fungsi informasi tes dan kemampuan peserta tes ICAS IPA tahun 2010

Analisis $I(\theta)$	$I(\theta)$	(θ)
Seluruh sekolah	2,2269	-0.5
Kelompok 1	1,9512	-1.25
Kelompok 2	1,9024	0.125

g. SEM soal ICAS IPA tahun 2010 menurut IRT model logistik 1 parameter dihitung menggunakan program BILOG MG pada fase III dan diperoleh nilai $SEM(\theta)$ yang disajikan pada Tabel 3.12 sebagai berikut.

Tabel 3.12 SEM soal ICAS IPA tahun 2010 menurut IRT model logistik 1 parameter

Analisis SEM	Koefisien $SEM(\theta)$
Seluruh sekolah	0,6701
Kelompok 1	0,7159
Kelompok 2	0,7250

h. Kecocokan soal ICAS IPA tahun 2010 dengan model logistik 1 parameter dihitung menggunakan program BILOG MG pada fase II dan diperoleh hitung (χ^2) yang dibandingkan dengan χ^2 kritis (Lampiran 7, 8, 9). Soal yang memiliki χ^2 hitung $< \chi^2$ kritis dinyatakan cocok dengan model. Kecocokan soal

ICAS IPA tahun 2010 dengan model logistik 1 parameter disajikan pada Tabel 3.13 sebagai berikut.

Tabel 3.13 Kecocokan soal ICAS IPA tahun 2010 dengan model logistik 1 parameter

Analisis kecocokan soal	Cocok (%)	Tidak cocok (%)
Seluruh sekolah	55,55%	44,45%
Kelompok 1	68,89%	31,11%
Kelompok 2	84,44%	15,56%

i. Tingkat kesukaran soal ICAS IPA tahun 2010 menurut IRT model logistik 1 parameter dihitung menggunakan program BILOG MG pada fase II dan diperoleh indeks kesukaran soal (b). Indeks kesukaran soal dianggap baik apabila $-2 \leq b \leq 2$. Indeks kesukaran soal dianggap tidak baik apabila nilai b lebih besar dari 2 dan lebih kecil dari -2. Tingkat kesukaran soal ICAS IPA tahun 2010 menurut IRT disajikan dalam Tabel 3.14 sebagai berikut.

Tabel 3.14 Tingkat kesukaran soal ICAS IPA tahun 2010 menurut IRT

Analisis tingkat kesukaran soal (b)	Baik (%)	Tidak baik (%)
Seluruh sekolah	55,55%	44,45%
Kelompok 1	48,89%	51,11%
Kelompok 2	51,11%	48,89%

3.8 Kriteria kualitas soal

Kriteria kualitas soal diketahui menggunakan tabel klasifikasi kualitas soal. Klasifikasi kualitas soal ICAS IPA tahun 2010 menurut IRT disajikan pada Tabel 3.15, klasifikasi kualitas soal menurut CTT disajikan pada Tabel 3.16, keakuratan hasil analisis soal menurut CTT dan IRT pada Tabel 3.17 sebagai berikut.

Tabel 3.15 Klasifikasi kualitas soal ICAS IPA tahun 2010 menurut IRT

Kriteria	Kecocokan dengan model IRT	Tingkat kesukaran soal (b)
Baik	Cocok	$-2 \leq b \leq 2$
Cukup Baik	Cocok	$b > 2$ atau $b < -2$
Tidak Baik	Tidak cocok	$b > 2$ atau $b < -2$ $-2 \leq b \leq 2$

Tabel 3.16 Klasifikasi kualitas soal ICAS IPA tahun 2010 menurut CTT

Kriteria	Tingkat kesukaran soal (p)	Korelasi biserial	Korelasi Biserial
Baik	$0,3 \leq p \leq 0,7$	$\rho_{bis} \geq 0,3$	Negatif kecuali kunci jawaban
Cukup Baik	$p > 0,7$ atau $p < 0,3$	$\rho_{bis} \geq 0,3$	Negatif kecuali kunci jawaban
	$0,3 \leq p \leq 0,7$	$\rho_{bis} \geq 0,3$	Ada korelasi biserial positif selain kunci jawaban
	$0,3 \leq p \leq 0,7$	$0,2 \leq \rho_{bis} < 0,3$	Negatif kecuali kunci jawaban
Tidak baik	$p > 0,7$ atau $p < 0,3$	-	Ada korelasi biserial positif selain kunci jawaban
	-	$\rho_{bis} < 0,20$	-
	-	$\rho_{bis} < 0,20$	Ada korelasi biserial positif selain kunci jawaban

Tabel 3.17 Klasifikasi keakuratan soal ICAS IPA tahun 2010 menurut CTT dan IRT

Metode pengukuran	Reliabilitas	Informasi tes	SEM	Keakuratan
CTT	Tinggi	-	Rendah	Tinggi
	Rendah	-	Tinggi	Rendah
IRT	-	Tinggi	Rendah	Tinggi
	-	Rendah	Tinggi	Rendah

BAB 4

HASIL PENELITIAN DAN PEMBAHASAN

4.1 Hasil Penelitian

Berdasarkan hasil analisis soal ICAS IPA tahun 2010 menurut teori tes klasik (CTT) diketahui kualitas soal secara keseluruhan, kelompok 1, dan kelompok 2. Soal ICAS IPA tahun 2010 menurut CTT tidak hanya dianalisis secara keseluruhan, tetapi juga dilakukan analisis setiap kelompok. Tujuannya adalah untuk mengetahui ketergantungan kelompok peserta tes terhadap parameter soal penyusun kualitas soal. Kualitas soal ICAS IPA tahun 2010 menurut CTT disajikan pada Tabel 4.1 sebagai berikut.

Tabel 4.1 Kualitas soal ICAS IPA tahun 2010 menurut CTT

Kualitas soal	Baik (%)	Cukup baik (%)	Tidak baik (%)
Seluruh sekolah	24,44%	26,67%	48,89%
Kelompok 1	22,22%	26,67%	51,11%
Kelompok 2	11,11%	20%	68,89%

*data selengkapnya disajikan pada Lampiran 5-7

Berdasarkan Tabel 4.1 diketahui bahwa semakin kecil jumlah peserta tes dalam suatu kelompok maka semakin besar jumlah soal yang berkualitas tidak baik. Kelompok 1 memiliki jumlah soal berkualitas baik dua kali lebih besar daripada kelompok 2. Kelompok 1 dan seluruh sekolah memiliki prosentase jumlah soal berkualitas cukup baik sama besar. Kelompok 2 memiliki jumlah soal dengan kriteria cukup baik paling kecil.

Berdasarkan hasil analisis soal ICAS IPA tahun 2010 menurut teori respons butir (IRT) model logistik 1 parameter diketahui kualitas soal secara keseluruhan, kelompok 1, dan kelompok 2. Soal ICAS IPA tidak hanya dianalisis secara keseluruhan, tetapi juga dilakukan analisis setiap kelompok. Tujuannya adalah untuk mengetahui invariansi tingkat kesukaran soal terhadap kemampuan peserta tes. Kualitas soal ICAS IPA tahun 2010 menurut IRT model logistik 1 parameter disajikan pada Tabel 4.2 sebagai berikut.

Tabel 4.2 Kualitas soal ICAS IPA tahun 2010 menurut IRT model logistik 1 parameter

Kualitas soal	Baik (%)	Cukup baik (%)	Tidak baik (%)
Seluruh sekolah	48%	28,89%	44,44%
Kelompok 1	28,89%	40%	31,11%
Kelompok 2	42,22%	42,22%	15,56%

*data selengkapnya disajikan pada Lampiran 8-11

Kualitas soal ICAS IPA tahun 2010 menurut IRT model logistik 1 parameter diketahui bahwa semakin kecil jumlah peserta tes dalam suatu kelompok maka semakin kecil jumlah soal yang tidak baik. Jika kriteria soal berkualitas baik dan cukup baik dijumlahkan, maka urutan jumlah terbesar ke jumlah terkecil soal yang berkualitas baik adalah kelompok 2, seluruh sekolah, dan kelompok 1. Ada kecenderungan bahwa semakin besar jumlah peserta tes dalam suatu kelompok maka semakin besar jumlah soal yang baik.

Berdasarkan hasil analisis soal ICAS IPA tahun 2010 menurut CTT dan IRT diketahui keakuratan soal secara keseluruhan, kelompok 1, dan kelompok 2. Pada kelompok 1, kelompok 2, maupun secara keseluruhan, analisis soal ICAS IPA tahun 2010 menggunakan pendekatan IRT model logistik 1 parameter lebih akurat daripada CTT. Hal ini ditunjukkan oleh koefisien kesalahan baku pengukuran menurut IRT model logistik 1 parameter lebih kecil daripada menurut CTT. Keakuratan soal ICAS IPA tahun 2010 menurut CTT dan IRT model logistik 1 parameter disajikan pada Tabel 4.3 sebagai berikut.

Tabel 4.3 Keakuratan soal ICAS IPA tahun 2010 menurut CTT dan IRT model logistik 1 parameter

Keakuratan soal	r	I(θ)	(θ)	SEM	SEM(θ)	Keterangan
Seluruh sekolah	0,683	2,2269	-0.5	2,881	0,6701	IRT lebih akurat daripada CTT
Kelompok 1	0,651	1,9512	-1.25	2,812	0,7159	IRT lebih akurat daripada CTT
Kelompok 2	0,648	1,9024	0.125	2,929	0,7250	IRT lebih akurat daripada CTT

*data selengkapnya disajikan pada Lampiran 3 dan 4

Berdasarkan Tabel 4.3 diketahui koefisien SEM(θ) berbanding terbalik kuadrat dengan nilai fungsi informasi tes dan koefisien SEM berbanding terbalik dengan koefisien reliabilitas. Koefisien kesalahan baku pengukuran pada CTT bergantung pada koefisien reliabilitas sedangkan pada IRT bergantung pada

nilai fungsi informasi tes. Kesalahan baku pengukuran pada analisis soal ICAS IPA tahun 2010 di atas, dapat dicapai apabila peserta tes memiliki kemampuan sedang. Hal ini dapat dibuktikan dengan diperolehnya kemampuan peserta tes dalam rentang -1 sampai dengan 1. Kemampuan peserta tes dalam rentang -1 sampai dengan 1 adalah dalam kriteria kemampuan sedang.

4.2 Pembahasan

Kualitas soal ICAS IPA tahun 2010 menurut CTT diketahui bahwa semakin kecil jumlah peserta tes dalam suatu kelompok maka semakin besar jumlah soal yang berkualitas tidak baik. Kelompok 1 memiliki jumlah soal berkualitas baik dua kali lebih besar daripada kelompok 2. Kelompok 1 dan seluruh sekolah memiliki prosentase jumlah soal berkualitas cukup baik sama besar. Kelompok 2 memiliki jumlah soal dengan kriteria cukup baik paling kecil. Perbedaan kualitas soal ICAS IPA tahun 2010 berdasarkan CTT secara keseluruhan, kelompok 1, dan kelompok 2 disebabkan oleh parameter soal penyusunnya. Menurut Depdikbud (1997) kualitas soal menurut CTT ditentukan oleh parameter soal yaitu tingkat kesukaran soal, daya pembeda soal dan efektivitas pengecoh. Apabila suatu soal yang sama diujikan pada kelompok peserta tes berbeda dan memiliki kriteria parameter soal yang berbeda, maka soal akan memiliki kualitas berbeda pula. Misalkan soal nomor 1. Kutipan soal nomor 1 dan pola jawaban peserta tes disajikan sebagai berikut.

1. *Radioisotopes are unstable atoms that decay by emitting alpha, beta and/or gamma radiation. Pure emitters are radioisotopes that emit only one particular type of radiation.*

The table shows a number of radioisotopes and the radiation they emit.

<i>Isotope</i>	<i>Type of radiation emitted</i>
^{241}Am	<i>Alpha, gamma</i>
^{90}Sr	<i>Beta</i>
^{60}Co	<i>Beta, gamma</i>
^{99m}Tc	<i>Gamma</i>
^{238}U	<i>Alpha, gamma</i>
^{137}Cs	<i>Beta, gamma</i>

Which of the following options only lists pure emitters of beta radiation?

- (A) ^{90}Sr , ^{60}Co , and ^{137}Cs
 (B) ^{60}Co and ^{137}Cs
 (C) ^{99m}Tc only
 (D) ^{90}Sr only

Tabel 4.4 Pola jawaban peserta tes pada soal ICAS IPA tahun 2010

	Peserta tes											
	Seluruh kelas				Kelompok 1				Kelompok 2			
Pilihan jawaban	a	b	c	d*	a	b	c	d*	a	b	c	d*
Jumlah peserta	66	9	7	167	28	3	5	102	38	6	2	65
Kelompok atas	6	1	0	66	0	0	0	38	10	0	0	26
Kelompok bawah	30	6	2	41	18	1	4	27	13	6	0	16
Selisih	-24	-5	-2	25	-18	-1	-4	11	-3	-6	0	10

Soal ICAS IPA nomor 1 tahun 2010 secara keseluruhan memiliki kualitas soal baik, pada kelompok 1 cukup baik dan pada kelompok 2 tidak baik. Soal ICAS IPA nomor 1 secara keseluruhan memiliki tingkat kesukaran sedang, daya pembeda lumayan baik, dan korelasi biserial pengecoh negatif. Soal ICAS IPA nomor 1 pada kelompok 1 memiliki tingkat kesukaran mudah, daya pembeda sangat baik, dan korelasi biserial negatif. Soal ICAS IPA nomor 1 pada kelompok 2 memiliki tingkat kesukaran sedang, daya pembeda perlu direvisi, korelasi biserial 2 pengecoh negatif, dan korelasi biserial 1 pengecoh sama dengan nol.

Perbedaan kriteria tingkat kesukaran soal ICAS IPA nomor 1 tahun 2010 secara keseluruhan, pada kelompok 1, dan kelompok 2 disebabkan oleh teknik menghitung indeks kesukaran soal. Indeks kesukaran soal bergantung pada perbandingan banyaknya peserta tes yang menjawab benar suatu soal dengan jumlah seluruh peserta tes. Semakin banyak peserta tes yang menjawab soal dengan benar, maka semakin tinggi indeks kesukaran dalam suatu kelompok atau sebaliknya. Hal ini diperkuat dengan pendapat Hambleton & Swaminathan (1985) yang menjelaskan bahwa pengukuran menggunakan teori klasik memiliki kelemahan yaitu bersifat *group dependent* dan *item dependent*. *Group dependent* artinya hasil pengukuran bergantung pada kelompok peserta yang mengerjakan tes. Jika tes diujikan kepada kelompok peserta dengan kemampuan tinggi, tingkat kesukaran soal akan rendah atau sebaliknya. *Item dependent* artinya hasil pengukuran bergantung pada soal yang diujikan. Jika soal yang diujikan memiliki tingkat kesukaran tinggi, estimasi kemampuan peserta akan rendah atau sebaliknya.

Perbedaan daya pembeda soal ICAS IPA nomor 1 tahun 2010 secara keseluruhan, pada kelompok 1, dan kelompok 2 kemungkinan disebabkan oleh homogenitas kemampuan peserta tes. Sebab soal ICAS IPA tahun 2010 baik secara keseluruhan, pada kelompok 1 maupun pada kelompok 2, hanya memiliki sedikit soal dengan kriteria daya pembeda yang sangat baik. Hal ini sejalan dengan pendapat Hambleton & Swaminathan (1985) yang menjelaskan bahwa besarnya daya pembeda soal bergantung pada homogenitas kelompok peserta tes. Semakin tidak homogen kemampuan peserta tes maka semakin tinggi koefisien daya pembeda soalnya.

Berdasarkan pola jawaban peserta tes pada soal ICAS IPA nomor 1 tahun 2010 diketahui bahwa pada analisis secara keseluruhan dan kelompok 1 memiliki karakteristik pilihan jawaban yang sama (korelasi biserial negatif). Kunci jawaban telah dipilih lebih banyak oleh peserta tes kelompok berkemampuan tinggi daripada kelompok berkemampuan rendah. Opsi jawaban a yang berperan sebagai pengecoh telah berfungsi dengan baik. Pengecoh dipilih lebih banyak oleh kelompok berkemampuan tinggi daripada kelompok berkemampuan rendah. Pengecoh b dan c belum berfungsi dengan baik karena belum dipilih oleh 5% dari jumlah peserta tes. Menurut Surapranata (2006) pengecoh dianggap baik apabila jumlah peserta tes yang memilih pengecoh itu paling sedikit 5% dari jumlah peserta tes. Kelompok 2 memiliki kunci jawaban dan pengecoh a yang telah berfungsi sebagaimana mestinya. Pengecoh b belum berfungsi dengan baik karena tidak dipilih minimal 5% dari jumlah peserta tes, sedangkan pengecoh c tidak berfungsi karena tidak dipilih sama sekali oleh peserta tes.

Berdasarkan uraian diatas, maka dapat diketahui bahwa kegiatan menentukan kualitas soal menurut CTT dapat dilakukan dengan membandingkan parameter tingkat kesukaran, daya pembeda, dan efektivitas pengecoh pada kriteria kualitas soal menurut CTT. Suatu soal yang diujikan pada kelompok peserta tes berbeda akan memiliki kualitas soal berbeda karena parameter soal penyusun kualitas soal bergantung pada kelompok peserta tes yang mengerjakan.

Berdasarkan Tabel 4.2 diketahui kualitas soal ICAS IPA tahun 2010 menurut IRT model logistik 1 parameter secara keseluruhan, kelompok 1, dan kelompok 2. Kualitas soal ICAS IPA tahun 2010 menurut IRT model logistik 1

parameter secara keseluruhan diketahui bahwa semakin kecil jumlah peserta tes dalam suatu kelompok maka semakin kecil jumlah soal yang tidak baik. Jika kriteria soal berkualitas baik dan cukup baik dijumlahkan, maka urutan jumlah terbesar ke jumlah terkecil soal yang berkualitas baik adalah kelompok 2, seluruh sekolah, dan kelompok 1. Ada kecenderungan bahwa semakin besar jumlah peserta tes dalam suatu kelompok maka semakin besar jumlah soal yang baik dan yang tidak baik.

Kualitas soal ICAS IPA tahun 2010 menurut IRT model logistik 1 parameter ditentukan oleh tingkat kesukaran soal dan kecocokan soal dengan model pengukuran yang dipakai. Soal dianggap baik apabila memiliki tingkat kesukaran $-2 \leq b \leq 2$. Soal dianggap tidak baik apabila soal memiliki tingkat kesukaran lebih dari 2 atau kurang dari -2 dan tidak cocok dengan model pengukuran yang dipakai. Soal dianggap cukup baik apabila soal cocok dengan model pengukuran yang dipakai tetapi memiliki tingkat kesukaran soal lebih dari 2 atau kurang dari -2.

Soal ICAS IPA tahun 2010 menurut IRT model logistik 1 parameter tidak hanya dianalisis secara keseluruhan, tetapi juga dilakukan analisis setiap kelompok. Tujuannya adalah untuk mengetahui invariansi tingkat kesukaran soal terhadap kemampuan peserta tes. Tingkat kesukaran soal ICAS IPA tahun 2010 invarian atau tetap terhadap kemampuan peserta tes. Kemiringan kurva *scatter plot* dan garis regresi estimasi kemampuan peserta tes berdasarkan 10 soal termudah dan 10 soal tersukar pada IRT model logistik 1 parameter, menunjukkan adanya korelasi positif antar parameter soal. Hal ini berarti tingkat kesukaran soal tidak tergantung pada kemampuan peserta tes yang digunakan untuk proses kalibrasi tingkat kesukaran soal. Invariansi tingkat kesukaran soal dengan peserta tes merupakan tujuan IRT. Menurut Hambleton & Swaminathan (1985) tujuan IRT adalah membentuk parameter soal dan parameter peserta tes yang bersifat invarian sehingga dapat mengatasi kelemahan-kelemahan CTT yaitu *group dependent* dan *item dependent*. Parameter soal pada IRT model logistik 1 parameter adalah tingkat kesukaran soal.

Berdasarkan Tabel 4.3 diketahui keakuratan soal secara keseluruhan, kelompok 1, dan kelompok 2. Pada kelompok 1, kelompok 2, maupun secara

keseluruhan, analisis soal ICAS IPA tahun 2010 menggunakan pendekatan IRT lebih akurat daripada CTT. Hal ini ditunjukkan oleh koefisien kesalahan baku pengukuran menurut IRT lebih kecil daripada menurut CTT. Koefisien kesalahan baku pengukuran pada CTT bergantung pada koefisien reliabilitas sedangkan pada IRT bergantung pada nilai fungsi informasi tes. Kesalahan baku pengukuran pada analisis soal ICAS IPA tahun 2010 di atas, dapat dicapai apabila peserta tes memiliki kemampuan sedang. Hal ini dapat dibuktikan dengan diperolehnya kemampuan peserta tes dalam rentang -1 sampai dengan 1 . Kemampuan peserta tes dalam rentang -1 sampai dengan 1 adalah dalam kriteria kemampuan sedang.

Jumlah sampel yang digunakan secara keseluruhan lebih besar daripada jumlah sampel pada kelompok 1 dan kelompok 2. Koefisien kesalahan baku pengukuran paling kecil dan nilai fungsi informasi paling besar dimiliki oleh soal yang dianalisis menggunakan IRT secara keseluruhan. Oleh karena itu analisis soal ICAS IPA tahun 2010 menurut IRT pada seluruh sekolah dinyatakan paling akurat dibandingkan analisis pada sampel yang lain. Hal ini membuktikan pendapat Hidayati (2010) yang menjelaskan bahwa secara keseluruhan kesalahan baku pengukuran yang paling kecil diperoleh jika melakukan analisis menggunakan teori respons butir. Klasifikasi kualitas soal yang baik menurut teori respons butir adalah memiliki fungsi informasi tes tinggi dan nilai SEM kecil. Analisis soal menurut teori respons butir akan akurat apabila menggunakan ukuran sampel besar.

Berdasarkan uraian di atas dapat diketahui bahwa soal ICAS IPA tahun 2010 menurut IRT dapat dikerjakan secara maksimal paling tidak oleh peserta tes yang memiliki kemampuan sedang. Soal ICAS IPA tahun 2010 secara keseluruhan, pada kelompok 1 dan kelompok 2 berturut-turut memiliki rata-rata tingkat kesukaran yaitu -0.567 , -0.950 , dan -0.181 . Tingkat kesukaran soal menurut IRT dalam rentang $-1 < b \leq 1$ termasuk dalam kriteria sedang. Hal ini menunjukkan bahwa soal-soal ICAS IPA tahun 2010 dari segi tingkat kesukaran soal layak untuk dijadikan perangkat tes dalam kompetisi internasional karena memiliki tingkat kesukaran sedang dan hanya dapat dikerjakan dengan maksimal oleh peserta tes yang memiliki kemampuan sedang. Adapun kualitas soal ICAS IPA tahun 2010 yang belum sepenuhnya baik menurut Tabel 4.2 dipengaruhi oleh

banyak hal. Salah satunya adalah kecocokan soal dengan IRT model logistik 1 parameter yang digunakan pada penelitian ini. Kecocokan soal dengan IRT model model logistik 1 parameter dapat dicapai dengan jumlah peserta tes tertentu. Menurut Mardapi (1991) sampel atau data respons peserta tes yang dibutuhkan paling tidak 100 orang untuk model Rasch. Selain itu, syarat suatu model agar cocok dengan data tes yaitu harus memenuhi asumsi-asumsi dalam IRT. Menurut Bond & Fox (2007) ukuran sampel untuk model logistik 1 parameter atau model Rasch antara 30 sampai 300. Dari segi jumlah sampel, penelitian analisis soal ICAS IPA tahun 2010 telah memenuhi syarat dan telah memenuhi salah satu asumsi IRT yaitu invarian. Oleh karena itu, kompetisi ICAS dapat menjadi salah satu pilihan kompetisi internasional yang dapat diikuti oleh siswa SMP di Indonesia.

BAB 5

PENUTUP

5.1 Simpulan

Berdasarkan hasil penelitian dan pembahasan, maka dapat disimpulkan bahwa:

- a. Jumlah soal ICAS IPA yang baik menurut IRT lebih banyak daripada CTT.
- b. Analisis soal ICAS IPA tahun 2010 menggunakan pendekatan IRT lebih akurat daripada CTT.

5.2 Saran

Berdasarkan kesimpulan di atas, maka guru atau siapapun yang akan menganalisis soal disarankan sebaiknya menggunakan IRT.

DAFTAR PUSTAKA

- Depdikbud. 1997. *Manual Item And Test Analysis (Iteman)*. Jakarta: Badan Penelitian dan Pengembangan Pendidikan dan Kebudayaan.
- Diknas. 2008. *Panduan Analisis Butir Soal*. Jakarta: Departemen Pendidikan Nasional.
- EAA. 2011. ICAS Science 2010. Australia. On line. Tersedia di <http://www.eaa.unsw.edu.au/icas/subjects/science> [diakses 10-12-2011].
- EAA. 2013. ICAS Science 2010. Australia. On line. Tersedia di <http://www.eaa.unsw.edu.au/icas/subjects/science> [diakses 10-1-2013].
- Azwar, S. 2003. *Tes Prestasi Fungsi dan Pengembangan Pengukuran Prestasi Belajar*. Yogyakarta: Pustaka Pelajar.
- _____. 2007. *Reliabilitas dan Validitas*. Yogyakarta: Pustaka Pelajar.
- Allen, M.J. & W.M. Yen. 2002. *Introduction to measurement theory*. USA. On line. Tersedia di http://books.google.co.id/books?id=MNUyY_csc6cC&printsec=frontcover&dq=allan+introduction+to+measurement+theory+.&hl=en&sa=X&ei=BidCU7ZShYiuB9TxgYAK&ved=0CCsQ6AEwAA#v=onepage&q=allan%20introduction%20to%20measurement%20theory%20.&f=false [diakses 7-4-2014].
- Arikunto, S. 2006. *Dasar-dasar Evaluasi Pendidikan*. Jakarta: Bumi Aksara.
- Baker, F.B. 2001. *The Basics of Item Response Theory*. USA: Educational Resources Information Center.
- Bond, T.G. & Ch.M.Fox. 2007. *Applying the rasch model: Fundamental measurement in the human sciences. 2-nd ed.* New Jersey: Lawrence Erlbaum Associates Publishers.
- Brooks, G.P. & G.A. Johanson. 2003. *Test analysis program. Applied psychological Measurement*. Chicago: AERA.
- Crocker, L. & J. Algina. 2006. *Introduction to classical and modern test theory*. USA: Cengage Learning.
- Dali, S.N. 1992. *Pengantar Teori Sekor Pada Pengukuran Pendidikan*. Jakarta: Gunadarma.
- Ebel, R.L. & D.A. Frisbie. 1991. *Essentials of educational measurement*. New Jersey: Prentice Hall.

- Hambleton, R.K. & H. Swaminathan. 1985. *Fundamentals of item Response Theory*. London: Sage Publications.
- Hambleton, R.K., H. Swaminathan, & H.J. Rogers. 1991. *Fundamentals of item response theory*. Newbury Park, CA: Sage.
- Hayat, B. Pengantar model rasch. Di dalam: Rustam. 1998. *Karakteristik Tes Program Penyetaraan D2PGSD Universitas Terbuka: Implementasi Model Rasch*. Yogyakarta: IKIP Yogyakarta. 103.
- Hidayati, K. 2010. Keakuratan hasil analisis butir menurut teori tes klasik dan teori respons butir ditinjau dari ukuran sampel. Yogyakarta. On line. Tersedia di <http://staff.uny.ac.id/sites/default/files/penelitian/Kana%20Hidayati,%20M.Pd./Keakuratan%20Hasil%20Analisis%20Butir.pdf> [diakses 10-1-2013].
- Hidayatulloh, M.A. 2013. *Analisis kualitas butir soal ulangan semester genap bahasa arab kelas VIII tahun pelajaran 2012/2013 dengan program quest di SMP Muhammadiyah 3 Depok Sleman Daerah Istimewa Yogyakarta*. Skripsi. Yogyakarta. UIN Sunan Kalijaga.
- Kustriyono. 2004. Penyusunan perangkat soal ujian akhir mata pelajaran sains-biologi dalam rangka pengembangan bank soal. *Jurnal Penelitian dan Evaluasi Pendidikan* 2 (6):175-198.
- Lababa, D. 2008. Analisis butir soal dengan teori tes klasik. *IQRA* ' 5 (1):29-37.
- Mardapi, D. 1991. Konsep dasar teori respon butir: Perkembangan dalam pengukuran pendidikan. *Cakrawala Pendidikan* 3(10):1-16.
- _____. 2002. Bukti kesahihan dan keandalan alat ukur: Tanggapan atas artikel tes keterampilan olahraga judo bagi mahasiswa. *Jurnal kependidikan* 1(1):116.
- Nitko, A.J. 2001. *Educational assessment of students second edition*. Merrill: Indianapolis.
- Purwanto, M.N. 2009. *Prinsip-prinsip dan Tehnik Evaluasi Pengajaran*. Bandung: Remaja Rosdakarya.
- Retnawati, H. 2003. *Keberfungsian butir diferensial pada perangkat tes seleksi masuk sekolah lanjutan tingkat pertama (SLTP) mata pelajaran matematika*. Thesis. Yogyakarta: Universitas Negeri Yogyakarta.
- Rudner, L.M. 2001. Item response theory. USA. On line. Tersedia di <http://echo.edres.org:8080/irt/> [diakses 11-10-2013].

- Stocking, M.L. 1999. *Item Response Theory*. In G. Masters and J. Keeves (Eds). *Advances in Measurement in Educational Research and Assessment*. Amsterdam: Pergamon.
- Sudaryono. 2013. *Toeri Responsi Butir*. Yogyakarta: GrahaIlmu.
- Sudjana, N. 2001. *Penilaian Hasil Proses Belajar Mengajar*. Bandung : Remaja Rosdakarya.
- Sulistiyono, S. 2011. *Buku Ajar : Psikometri*. Jakarta : Fakultas Psikologi Universitas Mercubuana.
- Surapranata, S. 2006. *Analisis, Validitas, Reliabilitas dan Interpretasi Hasil Tes*. Bandung: Remaja Rosdakarya.
- Syarifah. 2007. Persyaratan analisis instrument sebagai prasyarat ketepatan hasil analisis dalam penelitian pendidikan. *Jurnal Penelitian MIPA* 1 (1):12-19
- Tayibnapi, F.Y. 2000. *Evaluasi Program*. Jakarta: Rineka Cipta.
- Zimoski, M.F., E. Muraki, R.J. Mislevy, & R.D. Bock. 2003. *Bilog MG*. Chicago: University of Chicago.

LAMPIRAN

Lampiran 1
Perangkat Tes ICAS IPA Tahun 2010



UNSW Global
 THE UNIVERSITY OF NEW SOUTH WALES
 SYDNEY • AUSTRALIA

PAPER
E



2010
 ICAS

International Competitions
 and Assessments for Schools

SCIENCE

**Educational
 Assessment
 Australia**
 eaa.unsw.edu.au

DO NOT OPEN THIS BOOKLET UNTIL INSTRUCTED.

45 QUESTIONS

TIME ALLOWED: 1 HOUR

STUDENT'S NAME:

Read the instructions on the **ANSWER SHEET** and fill in your **NAME, SCHOOL** and **OTHER INFORMATION**.

Use a 2B or B pencil.

Do **NOT** use a pen.

Rub out any mistakes completely.

You **MUST** record your answers on the **ANSWER SHEET**.

Mark only **ONE** answer for each question.

Your score will be the number of correct answers.

Marks are **NOT** deducted for incorrect answers.

There are **45 MULTIPLE-CHOICE QUESTIONS** (1–45).

Use the information provided to choose the **BEST** answer from the four possible options.

On your **ANSWER SHEET** fill in the oval that matches your answer.

You may use a calculator and a ruler.

1. Radioisotopes are unstable atoms that decay by emitting alpha, beta and/or gamma radiation. Pure emitters are radioisotopes that emit only one particular type of radiation.

The table shows a number of radioisotopes and the radiation they emit.

Isotope	Type of radiation emitted
^{241}Am	alpha, gamma
^{90}Sr	beta
^{60}Co	beta, gamma
$^{99\text{m}}\text{Tc}$	gamma
^{238}U	alpha, gamma
^{137}Cs	beta, gamma

Which of the following options only lists pure emitters of beta radiation?

- (A) ^{90}Sr , ^{60}Co and ^{137}Cs
 (B) ^{60}Co and ^{137}Cs
 (C) $^{99\text{m}}\text{Tc}$ only
 (D) ^{90}Sr only
2. The table shows the length of copper wire of different gauges that will produce a resistance equal to 1 ohm.

Electrical Resistance of Copper Wire						
Gauge	6	8	10	12	14	16
Length of wire for 1 ohm of resistance (cm)	81.7	51.4	32.3	20.3	12.8	8.0

Which of the following lengths of copper wire has a resistance of 0.5 ohms for the listed gauge?



- (A) 8 gauge
 (B) 10 gauge
 (C) 12 gauge
 (D) 14 gauge

- 3.

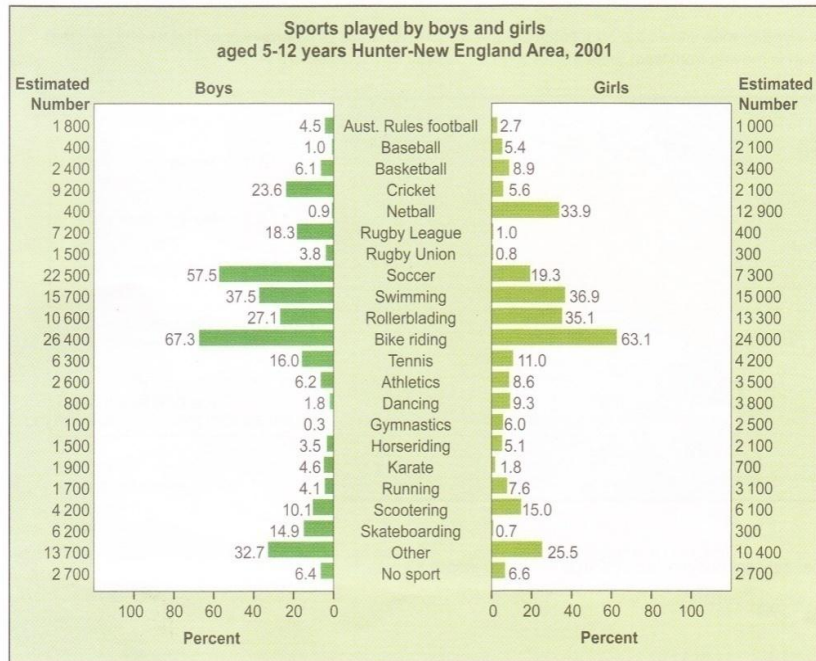


This footprint would most accurately be described as

- (A) left foot, 28 cm long.
 (B) left foot, 33 cm long.
 (C) right foot, 28 cm long.
 (D) right foot, 33 cm long.

For questions 4 and 5 use the information below.

The table shows the estimated number and percentage of boys and girls participating in different sports.

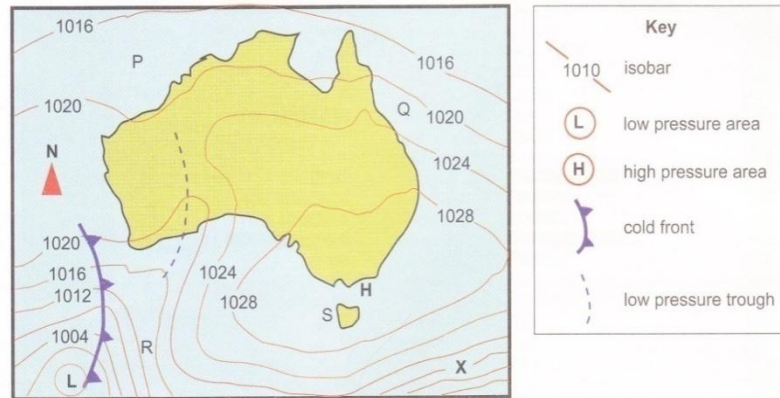


4. According to this information, what was the second most popular sport for girls?
- (A) Bike riding (B) Swimming (C) Netball (D) Soccer
5. Which is the most likely reason why the percentages add up to more than 100%?
- (A) There was a calculation error when converting from number to percentage.
 (B) Children may participate in more than one sport.
 (C) There is an overestimation of numbers.
 (D) Surveys are often wrong.

For questions 6 and 7 use the information below.

An isobar is a line on a weather map joining positions of equal atmospheric pressure. Atmospheric pressure is measured in hectopascals (hPa).

The weather map shows the atmospheric pressure over Australia at a certain time. The weather system shown is moving from west to east.



6. What atmospheric pressure (hPa) is shown by isobar X?

- (A) 1020
- (B) 1016
- (C) 1012
- (D) 1008

7. The closer adjacent isobars are to each other, the higher the wind speed will be.

Four positions P, Q, R and S are shown on the map.

Which position is most likely to have the highest wind speed?

- (A) P
- (B) Q
- (C) R
- (D) S

For questions 8 and 9 use the information below.

Majeda placed a syringe between two pieces of wood so that the trapped air inside the syringe could be compressed as weights were placed on the top piece of wood.



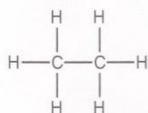
As she added identical weights, Majeda recorded the volume of the trapped air inside the syringe. The table shows her results.

Number of weights	1	2	3	4	5	6
Volume of trapped air (mL)	6.0	3.0	2.0	?	1.2	1.0

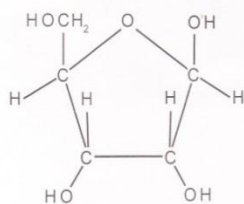
8. From the table it appears that, as the number of weights increased, the volume of trapped air
- (A) increased. (B) decreased.
 (C) stayed the same. (D) decreased then increased.
9. What value would you predict for the volume of air in the syringe when four weights were used?
- (A) 1.3 mL (B) 1.5 mL (C) 1.7 mL (D) 1.9 mL

10. The chemical formula of a substance tells us the type of atoms and the number of each type of atom present in one molecule of the substance.

For example, one molecule of ethane consists of two carbon atoms and six hydrogen atoms. Its chemical formula is C_2H_6 . This may be drawn as a structural formula.



The diagram below shows the structural formula of a molecule of a sugar called ribose.



What is the chemical formula of ribose?

- (A) $C_4H_9O_4$
 (B) $C_4H_{10}O_5$
 (C) $C_5H_9O_5$
 (D) $C_5H_{10}O_5$

For questions 11 and 12 use the information below.

The table gives information about various life stages of the brown trout.

Stage	Age range	Length	Name of stage
embryo	0–14 days	4 mm diameter	egg
hatching stage	14–30 days	4 mm diameter	hatching egg
larval stage	30 days – 2 months	16–26 mm	alevin
juvenile	2–4 months	less than 10 cm	fry
small trout	4–5 months	greater than 10 cm	fingerling
	22 months	13 cm	
large trout	2+ years	25–45 cm	mature fish
	5–6 years	45 cm	

11. Andrew caught a brown trout which was approximately 11 cm long. He gently returned it to the river.

How old was it and what was the name of the stage?

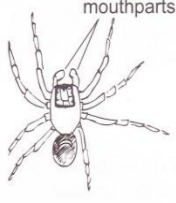

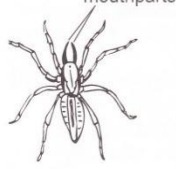

	Age range	Name of stage
(A)	1–2 months old	alevin
(B)	3 months old	fry
(C)	4–5 months old	fingerling
(D)	22 months old	fingerling

12. In some parts of the world, the legal limit for trout is 25 cm. This means that any trout less than 25 cm in length must be returned to the river.

At what age do brown trout typically reach the legal limit?

- (A) 2 months (B) 5 months (C) 2 years (D) 5 years

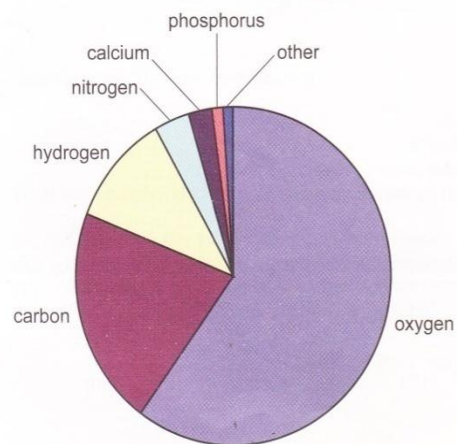
13. The drawings show four different kinds of spiders.

<p>spider 1</p> <p>mouthparts</p> 	<p>spider 2</p> <p>mouthparts</p> 	<p>1</p> <div style="border: 1px solid black; padding: 5px; width: fit-content;"> <p>The spider's mouthparts are rounded.</p> </div> <p>Go to 2</p> <div style="border: 1px solid black; padding: 5px; width: fit-content;"> <p>The spider's mouthparts are pointed.</p> </div> <p>Go to 3</p>
<p>spider 3</p> <p>mouthparts</p> 	<p>spider 4</p> <p>mouthparts</p> 	<p>2</p> <div style="border: 1px solid black; padding: 5px; width: fit-content;"> <p>The two main parts of the spider's body are different sizes.</p> </div> <p>jumping spider</p> <div style="border: 1px solid black; padding: 5px; width: fit-content;"> <p>The two main parts of the spider's body are similar in size.</p> </div> <p>house spider</p>
		<p>3</p> <div style="border: 1px solid black; padding: 5px; width: fit-content;"> <p>The spider's legs are similar in length to its body.</p> </div> <p>wolf spider</p> <div style="border: 1px solid black; padding: 5px; width: fit-content;"> <p>The spider's legs are longer than its body.</p> </div> <p>orb-weaving spider</p>

Using the drawings and the key, what is the name of spider 4?

- | | |
|--------------------|------------------------|
| (A) jumping spider | (B) house spider |
| (C) wolf spider | (D) orb-weaving spider |

14. The graph shows the proportions of different elements in living cells.

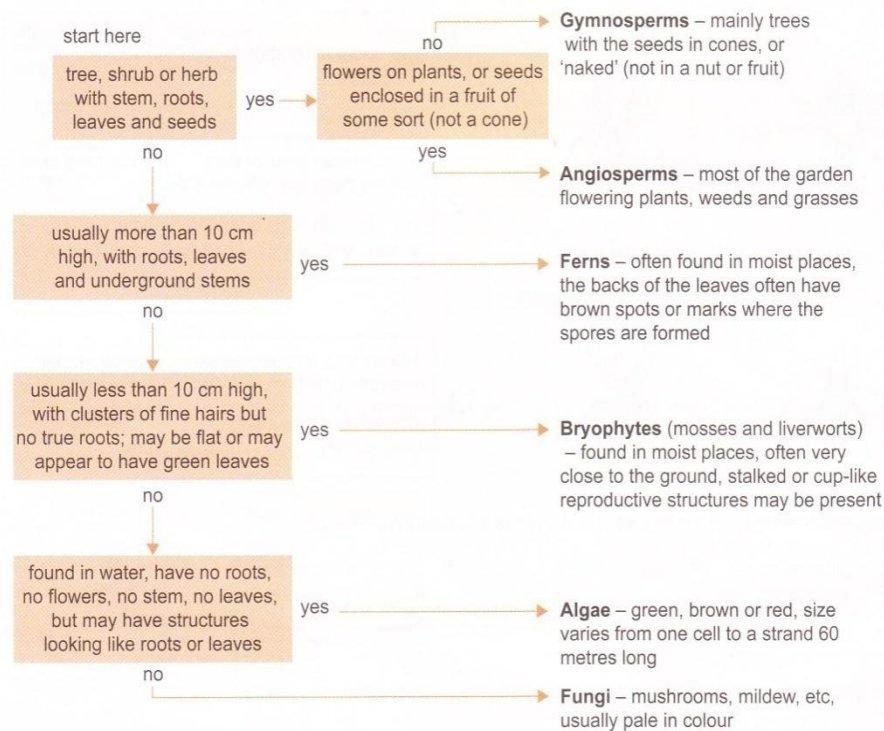


Approximately what proportion of the elements in living cells is either carbon or nitrogen?

- (A) 25%
 (B) 60%
 (C) 80%
 (D) 90%

For questions 15 and 16 use the information below.

The key shows one way to classify plants.



15. Using this key, the best description of a bryophyte is

- (A) a plant with a stem, roots, green leaves and seeds.
- (B) a plant that can usually grow to a height of two metres.
- (C) a plant with no true roots but which appears to have green leaves.
- (D) a plant which may be green, brown or red but with no roots or leaves.

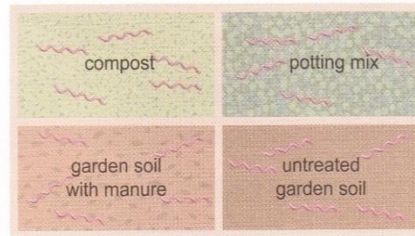
16. A plant was given this description: This green plant is usually about 50 cm high. It does not reproduce by seeds, and prefers a moist place in the garden.

What type of plant is it?

- (A) gymnosperm
- (B) angiosperm
- (C) algae
- (D) fern

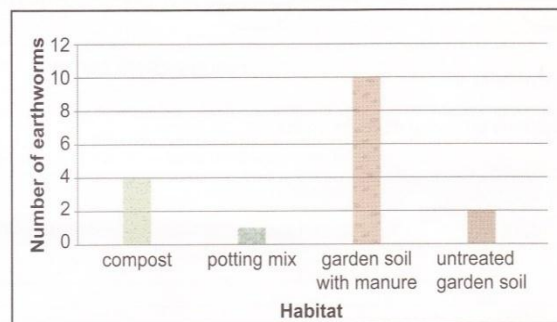
For questions 17 and 18 use the information below.

Jinan is investigating the preferred habitats of earthworms. She created four adjoining habitats and placed five earthworms in each habitat.



The earthworms were free to move between the habitats. After four weeks, Jinan counted the number of earthworms found in each habitat and noticed that some earthworms were missing.

The graph shows Jinan's results.



17. How many earthworms were missing?

- (A) 3 (B) 7 (C) 10 (D) 17




18. What can be concluded from Jinan's experiment?

- (A) The potting mix killed the missing earthworms.
 (B) The missing earthworms preferred habitats that Jinan did not provide.
 (C) Earthworms prefer to live in garden soil with manure rather than in the other habitats.
 (D) Garden soil with manure provides a more nutritious habitat for earthworms than potting mix.

19. Joanne is doing an experiment with vinegar and baking soda. She knows that carbon dioxide gas is produced when vinegar and baking soda are combined.

She is using three identical flasks with a balloon attached to each one to capture the gas. She is using the same type of balloon for each flask.

The table shows her setup and results after 10 minutes.

Flask I	Flask II	Flask III
0 grams baking soda 100 mL vinegar room temperature 21 °C	5 grams baking soda 100 mL vinegar room temperature 21 °C	10 grams baking soda 100 mL vinegar room temperature 21 °C
		

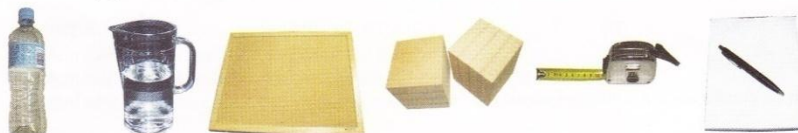
What hypothesis is Joanne testing?

The amount of carbon dioxide produced depends on the

- (A) amount of vinegar used.
- (B) pH of the vinegar solution.
- (C) amount of baking soda used.
- (D) temperature at which the reaction occurs.

20. Rebecca had the following equipment.

plastic bottle jug with water wooden board wooden blocks tape measure notepad and pen

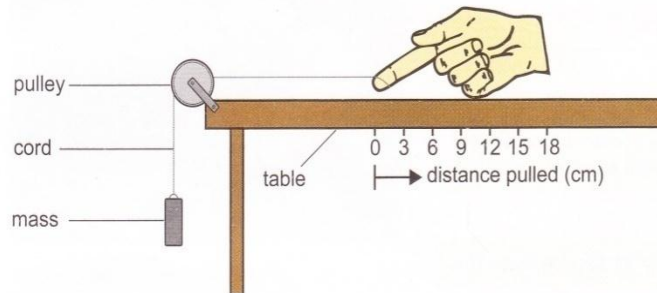


Which question **cannot** be answered using only the listed equipment?

- (A) "Does the type of bottle affect how far it will roll?"
- (B) "Does the amount of water in the bottle affect how far it will roll?"
- (C) "Does the steepness of the ramp affect how far the bottle will roll?"
- (D) "Does the wooden board break when supporting a full bottle of water?"

For questions 21 and 22 use the information below.

Some students investigated how long it took for their index finger muscles to fatigue (become tired). A mass was attached to the index finger as shown in the diagram. The hand was kept stationary on the desk and only the finger moved.



Each student doing a test pulled on the cord to lift the mass up and then released it. They did this as many times in a row as they could. The test was stopped when the student could not pull the mass a distance of more than 3 cm.

In Test 1, a 100 gram mass was attached to the cord.

After waiting 15 minutes, they carried out Test 2 using a 250 gram mass.

Students' results – time until muscles fatigue

Mass (g)	Time (seconds)				
	Student 1	Student 2	Student 3	Student 4	Average
100	80	92	45	35	63
250	55	51	35	35	44

21. Which student recorded the same fatigue time for both masses?

- (A) Student 1 (B) Student 2 (C) Student 3 (D) Student 4

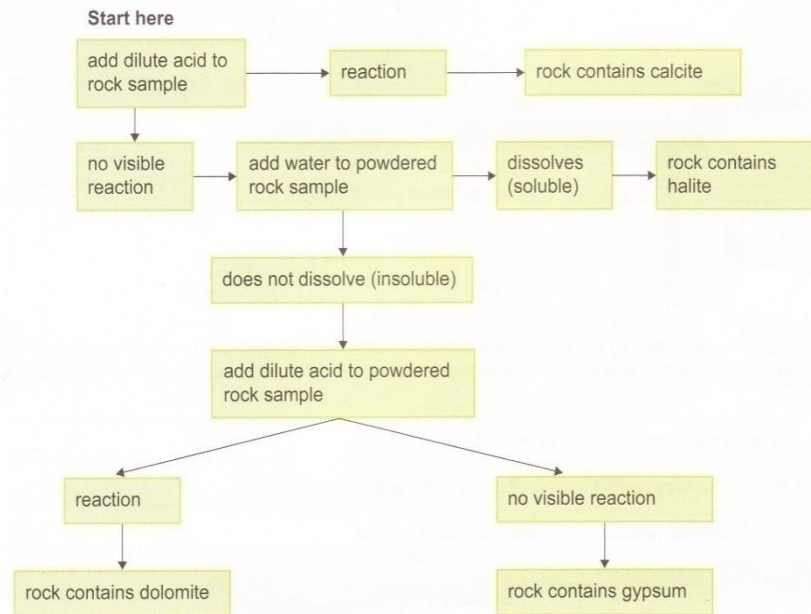
22. What hypothesis could the class have been investigating?

The average fatigue time depends on the

- (A) mass attached to the cord.
 (B) rest time between Tests 1 and 2.
 (C) length of cord attached to the finger.
 (D) distance at which the test is stopped.

23. Calcite, dolomite, gypsum and halite are minerals that are found in some sedimentary rocks.

The diagram shows how simple tests can be used to identify these minerals. If the rock does not initially react with the acid, it is ground into a powder which is then tested.



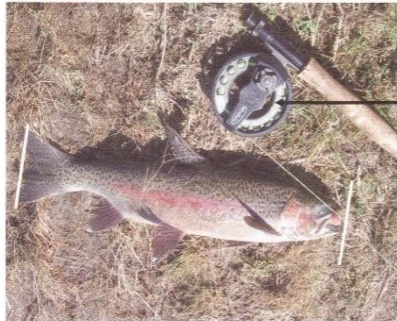
Harry tested four sedimentary rocks that he collected. Here are his results.

Rock	Acid added directly to rock sample	Water added to powder	Acid added to powder
I	no visible reaction	insoluble	reaction
II	no visible reaction	soluble	not tested
III	no visible reaction	insoluble	no visible reaction
IV	reaction	not tested	not tested

From the results of these tests, which rock must contain dolomite?

- (A) Rock I (B) Rock II (C) Rock III (D) Rock IV

24. The photograph shows a trout that Carlyne caught.

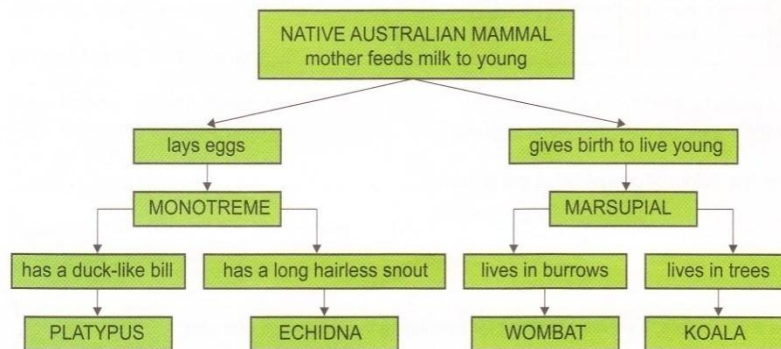


The radius of the reel is 5.0 cm.

What is the approximate length of the trout?

- (A) 75 cm
- (B) 42 cm
- (C) 21 cm
- (D) 10.5 cm

25. The key shows one way to distinguish between four native Australian mammals.

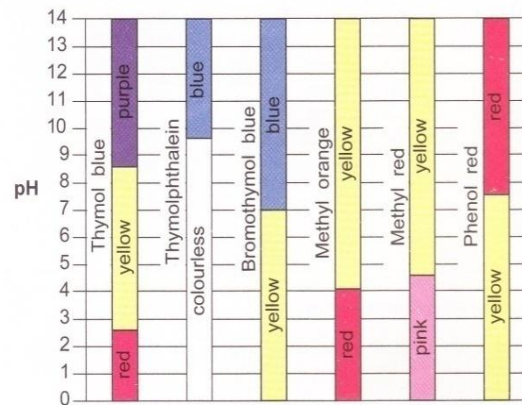


Which characteristic distinguishes a MONOTREME from a MARSUPIAL?

- (A) where it lives
- (B) its appearance
- (C) how it reproduces
- (D) how it feeds its young

The pH of a substance is a measure of how acidic or basic a substance is. Indicators are chemicals that can be used to measure the pH. The indicator is added to a solution of the substance and the final colour shows the pH range for that substance.

The chart shows the colour of some indicators at different pH values.

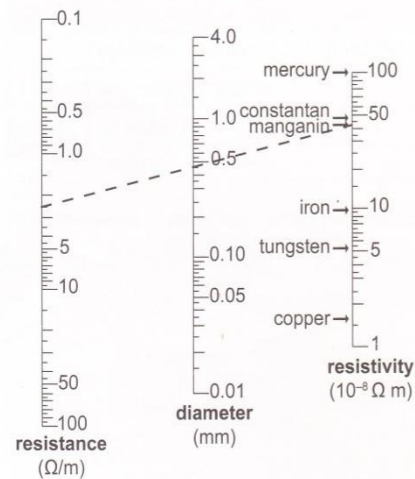


26. A solution with pH 5 is tested with phenol red indicator.

What will be the final colour of phenol red in this solution?

- (A) red (B) blue (C) yellow (D) colourless

27. This diagram is a nomogram. It shows a method for calculating the resistance of a wire.

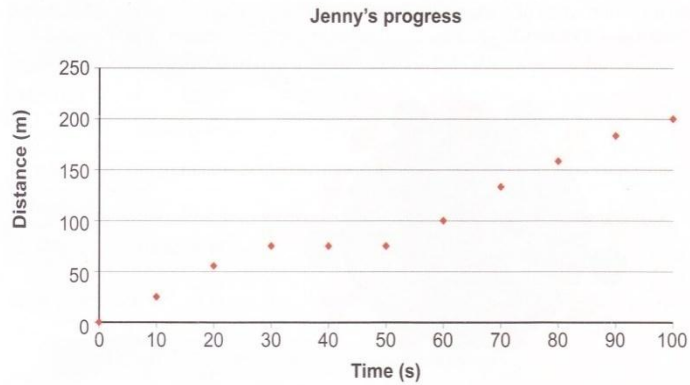


The diagram shows that the resistance for manganin wire with a diameter of 0.45 mm is 2.5 Ω/m .

What diameter of copper wire has a resistance of 0.2 Ω/m ?

- (A) 0.30 mm (B) 0.45 mm (C) 0.70 mm (D) 1.50 mm

28. Jenny took part in a running race. This graph shows Jenny's progress during the race.

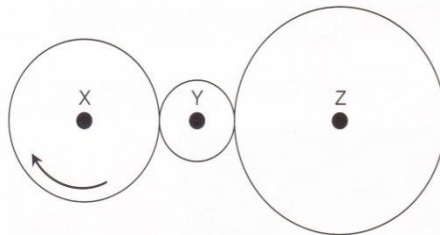


Jenny fell over at one stage of the race, but then she got back up and continued running.

When did Jenny start running again after falling?

- (A) between 30 and 40 seconds (B) between 40 and 50 seconds
 (C) between 50 and 60 seconds (D) between 60 and 70 seconds

29. The photograph shows three gears.

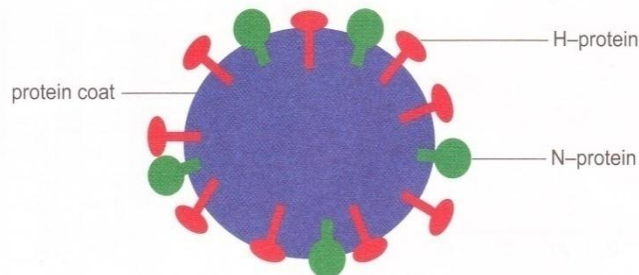


When gear X rotates in a clockwise direction, gear Z will rotate

- (A) clockwise faster than gear X.
 (B) clockwise slower than gear X.
 (C) anti-clockwise faster than gear X.
 (D) anti-clockwise slower than gear X.

30. Viruses consist of genetic material covered by a protein coat.

Influenza viruses are classified by the types of proteins ('H' proteins and 'N' proteins) found on their surfaces. There are 16 known H-proteins and 9 known N-proteins. Each different protein is known by a letter and a number.



'Swine flu' is caused by the H1N1 strain of influenza A, while avian or 'bird flu' is caused by the H5N1 strain of the virus.

Which of the following statements about these two influenza viruses is correct?

- (A) Both strains have different forms of both H-protein and N-protein on their protein coats.
- (B) Both strains have the same forms of H-protein and N-protein on their protein coats.
- (C) Both strains have the same form of H-protein but different forms of N-protein on their protein coats.
- (D) Both strains have the same form of N-protein but different forms of H-protein on their protein coats.

31. An observation is information obtained by sight, smell, touch, hearing or taste. An inference is a possible explanation for an observation.

The diagram shows some fossilised footprints of three early humans that have been preserved in mud.



Some scientists made four statements about the footprints:

- W The footprints of the early humans are similar to the footprints of modern humans.
- X The early humans walked in a way similar to modern humans.
- Y The footprints of the early humans were found in Africa.
- Z The three early humans had different sized feet.

Which statement is an inference?

- (A) W
- (B) X
- (C) Y
- (D) Z

32. The table shows the rising times for the Moon, Jupiter and Mercury over five consecutive days.

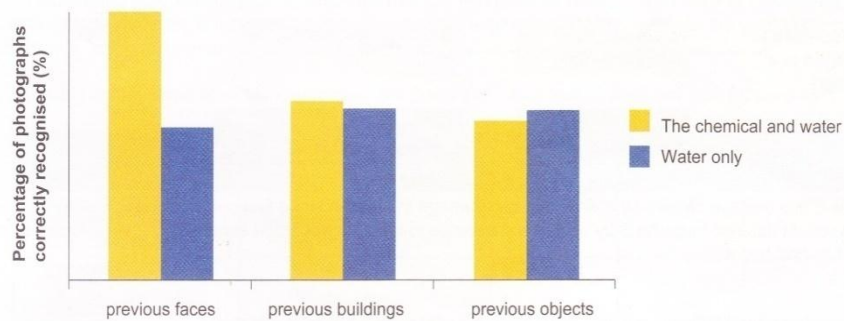
Day	Rising time (24-hour clock)		
	Moon	Jupiter	Mercury
Monday	03:06	15:16	06:36
Tuesday	03:51	15:12	06:31
Wednesday	04:30	15:08	06:25
Thursday	05:03	15:03	06:20
Friday	05:34	14:59	06:14

What conclusion can be made about the changing rising times for each of these astronomical bodies?

	Moon rose	Jupiter rose	Mercury rose
(A)	earlier each day	later each day	later each day
(B)	later each day	earlier each day	earlier each day
(C)	later each day	earlier each day	later each day
(D)	earlier each day	earlier each day	later each day

A group of scientists conducted an experiment to test the effects of a chemical on people's memory. One group was given a nasal spray containing the chemical mixed with water. The other group was given water only. Both groups were shown photos of faces, buildings and objects. On the next day, they were shown photos of faces, buildings and objects again; some of these photos were new while others were the same photos from the previous day. The people in both groups were asked whether they recognised the photos from the previous day.

The graph shows the percentage of photos shown on the first day that were correctly recognised by each of the two groups on the second day.

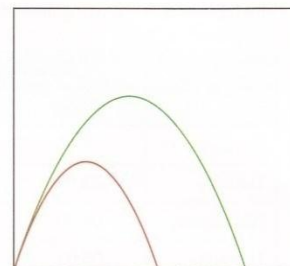


33. What hypothesis do the results support?

- (A) The chemical appears to significantly decrease the ability to remember faces only.
- (B) The chemical appears to significantly increase the ability to remember faces only.
- (C) The chemical appears to significantly decrease the ability to remember buildings and objects.
- (D) The chemical appears to significantly increase the ability to remember faces, buildings and objects.

34. Some students used a launcher to shoot a small rubber ball into the air at a constant angle. They recorded how long it took for the ball to land, for a number of different launch speeds.

Launch speed (m/s)	1	2	4	8
Flight time (s)	0.16	0.31	0.63	1.25



Predict the flight time for a launch speed of 16 m/s at the same angle.

The flight time will be closest to

- (A) 0.25 s. (B) 1.42 s. (C) 1.55 s. (D) 2.51 s.

35. John tested four metals in his laboratory. He found that all of them had low resistance, that is, they allowed electric current to flow easily through them.

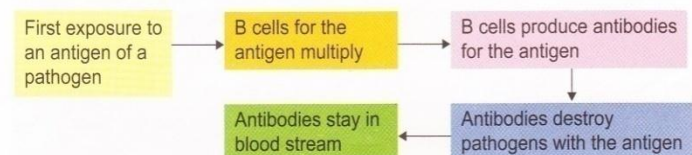
From **this** result it can be concluded that

- (A) all metals have low resistance.
 (B) some metals have low resistance.
 (C) all metals have low resistance and all non-metals have high resistance.
 (D) some metals have low resistance and some metals have high resistance.

A pathogen is an organism that causes disease.

An antigen is a marker on a pathogen. An antigen is used by the human immune system to recognise the pathogen. Then the immune system can use a type of white blood cell called B cells to fight the pathogens, after they have been recognised.

The flowchart shows a possible response from the human immune system on exposure to an antigen of a pathogen.



Hepatitis B is a disease caused by a virus that can damage the liver. A blood test can be done to determine if a healthy patient has had hepatitis B by adding a substance to the patient's blood sample. If the blood clumps, the patient has had hepatitis.

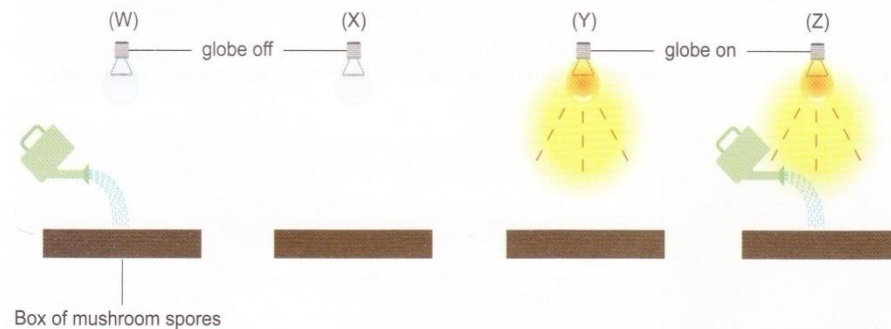
36. What does the added substance clump with to show evidence of a previous infection?

- (A) hepatitis antibodies (B) hepatitis antigens
 (C) DNA of pathogens (D) B cells

Donya is investigating the effect of some environmental factors on the growth of mushrooms. Some of her background research on mushrooms is shown:

- Mushrooms belong to the Kingdom Fungi.
- Fungi reproduce by wind-borne spores.
- Fungi mostly grow in moist and dark areas.

Donya set up four different conditions, W, X, Y and Z, for her experiment.



37. Which setup is most likely to produce the most mushroom growth?

- (A) W (B) X (C) Y (D) Z

38. An advertising light is made from a glass tube that has had all its air replaced with another gas mixture. When high voltage electricity is connected to the tube, the gas begins to glow. The colour of the glow in the glass tube depends on the gas or mixture of gases in the tube.

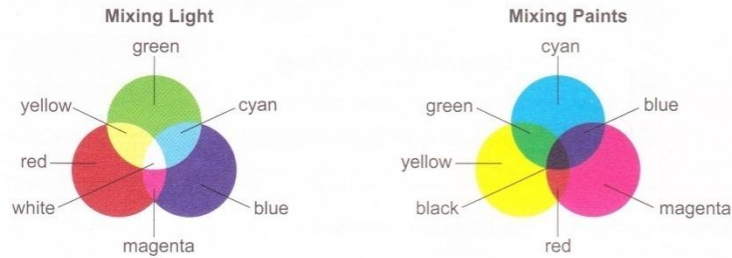
The table shows some advertising signs and the mixture of gases that fill the glass tube for each sign.

Sign	Colour of sign	Neon with mercury	Neon	Argon with phosphor	Carbon dioxide
	red, yellow and blue	✓	✓	✓	
	red		✓		
	white and blue	✓			✓

Which gas or gas mixture produces a blue colour in advertising signs?

- (A) neon only (B) carbon dioxide only
(C) neon with mercury (D) argon with phosphor

39. The diagrams show how the three primary colours of light and the three primary colours of paint can be mixed to produce a variety of other colours.



Alex examines a digital photograph of a red tomato on his computer screen and the same photograph printed by a colour printer. The computer screen mixes light to produce different colours while the colour printer mixes paint.

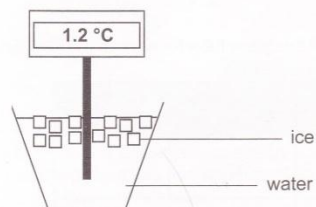


Which colours are used to produce the red colour of the tomato on the computer screen and in the printed photograph?

	Photograph on computer screen	Printed photograph
(A)	red only	red only
(B)	red only	yellow and magenta only
(C)	yellow and magenta only	red only
(D)	yellow and magenta only	yellow and magenta only

40. Cameron wanted to test the accuracy of his digital thermometer.

The temperature of a well-stirred mixture of pure ice and water is known to be exactly $0.0\text{ }^{\circ}\text{C}$. The diagram below shows Cameron's result.



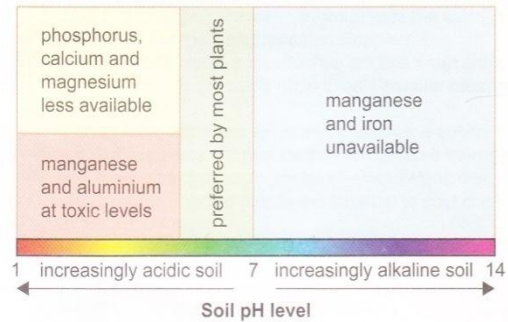
Cameron then measured the temperature in his freezer. The thermometer reading showed $-12.8\text{ }^{\circ}\text{C}$.

What is the actual temperature in the freezer?

- (A) $0.0\text{ }^{\circ}\text{C}$
 (B) $-11.6\text{ }^{\circ}\text{C}$
 (C) $-12.8\text{ }^{\circ}\text{C}$
 (D) $-14.0\text{ }^{\circ}\text{C}$

41. If a substance is present in soil at a toxic level, plants will not grow.

The diagram indicates the availability to plants of chemicals in soil at various levels of soil acidity (pH).



Hydrangea blooms are

- bright pink when the plant is grown in soil with a pH level of 4.5, and
- bright blue when grown in a soil with a pH level of 7.

A gardener who had obtained information from the diagram above added aluminium sulfate to the soil around a hydrangea plant.

What did the gardener believe would happen to the soil pH, and what colour was he trying to give his hydrangea blooms?

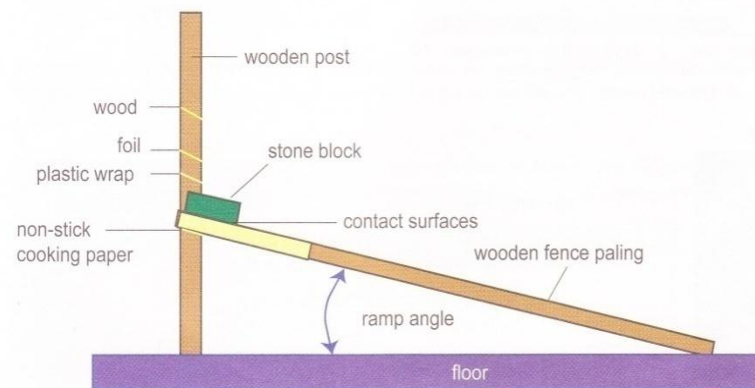
- (A) increased acidity, pink
 (B) increased acidity, blue
 (C) decreased acidity, blue
 (D) decreased acidity, pink

42. Mark investigated the friction force between a stone block and different surfaces.

Mark knew that if you put a solid block on a ramp and start to tilt the ramp upward, the block will be held in place on the ramp by a friction force. As the ramp angle becomes larger, there is a point where the friction force can no longer hold the block in place and the block will start to move.

The friction force is affected by both the surface of the ramp and the surface of the block. The rougher the surfaces the greater the friction force between the ramp and the block.

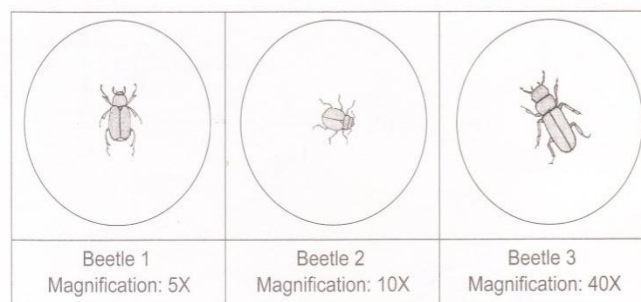
Mark used a fence paling as a ramp. The paling provided a wooden surface. He made the other ramp surfaces by wrapping materials around the paling where it came into contact with the stone block. For each ramp surface, he marked the wooden post at the ramp angle where the stone block began to move. His equipment and markings on the wooden post for each type of material are shown below.



What conclusion can be correctly drawn from **these** observations?

- (A) The friction force between a ramp and an object on the ramp is greater when the ramp angle is smaller.
- (B) The longer the wooden post, the greater the friction force between the ramp and the object on the ramp.
- (C) The stone block moves more quickly over foil than over plastic wrap.
- (D) Non-stick cooking paper has a smoother surface than foil.

43. A student looked at three adult beetles of different species through a microscope. He drew each beetle as it appeared. The diagram shows the magnification used to view each beetle.



Which option lists the beetles in order, from shortest to longest?

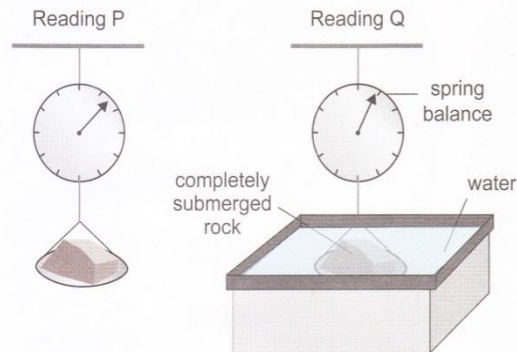
- (A) Beetle 1, Beetle 2, Beetle 3
- (B) Beetle 1, Beetle 3, Beetle 2
- (C) Beetle 3, Beetle 1, Beetle 2
- (D) Beetle 3, Beetle 2, Beetle 1

Jim wanted to determine the density of a piece of rock. He knows that the density of water = 1.0 g/cm^3 and that

$$\text{density} = \frac{\text{mass}}{\text{volume}}$$

He first weighed the rock, and recorded Reading P.

He then lifted a tray full of water up to the rock, so that the level of water completely covered the rock; he then recorded Reading Q.



He also knew that the apparent weight loss of the rock is the same as the weight of water that is displaced by the rock.

44. To determine the density of the rock, Jim would calculate the value of

- (A) $\frac{P - Q}{P}$ (B) $\frac{Q - P}{Q}$ (C) $\frac{Q}{Q - P}$ (D) $\frac{P}{P - Q}$

45. The pressure exerted on the ground by an object is given by the formula:

$$\text{pressure} = \frac{\text{weight}}{\text{area}}$$

Sue took some measurements to find out the pressure exerted on the ground by her pets.

Pet	Weight (newton)	Total area of
Manny (dog)	117	4 paws = 30 cm^2
Sooky (cat)	39	4 paws = 8 cm^2
Boots (guinea pig)	9	4 paws = 3 cm^2
Tweet (canary)	0.25	2 feet = 1 cm^2

Using Sue's measurements, which one of her pets exerts the greatest pressure when all feet are on the ground?

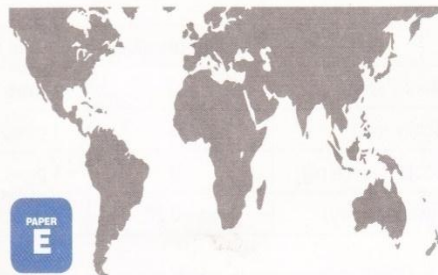
- (A) Manny (B) Sooky (C) Boots (D) Tweet

Acknowledgment

Copyright in this booklet is owned by Educational Assessment Australia, UNSW Global Pty Limited, unless otherwise indicated. Every effort has been made to trace and acknowledge copyright. Educational Assessment Australia apologises for any accidental infringement and welcomes information to redress the situation.

The following year levels should sit THIS Paper:

Australia	Year 7
Brunei	Form 1
Indonesia	Year 8
Malaysia	Form 1
New Zealand	Year 8
Pacific	Year 7
Singapore	Primary 6
South Africa	Grade 7



UNSW Global
THE UNIVERSITY OF NEW SOUTH WALES
SYDNEY • AUSTRALIA

**Educational
Assessment
Australia**
eaa.unsw.edu.au

© 2010 Educational Assessment Australia.
EAA is an education group of UNSW Global Pty Limited, a not-for-profit provider of education, training and consulting services and a wholly owned enterprise of the University of New South Wales. ABN 62 086 418 582

Lampiran 2
Hasil Keluaran (output) Program TAP versi 14.7.4

Item	Key	Number Correct	Item Diff	Disc. Index	# Correct in High Grp	# Correct in Low Grp	Point Biser	Adj PtBis
Item 01	(4)	167	0.66	0.39	66 (0.89)	41 (0.51)	0.35	0.27
Item 02	(3)	74	0.29	0.31	37 (0.50)	15 (0.19)	0.31	0.23
Item 03	(3)	187	0.74	0.22	63 (0.85)	51 (0.63)	0.20	0.12
Item 04	(2)	205	0.81	0.18	68 (0.92)	60 (0.74)	0.23	0.16
Item 05	(2)	208	0.83	0.27	71 (0.96)	56 (0.69)	0.31	0.24
Item 06	(2)	195	0.77	0.29	67 (0.91)	50 (0.62)	0.25	0.17
Item 07	(3)	85	0.34	0.26	33 (0.45)	15 (0.19)	0.31	0.22
Item 08	(2)	217	0.86	0.24	70 (0.95)	57 (0.70)	0.31	0.24
Item 09	(2)	187	0.74	0.39	69 (0.93)	44 (0.54)	0.38	0.30
Item 10	(4)	191	0.76	0.29	66 (0.89)	49 (0.60)	0.31	0.24
Item 11	(3)#	242	0.96	0.06	73 (0.99)	75 (0.93)	0.12	0.09
Item 12	(3)	217	0.86	0.16	72 (0.97)	66 (0.81)	0.18	0.11

marks potential problems (p<0.2 or p>0.95, D<0, pbis<0, adjpbis<0)
 These results have been sorted by item number

```

=====
Number of Items Excluded      = 0
Number of Items Analyzed     = 45
Mean Item Difficulty          = 0.552
Mean Discrimination Index     = 0.261
Mean Point Biserial          = 0.256
Mean Adj. Point Biserial     = 0.175
KR20 (Alpha)                 = 0.683
KR21                         = 0.588
SEM (from KR20)              = 2.881
High Grp Min Score (n=74)    = 28.000
Low Grp Max Score (n=81)     = 22.000
=====

```

```

# Potential Problem Items = 8
defined as:  difficulty <= 0.20(4)
             or:    difficulty >= 0.95(1)
             or:    D index <= 0.00(0)
             or:    AdjPtBiserial <= 0.00(3)
=====

```

 quick options Analysis

 * is keyed answer, # is option that discriminates better than keyed answer
 ~~~~~~

| Item Group | option 1    | option 2   | option 3   | option 4    | option 5  |
|------------|-------------|------------|------------|-------------|-----------|
| 1 TOTAL    | 66 (0.262)  | 9 (0.036)  | 7 (0.028)  | 167*(0.663) | 0 (0.000) |
| High       | 6 (0.081)   | 1 (0.014)  | 0 (0.000)  | 66 (0.892)  | 0 (0.000) |
| Low        | 30 (0.370)  | 6 (0.074)  | 2 (0.025)  | 41 (0.506)  | 0 (0.000) |
| Diff       | -24(-0.289) | -5(-0.061) | -2(-0.025) | 25 (0.386)  | 0 (0.000) |

## Lampiran 3

Hasil keluaran (*output*) BILOG-MG 3 Fase II untuk Mengestimasi Parameter Soal, Fase III untuk Menaksir Fungsi Informasi, Kesalahan Baku Pengukuran, dan Kemampuan Peserta Tes ICAS IPA Tahun 2010

| ITEM     | INTERCEPT<br>S.E. | SLOPE<br>S.E.   | THRESHOLD<br>S.E. | LOADING<br>S.E. | ASYMPTOTE<br>S.E. | CHISQ<br>(PROB) | DF  |
|----------|-------------------|-----------------|-------------------|-----------------|-------------------|-----------------|-----|
| ICASIPA0 | 0.611<br>0.119*   | 0.285<br>0.022* | -2.147<br>0.416*  | 0.274<br>0.022* | 0.000<br>0.000*   | 5.9<br>(0.4320) | 6.0 |
| ICASIPA1 | -0.384<br>0.110*  | 0.285<br>0.022* | 1.350<br>0.386*   | 0.274<br>0.022* | 0.000<br>0.000*   | 9.0<br>(0.1098) | 5.0 |
| ICASIPA2 | 0.911<br>0.126*   | 0.285<br>0.022* | -3.203<br>0.443*  | 0.274<br>0.022* | 0.000<br>0.000*   | 8.0<br>(0.1577) | 5.0 |

.....  
ICAS IPA 45

| GROUP<br>WEIGHT | SUBJECT<br>TEST | IDENTIFICATION |       |       | PERCENT | ABILITY | S.E.     | MARGINAL<br>PROB |
|-----------------|-----------------|----------------|-------|-------|---------|---------|----------|------------------|
|                 |                 | TRIED          | RIGHT |       |         |         |          |                  |
| 1               | 1               |                |       |       |         |         |          |                  |
| 1.00            | ICASIPA         | 45             | 36    | 80.00 | 1.9015  | 0.6085  | 0.000000 |                  |
| 1               | 2               |                |       |       |         |         |          |                  |
| 1.00            | ICASIPA         | 45             | 22    | 48.89 | -0.4711 | 0.5691  | 0.000000 |                  |

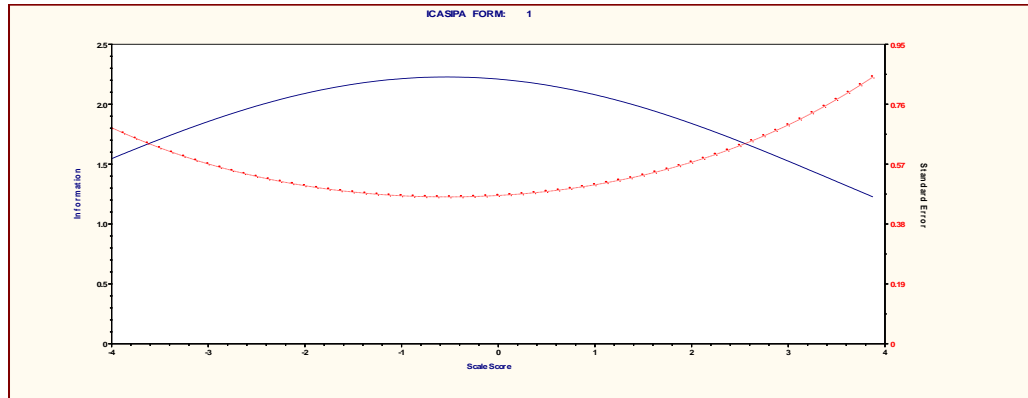
| GROUP<br>WEIGHT | SUBJECT<br>TEST | IDENTIFICATION |       |       | PERCENT | ABILITY | S.E.     | MARGINAL<br>PROB |
|-----------------|-----------------|----------------|-------|-------|---------|---------|----------|------------------|
|                 |                 | TRIED          | RIGHT |       |         |         |          |                  |
| 1               | 1               |                |       |       |         |         |          |                  |
| 1.00            | ICASIPA         | 45             | 36    | 80.00 | 1.7389  | 0.6384  | 0.000000 |                  |
| 1               | 2               |                |       |       |         |         |          |                  |
| 1.00            | ICASIPA         | 45             | 35    | 77.78 | 1.5485  | 0.6325  | 0.000000 |                  |

| GROUP<br>WEIGHT | SUBJECT<br>TEST | IDENTIFICATION |       |       | PERCENT | ABILITY | S.E.     | MARGINAL<br>PROB |
|-----------------|-----------------|----------------|-------|-------|---------|---------|----------|------------------|
|                 |                 | TRIED          | RIGHT |       |         |         |          |                  |
| 1               | 1               |                |       |       |         |         |          |                  |
| 1.00            | ICASIPA         | 45             | 22    | 48.89 | -0.1644 | 0.6055  | 0.000000 |                  |
| 1               | 2               |                |       |       |         |         |          |                  |
| 1.00            | ICASIPA         | 45             | 27    | 60.00 | 0.6646  | 0.6056  | 0.000000 |                  |

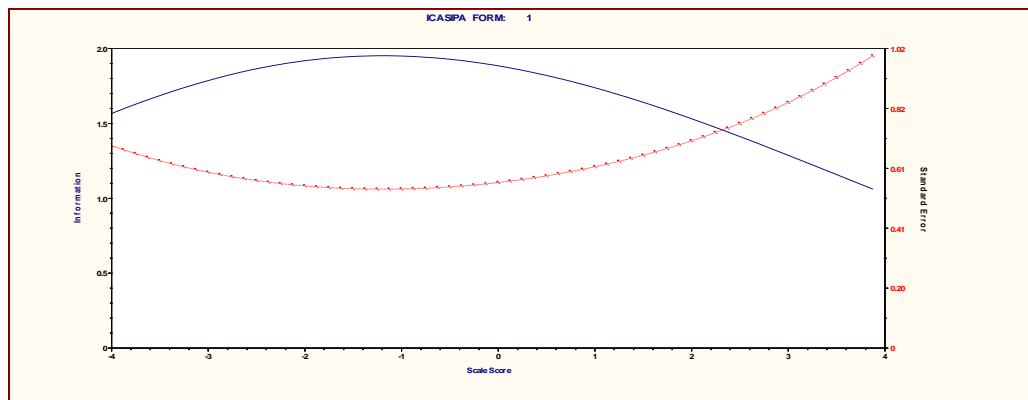
## Lampiran 4

### Kurva Fungsi Informasi Tes dan Kesalahan Baku Pengukuran

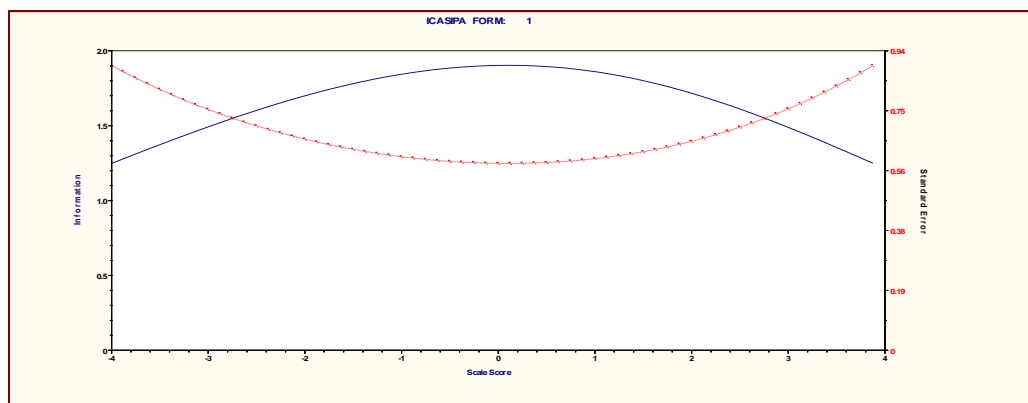
#### Seluruh Sekolah



#### Kelompok 1



#### Kelompok 2



## Lampiran 5

### Kualitas Soal ICAS IPA Tahun 2010 Menurut CTT Secara Keseluruhan

| No | p    | Keterangan | pbiserial | Keterangan              | pbiserial pengecoh                  | Kriteria soal |
|----|------|------------|-----------|-------------------------|-------------------------------------|---------------|
| 1  | 0.66 | Sedang     | 0.35      | Lumayan baik            | Negatif                             | Baik          |
| 2  | 0.29 | Sukar      | 0.31      | Lumayan baik            | Negatif                             | Cukup baik    |
| 3  | 0.74 | Mudah      | 0.20      | Perlu revisi            | Negatif                             | Tidak baik    |
| 4  | 0.81 | Mudah      | 0.23      | Perlu revisi            | 1 negatif, 2 positif                | Tidak baik    |
| 5  | 0.83 | Mudah      | 0.31      | Lumayan baik            | Negatif                             | Cukup baik    |
| 6  | 0.77 | Mudah      | 0.25      | Perlu revisi            | Negatif                             | Tidak baik    |
| 7  | 0.34 | Sedang     | 0.31      | Lumayan baik            | Negatif                             | Baik          |
| 8  | 0.86 | Mudah      | 0.31      | Lumayan baik            | Negatif                             | Cukup baik    |
| 9  | 0.74 | Mudah      | 0.38      | Lumayan baik            | Negatif                             | Cukup baik    |
| 10 | 0.76 | Mudah      | 0.31      | Lumayan baik            | Negatif                             | Cukup baik    |
| 11 | 0.96 | Mudah      | 0.12      | Jelek dan perlu dibuang | 2 negatif, 1 positif, kunci negatif | Tidak baik    |
| 12 | 0.86 | Mudah      | 0.18      | Jelek dan perlu dibuang | Negatif                             | Tidak baik    |
| 13 | 0.83 | Mudah      | 0.28      | Perlu revisi            | Negatif                             | Tidak baik    |
| 14 | 0.39 | Sedang     | 0.07      | Jelek dan perlu dibuang | 2 negatif, 1 positif                | Tidak baik    |
| 15 | 0.88 | Mudah      | 0.33      | Lumayan baik            | Negatif                             | Cukup baik    |
| 16 | 0.57 | Sedang     | 0.27      | Perlu revisi            | 2 negatif, 1 positif                | Tidak baik    |
| 17 | 0.59 | Sedang     | 0.46      | Sangat baik             | Negatif                             | Baik          |
| 18 | 0.54 | Sedang     | 0.22      | Perlu revisi            | Negatif                             | Cukup baik    |
| 19 | 0.85 | Mudah      | 0.25      | Perlu revisi            | Negatif                             | Tidak baik    |
| 20 | 0.19 | Sukar      | 0.19      | Jelek dan perlu dibuang | Negatif                             | Tidak baik    |
| 21 | 0.84 | Mudah      | 0.09      | Jelek dan perlu dibuang | Negatif                             | Tidak baik    |
| 22 | 0.48 | Sedang     | 0.30      | Lumayan baik            | Negatif                             | Baik          |
| 23 | 0.68 | Sedang     | 0.37      | Lumayan baik            | Negatif                             | Baik          |
| 24 | 0.30 | Sedang     | 0.26      | Perlu revisi            | Negatif                             | Cukup baik    |
| 25 | 0.47 | Sedang     | 0.36      | Lumayan baik            | Negatif                             | Baik          |
| 26 | 0.75 | Mudah      | 0.24      | Perlu revisi            | Negatif                             | Tidak baik    |
| 27 | 0.28 | Sukar      | 0.06      | Jelek dan perlu dibuang | 1 negatif, 2 positif                | Tidak baik    |
| 28 | 0.63 | Sedang     | 0.25      | Perlu revisi            | Negatif                             | Cukup baik    |
| 29 | 0.62 | Sedang     | 0.17      | Jelek dan perlu dibuang | Negatif                             | Tidak baik    |
| 30 | 0.51 | Sedang     | 0.29      | Perlu revisi            | Negatif                             | Cukup baik    |
| 31 | 0.12 | Sukar      | 0.08      | Jelek dan perlu dibuang | 2 negatif, 1 positif                | Tidak baik    |
| 32 | 0.42 | Sedang     | 0.24      | Perlu revisi            | 2 negatif, 1 positif                | Tidak baik    |
| 33 | 0.35 | Sedang     | 0.36      | Lumayan baik            | Negatif                             | Baik          |
| 34 | 0.58 | Sedang     | 0.40      | Sangat baik             | Negatif                             | Baik          |
| 35 | 0.28 | Sukar      | 0.08      | Jelek dan perlu dibuang | 1 negatif, 2 positif                | Tidak baik    |
| 36 | 0.22 | Sukar      | 0.25      | Perlu revisi            | 2 negatif, 1 positif                | Tidak baik    |
| 37 | 0.61 | Sedang     | 0.36      | Lumayan baik            | Negatif                             | Baik          |
| 38 | 0.78 | Mudah      | 0.32      | Lumayan baik            | Negatif                             | Cukup baik    |
| 39 | 0.60 | Sedang     | 0.32      | Lumayan baik            | 2 negatif, 1 positif                | Cukup baik    |
| 40 | 0.16 | Sukar      | 0.14      | Jelek dan perlu dibuang | Negatif                             | Tidak baik    |
| 41 | 0.29 | Sukar      | 0.29      | Perlu revisi            | 2 negatif, 1 positif                | Tidak baik    |
| 42 | 0.20 | Sukar      | 0.17      | Jelek dan perlu dibuang | 2 negatif, 1 positif                | Tidak baik    |
| 43 | 0.43 | Sedang     | 0.32      | Lumayan baik            | Negatif                             | Baik          |
| 44 | 0.34 | Sedang     | 0.18      | Jelek dan perlu dibuang | 2 negatif, 1 positif                | Tidak baik    |
| 45 | 0.42 | Sedang     | 0.30      | Lumayan baik            | Negatif                             | Baik          |

**Lampiran 6**  
**Kualitas Soal ICAS IPA Tahun 2010 Menurut CTT pada Kelompok 1**

| No | p    | Keterangan | $\rho_{\text{biserial}}$ | Keterangan              | $\rho_{\text{biserial}}$ pengecoh     | Kriteria soal |
|----|------|------------|--------------------------|-------------------------|---------------------------------------|---------------|
| 1  | 0.73 | Mudah      | 0.41                     | Sangat baik             | Negatif                               | Cukup baik    |
| 2  | 0.35 | Sedang     | 0.37                     | Lumayan baik            | Negatif                               | Baik          |
| 3  | 0.81 | Mudah      | 0.00                     | Jelek dan perlu dibuang | 1 negatif 1 positif 1 tidak berfungsi | Tidak baik    |
| 4  | 0.85 | Mudah      | 0.10                     | Jelek dan perlu dibuang | 2 negatif 1 tidak berfungsi           | Tidak baik    |
| 5  | 0.88 | Mudah      | 0.21                     | Perlu revisi            | Semua negatif termasuk kunci          | Tidak baik    |
| 6  | 0.82 | Mudah      | 0.20                     | Perlu revisi            | Semua negatif termasuk kunci          | Tidak baik    |
| 7  | 0.35 | Sedang     | 0.30                     | Lumayan baik            | Negatif                               | Baik          |
| 8  | 0.86 | Mudah      | 0.27                     | Perlu revisi            | Semua negatif termasuk kunci          | Tidak baik    |
| 9  | 0.75 | Mudah      | 0.45                     | Sangat baik             | Negatif                               | Cukup baik    |
| 10 | 0.79 | Mudah      | 0.32                     | Lumayan baik            | Negatif                               | Cukup baik    |
| 11 | 0.97 | Mudah      | -0.01                    | Jelek dan perlu dibuang | 1 negatif 1 positif 1 tidak berfungsi | Tidak baik    |
| 12 | 0.84 | Mudah      | 0.27                     | Perlu revisi            | Semua negatif termasuk kunci          | Tidak baik    |
| 13 | 0.89 | Mudah      | 0.32                     | Lumayan baik            | Semua negatif termasuk kunci          | Tidak baik    |
| 14 | 0.38 | Sedang     | 0.11                     | Jelek dan perlu dibuang | 1 negatif 2 nol                       | Tidak baik    |
| 15 | 0.90 | Mudah      | 0.28                     | Perlu revisi            | Semua negatif termasuk kunci          | Tidak baik    |
| 16 | 0.67 | Sedang     | 0.22                     | Perlu revisi            | 2 negatif 1 positif                   | Tidak baik    |
| 17 | 0.70 | Sedang     | 0.39                     | Lumayan baik            | Negatif                               | Baik          |
| 18 | 0.60 | Sedang     | 0.23                     | Perlu revisi            | Negatif                               | Cukup baik    |
| 19 | 0.89 | Mudah      | 0.28                     | Perlu revisi            | 2 negatif 1 nol                       | Tidak baik    |
| 20 | 0.11 | Sukar      | 0.21                     | Perlu revisi            | Negatif                               | Cukup baik    |
| 21 | 0.85 | Mudah      | 0.14                     | Jelek dan perlu dibuang | 2 negatif 1 nol                       | Tidak baik    |
| 22 | 0.51 | Sedang     | 0.32                     | Lumayan baik            | Negatif                               | Baik          |
| 23 | 0.71 | Mudah      | 0.27                     | Perlu revisi            | 2 negatif 1 positif kunci negatif     | Tidak baik    |
| 24 | 0.34 | Sedang     | 0.29                     | Perlu revisi            | Negatif                               | Cukup baik    |
| 25 | 0.50 | Sedang     | 0.38                     | Lumayan baik            | Negatif                               | Baik          |
| 26 | 0.74 | Mudah      | 0.26                     | Perlu revisi            | Negatif                               | Cukup baik    |
| 27 | 0.24 | Sukar      | -0.01                    | Jelek dan perlu dibuang | 2 negatif 1 positif                   | Tidak baik    |
| 28 | 0.71 | Mudah      | 0.16                     | Jelek dan perlu dibuang | Semua negatif termasuk kunci          | Tidak baik    |
| 29 | 0.71 | Mudah      | 0.02                     | Jelek dan perlu dibuang | 2 negatif 1 positif kunci negatif     | Tidak baik    |
| 30 | 0.55 | Sedang     | 0.20                     | Perlu revisi            | 2 negatif 1 positif kunci negatif     | Tidak baik    |
| 31 | 0.13 | Sukar      | 0.16                     | Jelek dan perlu dibuang | 1 negatif 2 positif                   | Tidak baik    |
| 32 | 0.48 | Sedang     | 0.16                     | Jelek dan perlu dibuang | Negatif                               | Tidak baik    |
| 33 | 0.42 | Sedang     | 0.37                     | Lumayan baik            | Negatif                               | Baik          |
| 34 | 0.62 | Sedang     | 0.42                     | Sangat baik             | Negatif                               | Baik          |
| 35 | 0.22 | Sukar      | -0.03                    | Jelek dan perlu dibuang | 1 negatif 2 positif kunci negatif     | Tidak baik    |
| 36 | 0.24 | Sukar      | 0.29                     | Perlu revisi            | Negatif                               | Cukup baik    |
| 37 | 0.68 | Sedang     | 0.31                     | Lumayan baik            | 2 negatif 1 tidak berfungsi           | Cukup baik    |
| 38 | 0.82 | Mudah      | 0.34                     | Lumayan baik            | Negatif                               | Cukup baik    |
| 39 | 0.65 | Sedang     | 0.33                     | Lumayan baik            | 2 negatif 1 positif                   | Cukup baik    |
| 40 | 0.16 | Sukar      | 0.15                     | Jelek dan perlu dibuang | 2 negatif 1 positif                   | Tidak baik    |
| 41 | 0.32 | Sedang     | 0.33                     | Lumayan baik            | Negatif                               | Baik          |
| 42 | 0.19 | Sukar      | 0.21                     | Perlu revisi            | 2 negatif 1 positif                   | Tidak baik    |
| 43 | 0.51 | Sedang     | 0.35                     | Lumayan baik            | Negatif                               | Baik          |
| 44 | 0.35 | Sedang     | 0.29                     | Perlu revisi            | Negatif                               | Cukup baik    |
| 45 | 0.53 | Sedang     | 0.30                     | Lumayan baik            | Negatif                               | Baik          |

**Lampiran 7**  
**Kualitas Soal ICAS IPA Tahun 2010 Menurut CTT pada Kelompok 2**

| No | p    | Keterangan | pbiserial | Keterangan              | pbiserial pengecoh          | Kriteria soal |
|----|------|------------|-----------|-------------------------|-----------------------------|---------------|
| 1  | 0.58 | Sedang     | 0.22      | Perlu revisi            | 2 negatif 1 nol             | Tidak baik    |
| 2  | 0.22 | Sukar      | 0.17      | Jelek dan perlu dibuang | 2 negatif 1 nol             | Tidak baik    |
| 3  | 0.65 | Sedang     | 0.31      | Lumayan baik            | Negatif                     | Baik          |
| 4  | 0.77 | Mudah      | 0.32      | Lumayan baik            | 2 negatif 1 tidak berfungsi | Tidak baik    |
| 5  | 0.76 | Mudah      | 0.34      | Lumayan baik            | 2 negatif 1 nol             | Tidak baik    |
| 6  | 0.71 | Mudah      | 0.25      | Perlu revisi            | 2 negatif 1 positif         | Tidak baik    |
| 7  | 0.32 | Sedang     | 0.33      | Lumayan baik            | Negatif                     | Baik          |
| 8  | 0.87 | Mudah      | 0.40      | Sangat baik             | Negatif                     | Cukup baik    |
| 9  | 0.73 | Mudah      | 0.33      | Lumayan baik            | Negatif                     | Cukup baik    |
| 10 | 0.72 | Mudah      | 0.29      | Perlu revisi            | Negatif                     | Tidak baik    |
| 11 | 0.95 | Mudah      | 0.22      | Perlu revisi            | 2 negatif 1 positif         | Tidak baik    |
| 12 | 0.88 | Mudah      | 0.13      | Jelek dan perlu dibuang | 2 negatif 1 nol             | Tidak baik    |
| 13 | 0.77 | Mudah      | 0.17      | Jelek dan perlu dibuang | 2 negatif 1 nol             | Tidak baik    |
| 14 | 0.41 | Sedang     | 0.06      | Jelek dan perlu dibuang | 2 negatif 1 positif         | Tidak baik    |
| 15 | 0.87 | Mudah      | 0.38      | Lumayan baik            | Negatif                     | Cukup baik    |
| 16 | 0.45 | Sedang     | 0.21      | Perlu revisi            | 2 negatif 1 positif         | Tidak baik    |
| 17 | 0.45 | Sedang     | 0.43      | Sangat baik             | 2 negatif 1 nol             | Cukup baik    |
| 18 | 0.47 | Sedang     | 0.14      | Jelek dan perlu dibuang | 2 negatif 1 nol             | Tidak baik    |
| 19 | 0.79 | Mudah      | 0.17      | Jelek dan perlu dibuang | Negatif                     | Tidak baik    |
| 20 | 0.29 | Sukar      | 0.35      | Lumayan baik            | 2 negatif 1 positif         | Tidak baik    |
| 21 | 0.82 | Mudah      | 0.01      | Jelek dan perlu dibuang | 2 negatif 1 positif         | Tidak baik    |
| 22 | 0.43 | Sedang     | 0.25      | Perlu revisi            | 2 negatif 1 positif         | Tidak baik    |
| 23 | 0.65 | Sedang     | 0.48      | Sangat baik             | Negatif                     | Baik          |
| 24 | 0.24 | Sukar      | 0.18      | Jelek dan perlu dibuang | 2 negatif 1 positif         | Tidak baik    |
| 25 | 0.43 | Sedang     | 0.32      | Lumayan baik            | Negatif                     | Baik          |
| 26 | 0.75 | Mudah      | 0.26      | Perlu revisi            | Negatif                     | Tidak baik    |
| 27 | 0.32 | Sedang     | 0.21      | Perlu revisi            | Negatif                     | Tidak baik    |
| 28 | 0.54 | Sedang     | 0.27      | Perlu revisi            | 2 negatif 1 positif         | Tidak baik    |
| 29 | 0.52 | Sedang     | 0.22      | Perlu revisi            | Negatif                     | Tidak baik    |
| 30 | 0.46 | Sedang     | 0.37      | Lumayan baik            | 2 negatif 1 nol             | Cukup baik    |
| 31 | 0.11 | Sukar      | -0.05     | Jelek dan perlu dibuang | 1 negatif 2 positif         | Tidak baik    |
| 32 | 0.36 | Sedang     | 0.27      | Perlu revisi            | Negatif                     | Cukup baik    |
| 33 | 0.27 | Sukar      | 0.29      | Perlu revisi            | Negatif                     | Tidak baik    |
| 34 | 0.54 | Sedang     | 0.37      | Lumayan baik            | 2 negatif 1 nol             | Cukup baik    |
| 35 | 0.36 | Sedang     | 0.31      | Lumayan baik            | 2 negatif 1 nol             | Cukup baik    |
| 36 | 0.21 | Sukar      | 0.20      | Perlu revisi            | 2 negatif 1 positif         | Tidak baik    |
| 37 | 0.52 | Sedang     | 0.34      | Lumayan baik            | Negatif                     | Baik          |
| 38 | 0.73 | Mudah      | 0.27      | Perlu revisi            | Negatif                     | Tidak baik    |
| 39 | 0.53 | Sedang     | 0.26      | Perlu revisi            | Negatif                     | Cukup baik    |
| 40 | 0.17 | Sukar      | 0.15      | Jelek dan perlu dibuang | 1 negatif 1 positif 1 nol   | Tidak baik    |
| 41 | 0.25 | Sukar      | 0.22      | Perlu revisi            | 1 negatif 2 positif         | Tidak baik    |
| 42 | 0.21 | Sukar      | 0.15      | Jelek dan perlu dibuang | 2 negatif 1 positif         | Tidak baik    |
| 43 | 0.32 | Sedang     | 0.19      | Jelek dan perlu dibuang | 2 negatif 1 positif         | Tidak baik    |
| 44 | 0.32 | Sedang     | 0.05      | Jelek dan perlu dibuang | 1 negatif 2 positif         | Tidak baik    |
| 45 | 0.29 | Sukar      | 0.16      | Jelek dan perlu dibuang | 2 negatif 1 positif         | Tidak baik    |

## Lampiran 8

### Kualitas Soal ICAS IPA Tahun 2010 Menurut IRT secara Keseluruhan

| No | b      | Keterangan | chisq | dk | x kritis | status      | Kriteria soal |
|----|--------|------------|-------|----|----------|-------------|---------------|
| 1  | -1.404 | Baik       | 8.9   | 7  | 14.07    | Cocok       | Baik          |
| 2  | 1.823  | Baik       | 29    | 7  | 14.07    | Tidak cocok | Tidak baik    |
| 3  | -2.192 | Tidak baik | 13    | 8  | 15.51    | Cocok       | Cukup baik    |
| 4  | -3.045 | Tidak baik | 8.1   | 8  | 15.51    | Cocok       | Cukup baik    |
| 5  | -3.209 | Tidak baik | 10.3  | 8  | 15.51    | Cocok       | Cukup baik    |
| 6  | -2.548 | Tidak baik | 14    | 8  | 15.51    | Cocok       | Cukup baik    |
| 7  | 1.405  | Baik       | 12.3  | 6  | 12.59    | Cocok       | Baik          |
| 8  | -3.760 | Tidak baik | 16.8  | 8  | 15.51    | Tidak cocok | Tidak baik    |
| 9  | -2.192 | Tidak baik | 14.7  | 7  | 14.07    | Tidak cocok | Tidak baik    |
| 10 | -2.366 | Tidak baik | 6     | 8  | 15.51    | Cocok       | Cukup baik    |
| 11 | -6.487 | Tidak baik | 0.6   | 3  | 7.81     | Cocok       | Cukup baik    |
| 12 | -3.760 | Tidak baik | 6.4   | 6  | 12.59    | Cocok       | Cukup baik    |
| 13 | -3.323 | Tidak baik | 9.2   | 7  | 14.07    | Cocok       | Cukup baik    |
| 14 | 0.907  | Baik       | 14.4  | 9  | 16.92    | Cocok       | Baik          |
| 15 | -4.195 | Tidak baik | 10    | 7  | 14.07    | Cocok       | Cukup baik    |
| 16 | -0.598 | Baik       | 9.4   | 9  | 16.92    | Cocok       | Baik          |
| 17 | -0.734 | Baik       | 16.9  | 8  | 15.51    | Tidak cocok | Tidak baik    |
| 18 | -0.364 | Baik       | 6.4   | 9  | 16.92    | Cocok       | Baik          |
| 19 | -3.503 | Tidak baik | 8.5   | 7  | 14.07    | Cocok       | Cukup baik    |
| 20 | 2.992  | Tidak baik | 12.2  | 5  | 11.07    | Tidak cocok | Tidak baik    |
| 21 | -3.382 | Tidak baik | 3.9   | 8  | 15.51    | Cocok       | Cukup baik    |
| 22 | 0.199  | Baik       | 13.7  | 8  | 15.51    | Cocok       | Baik          |
| 23 | -1.590 | Baik       | 14.8  | 8  | 15.51    | Cocok       | Baik          |
| 24 | 1.784  | Baik       | 6.8   | 6  | 12.59    | Cocok       | Baik          |
| 25 | 0.266  | Baik       | 20    | 6  | 12.59    | Tidak cocok | Tidak baik    |
| 26 | -2.234 | Tidak baik | 2.2   | 8  | 15.51    | Cocok       | Cukup baik    |
| 27 | 1.984  | Baik       | 13.8  | 6  | 12.59    | Tidak cocok | Tidak baik    |
| 28 | -1.151 | Baik       | 5.7   | 9  | 16.92    | Cocok       | Baik          |
| 29 | -1.045 | Baik       | 32.9  | 9  | 16.92    | Tidak cocok | Tidak baik    |
| 30 | -0.065 | Baik       | 21.3  | 7  | 14.07    | Tidak cocok | Tidak baik    |
| 31 | 4.117  | Tidak baik | 8.2   | 7  | 14.07    | Cocok       | Cukup baik    |
| 32 | 0.633  | Baik       | 6     | 9  | 16.92    | Cocok       | Baik          |
| 33 | 1.259  | Baik       | 29.5  | 5  | 11.07    | Tidak cocok | Tidak baik    |
| 34 | -0.700 | Baik       | 21.2  | 8  | 15.51    | Tidak cocok | Tidak baik    |
| 35 | 1.943  | Baik       | 12.5  | 7  | 14.07    | Cocok       | Baik          |
| 36 | 2.595  | Tidak baik | 45.2  | 4  | 9.49     | Tidak cocok | Tidak baik    |
| 37 | -0.905 | Baik       | 21.9  | 8  | 15.51    | Tidak cocok | Tidak baik    |
| 38 | -2.642 | Tidak baik | 16.8  | 8  | 15.51    | Tidak cocok | Tidak baik    |
| 39 | -0.802 | Baik       | 10.2  | 7  | 14.07    | Cocok       | Baik          |
| 40 | 3.382  | Tidak baik | 14.6  | 5  | 11.07    | Tidak cocok | Tidak baik    |
| 41 | 1.863  | Baik       | 42.3  | 6  | 12.59    | Tidak cocok | Tidak baik    |
| 42 | 2.888  | Tidak baik | 24.4  | 4  | 9.49     | Tidak cocok | Tidak baik    |
| 43 | 0.600  | Baik       | 24.7  | 6  | 12.59    | Tidak cocok | Tidak baik    |
| 44 | 1.405  | Baik       | 20.6  | 6  | 12.59    | Tidak cocok | Tidak baik    |
| 45 | 0.633  | Baik       | 31.4  | 7  | 14.07    | Tidak cocok | Tidak baik    |



**Lampiran 9**  
**Kualitas Soal ICAS IPA Tahun 2010 Menurut IRT pada Kelompok 1**

| No | b      | Keterangan | chisq | dk | x kritis | status      | Kriteria soal |
|----|--------|------------|-------|----|----------|-------------|---------------|
| 1  | -2.147 | Tidak baik | 5.9   | 6  | 12.59    | Cocok       | Cukup baik    |
| 2  | 1.350  | Baik       | 9     | 5  | 11.07    | Cocok       | Baik          |
| 3  | -3.203 | Tidak baik | 8     | 5  | 11.07    | Cocok       | Cukup baik    |
| 4  | -3.751 | Tidak baik | 9.5   | 5  | 11.07    | Cocok       | Cukup baik    |
| 5  | -4.271 | Tidak baik | 2.8   | 5  | 11.07    | Cocok       | Cukup baik    |
| 6  | -3.306 | Tidak baik | 4.5   | 6  | 12.59    | Cocok       | Cukup baik    |
| 7  | 1.350  | Baik       | 5.1   | 5  | 11.07    | Cocok       | Baik          |
| 8  | -3.873 | Tidak baik | 4.1   | 6  | 12.59    | Cocok       | Cukup baik    |
| 9  | -2.387 | Tidak baik | 14.6  | 7  | 14.07    | Tidak cocok | Tidak baik    |
| 10 | -2.819 | Tidak baik | 4.4   | 7  | 14.07    | Cocok       | Cukup baik    |
| 11 | -7.518 | Tidak baik | 0     | 0  | 0        | Tidak cocok | Tidak baik    |
| 12 | -3.634 | Tidak baik | 1.4   | 6  | 12.59    | Cocok       | Cukup baik    |
| 13 | -4.416 | Tidak baik | 4.1   | 5  | 11.07    | Cocok       | Cukup baik    |
| 14 | 1.082  | Baik       | 5.6   | 7  | 14.07    | Cocok       | Baik          |
| 15 | -4.733 | Tidak baik | 4.4   | 4  | 9.49     | Cocok       | Cukup baik    |
| 16 | -1.556 | Baik       | 10.5  | 7  | 14.07    | Cocok       | Baik          |
| 17 | -1.843 | Tidak baik | 7.9   | 8  | 15.51    | Cocok       | Cukup baik    |
| 18 | -0.883 | Tidak baik | 8     | 8  | 15.51    | Cocok       | Cukup baik    |
| 19 | -4.416 | Tidak baik | 8.6   | 4  | 9.49     | Cocok       | Cukup baik    |
| 20 | 4.415  | Tidak baik | 10.1  | 2  | 5.99     | Tidak cocok | Tidak baik    |
| 21 | -3.751 | Tidak baik | 1.6   | 5  | 11.07    | Cocok       | Cukup baik    |
| 22 | -0.123 | Baik       | 10.8  | 6  | 12.59    | Cocok       | Baik          |
| 23 | -1.918 | Baik       | 8.7   | 8  | 15.51    | Cocok       | Baik          |
| 24 | 1.419  | Baik       | 5.9   | 7  | 14.07    | Cocok       | Baik          |
| 25 | 0.002  | Baik       | 11.4  | 6  | 12.59    | Cocok       | Baik          |
| 26 | -2.305 | Tidak baik | 3.1   | 6  | 12.59    | Cocok       | Cukup baik    |
| 27 | 2.471  | Tidak baik | 3.1   | 3  | 7.81     | Cocok       | Cukup baik    |
| 28 | -1.918 | Baik       | 2.4   | 7  | 14.07    | Cocok       | Baik          |
| 29 | -1.918 | Baik       | 17.7  | 7  | 14.07    | Tidak cocok | Tidak baik    |
| 30 | -0.436 | Baik       | 10.9  | 7  | 14.07    | Cocok       | Baik          |
| 31 | 4.131  | Tidak baik | 3.2   | 3  | 7.81     | Cocok       | Cukup baik    |
| 32 | 0.189  | Baik       | 10.2  | 8  | 15.51    | Cocok       | Baik          |
| 33 | 0.693  | Baik       | 14.3  | 4  | 9.49     | Tidak cocok | Tidak baik    |
| 34 | -1.079 | Baik       | 17.5  | 8  | 15.51    | Tidak cocok | Tidak baik    |
| 35 | 2.730  | Tidak baik | 6.5   | 4  | 9.49     | Cocok       | Cukup baik    |
| 36 | 2.556  | Tidak baik | 25.9  | 2  | 5.99     | Tidak cocok | Tidak baik    |
| 37 | -1.626 | Baik       | 4.8   | 8  | 15.51    | Cocok       | Baik          |
| 38 | -3.306 | Tidak baik | 13.2  | 5  | 11.07    | Tidak cocok | Tidak baik    |
| 39 | -1.348 | Baik       | 12    | 7  | 14.07    | Cocok       | Baik          |
| 40 | 3.634  | Tidak baik | 11.4  | 5  | 11.07    | Tidak cocok | Tidak baik    |
| 41 | 1.629  | Baik       | 20.2  | 5  | 11.07    | Tidak cocok | Tidak baik    |
| 42 | 3.104  | Tidak baik | 29.6  | 3  | 7.81     | Tidak cocok | Tidak baik    |
| 43 | -0.123 | Baik       | 21.2  | 6  | 12.59    | Tidak cocok | Tidak baik    |
| 44 | 1.350  | Tidak baik | 22.2  | 5  | 11.07    | Tidak cocok | Tidak baik    |
| 45 | -0.248 | Baik       | 19.1  | 7  | 14.07    | Tidak cocok | Tidak baik    |

**Lampiran 10**  
**Kualitas Soal ICAS IPA Tahun 2010 Menurut IRT pada Kelompok 2**

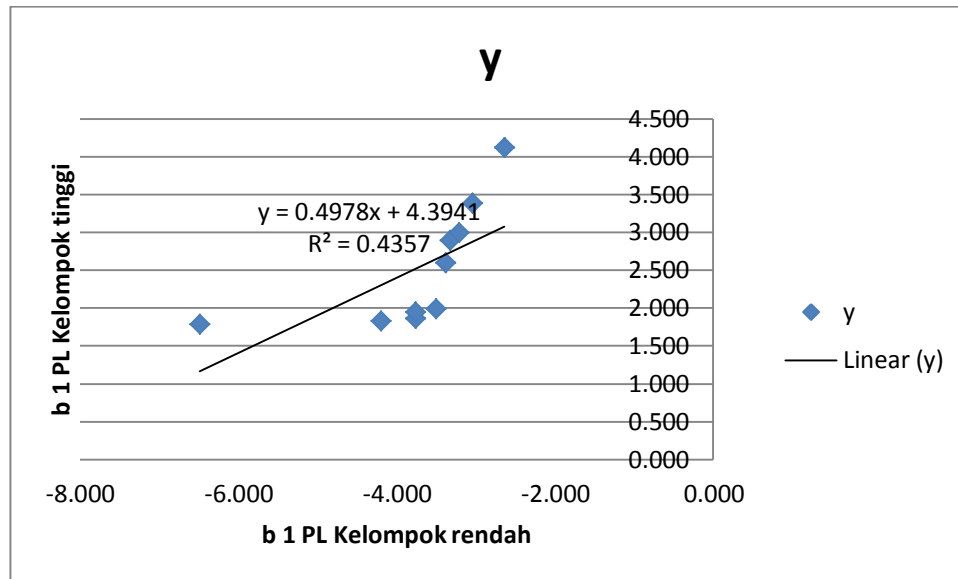
| No | b      | Keterangan | chisq | dk | x kritis | status      | Kriteria soal |
|----|--------|------------|-------|----|----------|-------------|---------------|
| 1  | -0.733 | Baik       | 9.9   | 6  | 12.59    | Cocok       | Baik          |
| 2  | 2.804  | Tidak baik | 14.8  | 4  | 9.49     | Tidak cocok | Tidak baik    |
| 3  | -1.415 | Baik       | 3.3   | 7  | 14.07    | Cocok       | Baik          |
| 4  | -2.691 | Tidak baik | 5.1   | 6  | 12.59    | Cocok       | Cukup baik    |
| 5  | -2.580 | Tidak baik | 5.2   | 6  | 12.59    | Cocok       | Cukup baik    |
| 6  | -2.065 | Tidak baik | 6.1   | 6  | 12.59    | Cocok       | Cukup baik    |
| 7  | 1.684  | Baik       | 2.8   | 5  | 11.07    | Cocok       | Baik          |
| 8  | -4.177 | Tidak baik | 9.2   | 4  | 9.49     | Cocok       | Cukup baik    |
| 9  | -2.265 | Tidak baik | 4.2   | 6  | 12.59    | Cocok       | Cukup baik    |
| 10 | -2.164 | Tidak baik | 3.9   | 6  | 12.59    | Cocok       | Cukup baik    |
| 11 | -6.376 | Tidak baik | 0.3   | 2  | 5.99     | Cocok       | Cukup baik    |
| 12 | -4.537 | Tidak baik | 2.7   | 2  | 5.99     | Cocok       | Cukup baik    |
| 13 | -2.691 | Tidak baik | 3.4   | 6  | 12.59    | Cocok       | Cukup baik    |
| 14 | 0.815  | Baik       | 8.6   | 6  | 12.59    | Cocok       | Baik          |
| 15 | -4.177 | Tidak baik | 14.1  | 4  | 9.49     | Tidak cocok | Tidak baik    |
| 16 | 0.485  | Baik       | 1.9   | 5  | 11.07    | Cocok       | Baik          |
| 17 | 0.485  | Baik       | 6.4   | 6  | 12.59    | Cocok       | Baik          |
| 18 | 0.242  | Baik       | 3     | 7  | 14.07    | Cocok       | Baik          |
| 19 | -3.040 | Tidak baik | 4     | 6  | 12.59    | Cocok       | Cukup baik    |
| 20 | 2.064  | Tidak baik | 4.7   | 4  | 9.49     | Cocok       | Cukup baik    |
| 21 | -3.424 | Tidak baik | 5.5   | 6  | 12.59    | Cocok       | Cukup baik    |
| 22 | 0.649  | Baik       | 6.5   | 6  | 12.59    | Cocok       | Baik          |
| 23 | -1.415 | Baik       | 15    | 5  | 11.07    | Tidak cocok | Tidak baik    |
| 24 | 2.580  | Tidak baik | 5.2   | 4  | 9.49     | Cocok       | Cukup baik    |
| 25 | 0.649  | Baik       | 7.9   | 6  | 12.59    | Cocok       | Baik          |
| 26 | -2.473 | Tidak baik | 2.5   | 6  | 12.59    | Cocok       | Cukup baik    |
| 27 | 1.684  | Baik       | 4.3   | 5  | 11.07    | Cocok       | Baik          |
| 28 | -0.405 | Baik       | 4.1   | 6  | 12.59    | Cocok       | Baik          |
| 29 | -0.162 | Baik       | 10.9  | 7  | 14.07    | Cocok       | Baik          |
| 30 | 0.404  | Baik       | 5.1   | 5  | 11.07    | Cocok       | Baik          |
| 31 | 4.736  | Tidak baik | 3.8   | 3  | 7.81     | Cocok       | Cukup baik    |
| 32 | 1.326  | Baik       | 5.5   | 6  | 12.59    | Cocok       | Baik          |
| 33 | 2.264  | Tidak baik | 9.2   | 4  | 9.49     | Cocok       | Cukup baik    |
| 34 | -0.324 | Baik       | 6.4   | 6  | 12.59    | Cocok       | Baik          |
| 35 | 1.326  | Baik       | 4.6   | 4  | 9.49     | Cocok       | Baik          |
| 36 | 3.040  | Tidak baik | 12.3  | 3  | 7.81     | Tidak cocok | Tidak baik    |
| 37 | -0.162 | Baik       | 13.5  | 5  | 11.07    | Tidak cocok | Tidak baik    |
| 38 | -2.265 | Tidak baik | 9.1   | 6  | 12.59    | Cocok       | Cukup baik    |
| 39 | -0.243 | Baik       | 3.7   | 6  | 12.59    | Cocok       | Baik          |
| 40 | 3.562  | Tidak baik | 4.3   | 3  | 7.81     | Cocok       | Cukup baik    |
| 41 | 2.472  | Tidak baik | 8     | 4  | 9.49     | Cocok       | Cukup baik    |
| 42 | 3.040  | Tidak baik | 10    | 4  | 9.49     | Tidak cocok | Tidak baik    |
| 43 | 1.684  | Baik       | 12.4  | 5  | 11.07    | Tidak cocok | Tidak baik    |
| 44 | 1.684  | Baik       | 5.7   | 5  | 11.07    | Cocok       | Baik          |
| 45 | 1.967  | Baik       | 8.1   | 4  | 9.49     | Cocok       | Baik          |

## Lampiran 11

### Invariansi parameter

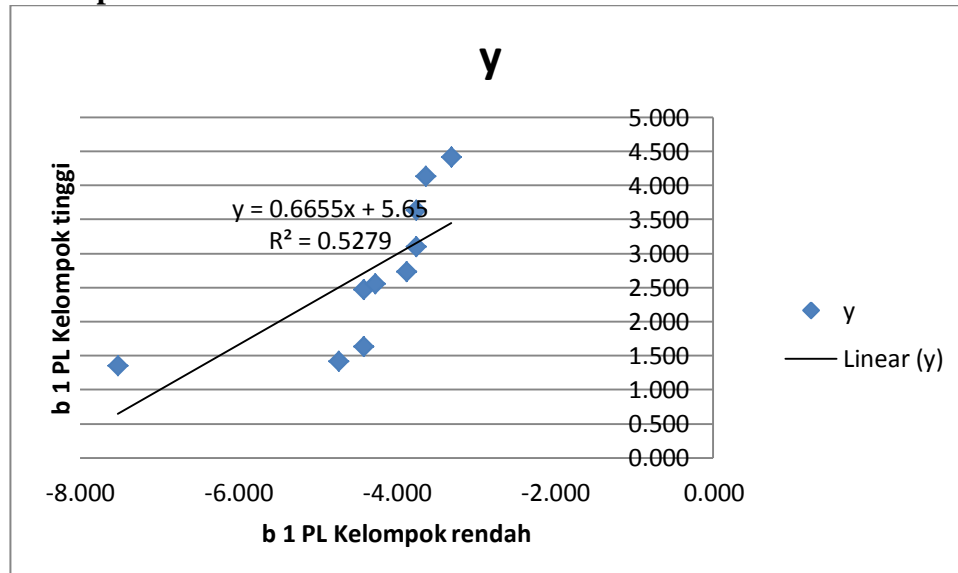
### Scatter Plot dan Garis Regresi Estimasi Kemampuan Berdasarkan 10 Soal Termudah dan 10 Soal Tersukar pada IRT Model Logistik 1 Parameter

#### Seluruh Sekolah

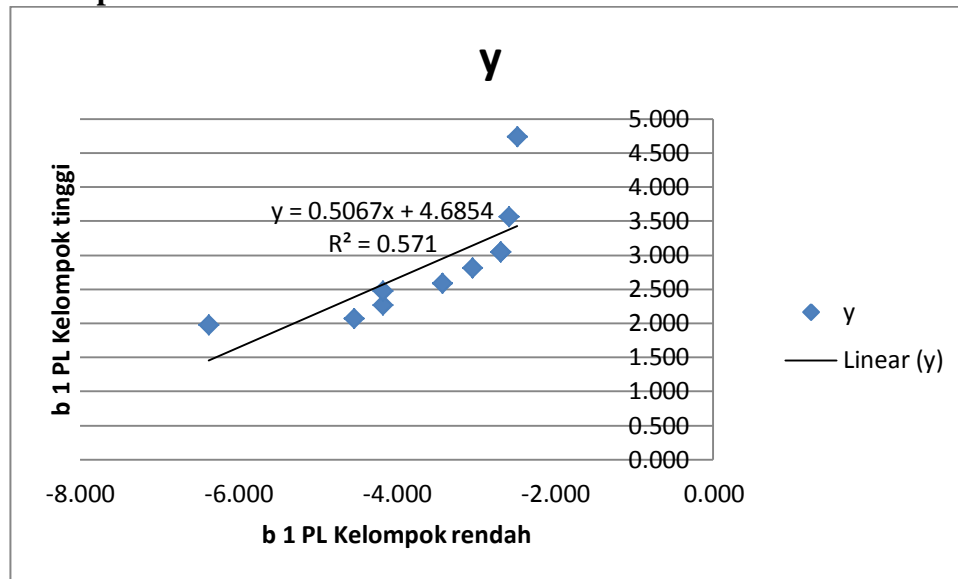


Nilai korelasi *pearson* adalah 0,660

#### Kelompok 1



Nilai korelasi *pearson* adalah 0,727

**Kelompok 2**

Nilai korelasi *pearson* adalah 0,756

## Lampiran 12

### Peluang Menjawab Benar Soal ICAS IPA Tahun 2010

Peluang Menjawab Benar Soal ICAS IPA Tahun 2010 secara Keseluruhan

| No | b      | P( $\theta$ ) | No | b      | P( $\theta$ ) | No | b      | P( $\theta$ ) |
|----|--------|---------------|----|--------|---------------|----|--------|---------------|
| 1  | -1.404 | 0.2582        | 16 | -0.598 | 0.3986        | 31 | 4.117  | 0.3185        |
| 2  | 1.823  | 0.0001        | 17 | -0.734 | 0.0316        | 32 | 0.633  | 0.7432        |
| 3  | -2.192 | 0.1136        | 18 | -0.364 | 0.6978        | 33 | 1.259  | 0.0000        |
| 4  | -3.045 | 0.4219        | 19 | -3.503 | 0.2867        | 34 | -0.700 | 0.0065        |
| 5  | -3.209 | 0.2439        | 20 | 2.992  | 0.0325        | 35 | 1.943  | 0.0859        |
| 6  | -2.548 | 0.0818        | 21 | -3.382 | 0.8696        | 36 | 2.595  | 0.0000        |
| 7  | 1.405  | 0.0546        | 22 | 0.199  | 0.0896        | 37 | -0.905 | 0.0051        |
| 8  | -3.760 | 0.0318        | 23 | -1.590 | 0.0639        | 38 | -2.642 | 0.0320        |
| 9  | -2.192 | 0.0401        | 24 | 1.784  | 0.3430        | 39 | -0.802 | 0.1759        |
| 10 | -2.366 | 0.6501        | 25 | 0.266  | 0.0028        | 40 | 3.382  | 0.0121        |
| 11 | -6.487 | 0.8880        | 26 | -2.234 | 0.9725        | 41 | 1.863  | 0.0000        |
| 12 | -3.760 | 0.3769        | 27 | 1.984  | 0.0320        | 42 | 2.888  | 0.0001        |
| 13 | -3.323 | 0.2381        | 28 | -1.151 | 0.7729        | 43 | 0.600  | 0.0004        |
| 14 | 0.907  | 0.1080        | 29 | -1.045 | 0.0001        | 44 | 1.405  | 0.0022        |
| 15 | -4.195 | 0.1878        | 30 | -0.065 | 0.0034        | 45 | 0.633  | 0.0001        |

Peluang Menjawab Benar Soal ICAS IPA Tahun 2010 pada Kelompok 1

| No | b      | P( $\theta$ ) | No | b      | P( $\theta$ ) | No | b      | P( $\theta$ ) |
|----|--------|---------------|----|--------|---------------|----|--------|---------------|
| 1  | -2.147 | 0.4320        | 16 | -1.556 | 0.1612        | 31 | 4.131  | 0.3658        |
| 2  | 1.350  | 0.1098        | 17 | -1.843 | 0.4461        | 32 | 0.189  | 0.2488        |
| 3  | -3.203 | 0.1577        | 18 | -0.883 | 0.4324        | 33 | 0.693  | 0.0064        |
| 4  | -3.751 | 0.0913        | 19 | -4.416 | 0.0714        | 34 | -1.079 | 0.0254        |
| 5  | -4.271 | 0.7310        | 20 | 4.415  | 0.0064        | 35 | 2.730  | 0.1677        |
| 6  | -3.306 | 0.6064        | 21 | -3.751 | 0.8984        | 36 | 2.556  | 0.0000        |
| 7  | 1.350  | 0.3990        | 22 | -0.123 | 0.0950        | 37 | -1.626 | 0.7763        |
| 8  | -3.873 | 0.6577        | 23 | -1.918 | 0.3709        | 38 | -3.306 | 0.0212        |
| 9  | -2.387 | 0.0408        | 24 | 1.419  | 0.5488        | 39 | -1.348 | 0.1004        |
| 10 | -2.819 | 0.7296        | 25 | 0.002  | 0.0758        | 40 | 3.634  | 0.0444        |
| 11 | -7.518 | 0.0000        | 26 | -2.305 | 0.7930        | 41 | 1.629  | 0.0012        |
| 12 | -3.634 | 0.9641        | 27 | 2.471  | 0.3805        | 42 | 3.104  | 0.0000        |
| 13 | -4.416 | 0.5409        | 28 | -1.918 | 0.9327        | 43 | -0.123 | 0.0017        |
| 14 | 1.082  | 0.5823        | 29 | -1.918 | 0.0136        | 44 | 1.350  | 0.0005        |
| 15 | -4.733 | 0.3538        | 30 | -0.436 | 0.1443        | 45 | -0.248 | 0.0078        |

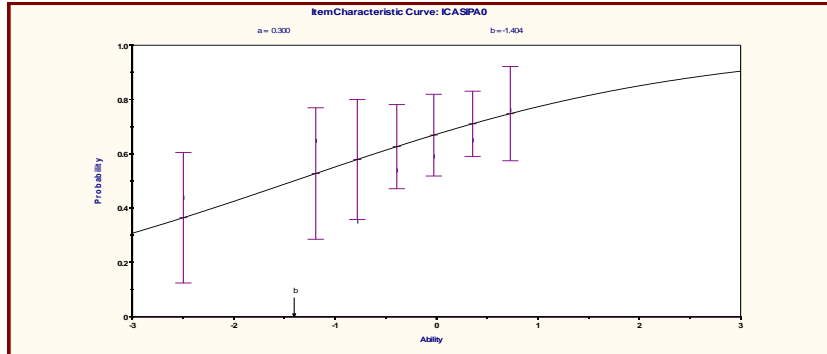
## Peluang Menjawab Benar Soal ICAS IPA Tahun 2010 pada Kelompok 2

| No | b      | P( $\theta$ ) | No | b      | P( $\theta$ ) | No | b      | P( $\theta$ ) |
|----|--------|---------------|----|--------|---------------|----|--------|---------------|
| 1  | -0.733 | 0.1290        | 16 | 0.485  | 0.8667        | 31 | 4.736  | 0.2839        |
| 2  | 2.804  | 0.0051        | 17 | 0.485  | 0.3844        | 32 | 1.326  | 0.4814        |
| 3  | -1.415 | 0.8545        | 18 | 0.242  | 0.8844        | 33 | 2.264  | 0.0565        |
| 4  | -2.691 | 0.5301        | 19 | -3.040 | 0.6817        | 34 | -0.324 | 0.3838        |
| 5  | -2.580 | 0.5139        | 20 | 2.064  | 0.3205        | 35 | 1.326  | 0.3315        |
| 6  | -2.065 | 0.4144        | 21 | -3.424 | 0.4855        | 36 | 3.040  | 0.0063        |
| 7  | 1.684  | 0.7345        | 22 | 0.649  | 0.3654        | 37 | -0.162 | 0.0188        |
| 8  | -4.177 | 0.0564        | 23 | -1.415 | 0.0105        | 38 | -2.265 | 0.1668        |
| 9  | -2.265 | 0.6465        | 24 | 2.580  | 0.2648        | 39 | -0.243 | 0.7137        |
| 10 | -2.164 | 0.6950        | 25 | 0.649  | 0.2448        | 40 | 3.562  | 0.2267        |
| 11 | -6.376 | 0.8427        | 26 | -2.473 | 0.8677        | 41 | 2.472  | 0.0915        |
| 12 | -4.537 | 0.2591        | 27 | 1.684  | 0.5065        | 42 | 3.040  | 0.0402        |
| 13 | -2.691 | 0.7609        | 28 | -0.405 | 0.6600        | 43 | 1.684  | 0.0300        |
| 14 | 0.815  | 0.1943        | 29 | -0.162 | 0.1437        | 44 | 1.684  | 0.3370        |
| 15 | -4.177 | 0.0070        | 30 | 0.404  | 0.3989        | 45 | 1.967  | 0.0879        |

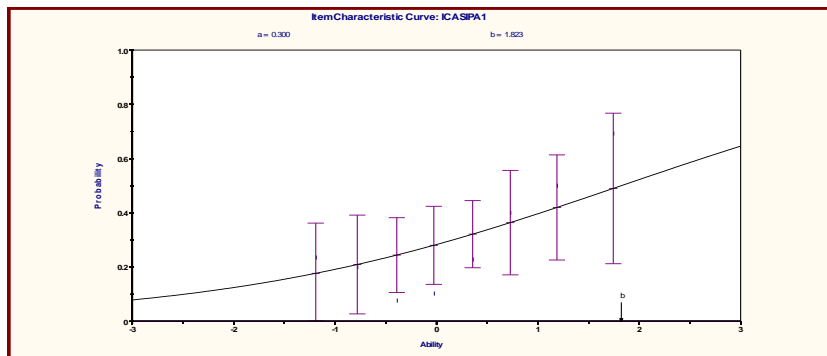
### Lampiran 13

### Kurva Karakteristik Butir Soal ICAS IPA Tahun 2010 secara Keseluruhan

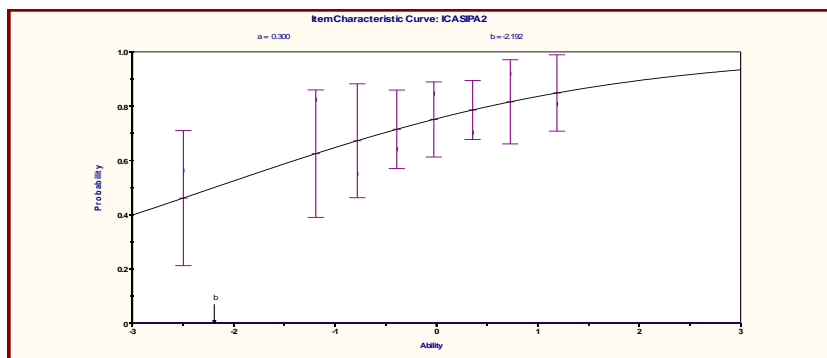
#### Soal Nomor 1



#### Soal Nomor 2



#### Soal Nomor 3



dan lain-lain.

**Lampiran 14**  
**Surat Ijin Penelitian Dinas Pendidikan Kota Semarang**



**PEMERINTAH KOTA SEMARANG**  
**DINAS PENDIDIKAN**

Jl. Dr. Wahidin 118 Semarang Telp. 8412180, Fax. 8317752, Kode Pos 50234

**SURAT IJIN KEPALA DINAS PENDIDIKAN KOTA SEMARANG**

Nomor : 070 / 2859

**TENTANG IJIN OBSERVASI**

Dasar : Surat dari Universitas Negeri Semarang  
 No. 3526/UN37.1.4/LT/2012, Tgl. 15 Mei 2012

Perihal : Ijin Observasi

Berdasarkan hal tersebut di atas, Kepala Dinas Pendidikan Kota Semarang mengijinkan Mahasiswa sebagai berikut :

Nama : Mahanani  
 NIM : 4401408104  
 Perguruan Tinggi : UNNES  
 Fakultas : Matematika dan Ilmu Pengetahuan Alam  
 Jurusan : Pend. Biologi  
 Judul : " Observasi "


Untuk melaksanakan observasi di SMP Negeri 5, SMP Negeri 2, SMP Negeri 21 Semarang.

Dengan memperhatikan hal-hal sebagai berikut

- 1 Observasi tidak mengganggu proses pembelajaran di sekolah.
- 2 Mentaati peraturan dan ketentuan yang berlaku di Sekolah tempat observasi.
- 3 Menyampaikan laporan kepada Kepala Dinas Pendidikan Kota Semarang segera setelah selesai pelaksanaan kegiatan tersebut.
- 4 Observasi dilaksanakan sejak dikeluarkannya surat ijin Kepala Dinas Pendidikan Kota Semarang sampai dengan selesai.

Semarang, 24 Mei 2012

An. Kepala Dinas Pendidikan  
 Kota Semarang  
 Kabid. Monitoring dan Pengembangan

  
 Dr. Ir. Nana Storada DM, SE, MM  
 Pembina  
 NIP. 196403091990031010

Tembusan Yth.

1. Walikota Semarang (sebagai laporan)
2. Kepala Sekolah ybs
3. Peringgal



**Lampiran 15**  
**Surat Keterangan Telah Melakukan Penelitian SMP N 2 Semarang**



PEMERINTAH KOTA SEMARANG  
 DINAS PENDIDIKAN  
**SMP NEGERI 2 SEMARANG**  
 Jl. Brigjen. Katamso No. 14 Telp. (024) 8414168 Fax. (024) 8411211 Semarang-50125  
 Website : [www.smpn2-smg.com](http://www.smpn2-smg.com) e-mail : [smpn2\\_semarang@yahoo.com](mailto:smpn2_semarang@yahoo.com)

**SURAT KETERANGAN**

Nomor : 481 / 423.1 / 2012

Yang bertanda tangan di bawah ini, Kepala SMP Negeri 2 Semarang menerangkan bahwa :

Nama : MAHANANI  
 N I M : 4401408104  
 Semester : 8  
 Jurusan : Biologi  
 Fakultas : FMIPA  
 Perguruan Tinggi : Universitas Negeri Semarang (UNNES)

telah melakukan Penelitian di SMP Negeri 2 Semarang dengan judul “Analisis Soal International Competitions and Assessments for School (ICAS) dengan menggunakan Metode Item Response Theory (IRT) dan Classical Test Theory (CTT)”

Pelaksanaan : 23 s.d. 31 Juli 2012

Demikian, Surat Keterangan ini dibuat untuk dipergunakan sebagaimana mestinya.

Semarang, 1 Agustus 2012  
 Kepala Sekolah,



Drs. Sutomo, A.Md., M.M.  
 Pembina  
 NIP. 19570227 198103 1 010

**Lampiran 16**  
**Surat Keterangan Telah Melakukan Penelitian SMP N 5 Semarang**



**PEMERINTAH KOTA SEMARANG**  
**DINAS PENDIDIKAN**  
**SMP 5**

Rintisan Sekolah Bertaraf Internasional  
 Website: <http://smp5-smg.sch.id> E-Mail: [smpn5smg@gmail.com](mailto:smpn5smg@gmail.com)

Jalan Sultan Agung Telp. (024) 8315140 Fax. (024) 8506183 Kode Pos 50252 Semarang

**SURAT KETERANGAN**

No. 070 / 018 / 2012

1. Dasar : 1. Surat Dekan UNNES No. 3526/UN37.1.4/LT/2012, tanggal: 15/05/2012, Hal. Permohonan Ijin Observasi;  
 2. Surat Kepala Dinas Pendidikan Kota Semarang, No. 070/2859, tanggal: 24/05/2012, tentang Ijin Observasi;

2. Kepala SMP Negeri 5 Semarang, menerangkan bahwa :

1. Nama : **MAHANANI**  
 2. N I M : 4401408104  
 3. Perguruan Tinggi : Universitas Negeri Semarang  
 4. Prodi / Jurusan / Smt. : Pend. Biologi S-1 / Biologi / VIII  
 5. Tahun Pelajaran : 2011-2012

Telah mengadakan / melaksanakan observasi / penelitian pada :

1. Sekolah : SMP Negeri 5 Semarang  
 2. Waktu Pelaksanaan : JUNI 2012  
 3. Lamanya : 1 (satu) hari  
 4. Sasaran : Peserta didik kelas VIII SMP 5 Semarang  
 5. Judul : *Analisis Soal International Competitions and Assessments for Schools (ICAS) dengan Menggunakan Metode Item Response Theory (IRT) dan Classical Test Theory (CTT)*

3. Demikian harap menjadikan maklum.

Semarang, 25 Juli 2012  
 Kepala  
  
 H. Suharto, S.Pd, MM.  
 NIP.19580312 197903 1 009

**Lampiran 17**  
**Surat Permohonan Ijin Observasi SMP N 1 Kudus**



KEMENTERIAN PENDIDIKAN DAN KEBUDAYAAN  
 UNIVERSITAS NEGERI SEMARANG  
 FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
 Gedung D5 Lt.1 Kampus Sekaran Gunungpati Semarang Kode Pos 50229, Telp. (024)8508112  
 Telp. Dekan (024)8508005; Jurusan: Matematika (024)8508032; Fisika (024)8508034; Kimia (024)8508035; Biologi (024)8508033  
 Fax. (024)8508005; Website: <http://mipa.unnes.ac.id>; Email: [mipa@unnes.ac.id](mailto:mipa@unnes.ac.id)

Nomor : 147/UN37.1.4/PP/2012  
 Lampiran : -  
 Hal : *Permohonan Ijin Observasi*

22 Februari 2012

Yth. Kepala SMP Negeri 1 Kudus  
 di Kudus

Kami memberitahukan dengan hormat, bahwa mahasiswa Fakultas Matematika dan Ilmu Pengetahuan Alam (FMIPA) Universitas Negeri Semarang tersebut di bawah ini :

Nama : Mahanani  
 NIM : 4401408104  
 Semester : 8  
 Jurusan : Biologi

dalam rangka tugas mata kuliah Skripsi dengan dosen pembimbing Dr. Saiful Ridlo, M.Si & Dra. Ely Rudyatmi, M.Si bermaksud akan mengadakan observasi di:

Tempat : SMP Negeri 1 Kudus  
 Waktu : 1 Maret 2012

Berkaitan dengan hal ini, kami mohon dapat diberikan ijin observasi kepada mahasiswa yang bersangkutan pada tempat dan jadwal waktu tersebut di atas.

Atas perhatian dan kerja sama Saudara, kami sampaikan terima kasih.



Prof. Dr. Wiyanto, M.Si  
 NIP. 19631012 198803 1 001

Tembusan :  
 1. Ketua Jurusan Biologi;  
 2. Dosen Pembimbing;  
 FMIPA Universitas Negeri Semarang.



**Lampiran 18**  
**Surat Keterangan Telah Melakukan Penelitian SMP N 1 Kudus**



PEMERINTAH KABUPATEN KUDUS  
 DINAS PENDIDIKAN PEMUDA DAN OLAHRAGA  
**SMP 1 KUDUS**  
**RINTISAN SEKOLAH BERTARAF INTERNASIONAL**  
 Jl. Sunan Muria No.10A Telp/Fax (0291) 437929 Kudus Kode Pos 59312  
 Website : www.smp1kudus.com



**SURAT KETERANGAN**

No. 070/361/14.06.03/2012

Kepala SMP 1 Kudus menerangkan bahwa :

| NO | NAMA     | NIM        | PROGRAM STUDI                                           |
|----|----------|------------|---------------------------------------------------------|
| 1  | Mahanani | 4401408104 | Fakultas Matematika dan Ilmu Pengetahuan Alam ( FMIPA ) |

Telah melakukan observasi di SMP 1 Kudus pada tanggal 1 Maret 2012 untuk memenuhi tugas mata kuliah Skripsi dengan Dosen Pembimbing Dr.Saiful Ridlo, M.Si & Dra.Ely Rudyatmi,M.Si  
 Demikian surat keterangan ini harap yang bersangkutan maklum dan dipergunakan sebagaimana mestinya.

Kudus, 2 Maret 2012  
 Kepala SMP 1 Kudus



H. Oki Sugarto, S.Pd  
 Pembina  
 NIP. 19571016 197803 1 002

**Lampiran 19**  
**Surat Permohonan Ijin Observasi Dinas Pendidikan Kota Yogyakarta**



KEMENTERIAN PENDIDIKAN DAN KEBUDAYAAN  
 UNIVERSITAS NEGERI SEMARANG  
 FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
 Gedung D5 Lt.1 Kampus Sekaran Gunungpati Semarang Kode Pos 50229, Telp. (024)8508112  
 Telp. Dekan (024)8508005; Jurusan: Matematika (024)8508032; Fisika (024)8508034; Kimia (024)8508035; Biologi (024)8508033  
 Fax. (024)8508005; Website: <http://mipa.unnes.ac.id>; Email: [mipa@unnes.ac.id](mailto:mipa@unnes.ac.id)

Nomor : 2358 /UN37.1.4/PP/2012  
 Lampiran : -  
 Hal : Permohonan Ijin Observasi

27 Maret 2012

Yth. Kepala Dinas Pendidikan Kota Yogyakarta  
 di Yogyakarta

Kami memberitahukan dengan hormat, bahwa mahasiswa Fakultas Matematika dan Ilmu Pengetahuan Alam (FMIPA) Universitas Negeri Semarang tersebut di bawah ini :

Nama : Mahanani  
 NIM : 4401408104  
 Semester : 8  
 Jurusan : Biologi

dalam rangka tugas mata kuliah Skripsi dengan dosen pembimbing Dr. Saiful Ridlo, M.Si & Dra. Ely Rudyatmi, M.Si bermaksud akan mengadakan observasi di:

Tempat : SMP Negeri 5 Yogyakarta  
 Waktu : bulan Maret – April 2012

Berkaitan dengan hal ini, kami mohon dapat diberikan ijin observasi kepada mahasiswa yang bersangkutan pada tempat dan jadwal waktu tersebut di atas.

Atas perhatian dan kerja sama Saudara, kami sampaikan terima kasih.



Prof. Dr. Wyanjo, M.Si  
 NIP. 196310121988031001

Tembusan :  
 1. Ketua Jurusan Biologi;  
 2. Dosen Pembimbing;  
 FMIPA Universitas Negeri Semarang.

**Lampiran 20**  
**Surat Permohonan Ijin Observasi SMP N 5 Yogyakarta**



KEMENTERIAN PENDIDIKAN DAN KEBUDAYAAN  
 UNIVERSITAS NEGERI SEMARANG  
 FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
 Gedung D5 Lt.1 Kampus Sekaran Gunungpati Semarang Kode Pos 50229, Telp. (024)8508112  
 Telp. Dekan (024)8508005; Jurusan: Matematika (024)8508032; Fisika (024)8508034; Kimia (024)8508035; Biologi (024)8508033  
 Fax. (024)8508005; Website: <http://fmipa.unnes.ac.id>; Email: [fmipa@unnes.ac.id](mailto:fmipa@unnes.ac.id)

Nomor : 343 /UN37.1.4/LT/2012  
 Lampiran : -  
 Hal : *Permohonan Ijin Observasi*

15 Mei 2012

Yth. Kepala SMP Negeri 5 Yogyakarta  
 di Yogyakarta

Kami memberitahukan dengan hormat, bahwa mahasiswa Fakultas Matematika dan Ilmu Pengetahuan Alam (FMIPA) Universitas Negeri Semarang tersebut di bawah ini :

Nama : Mahanani  
 NIM : 44014081104  
 Semester : 8  
 Jurusan : Pend. Biologi

dalam rangka tugas mata kuliah Skripsi dengan dosen pembimbing/pengampu :

1. Dr. Saiful Ridlo, M.Si
2. Dra. Ely Rudyatmi, M.Si

bermaksud akan mengadakan observasi pada :

Tempat : SMP N 5 Yogyakarta  
 Waktu : Mei 2012

Berkaitan dengan hal ini, kami mohon dapat diberikan ijin observasi kepada mahasiswa yang bersangkutan pada tempat dan jadwal waktu tersebut di atas.

Atas perhatian dan kerja sama Saudara, kami sampaikan terima kasih.



Prof. Dr. Wanto, M.Si  
 NIP. 19631012 198803 1 001

Tembusan :  
 1. Ketua Jurusan Biologi;  
 2. Dosen Pembimbing;  
 FMIPA Universitas Negeri Semarang.