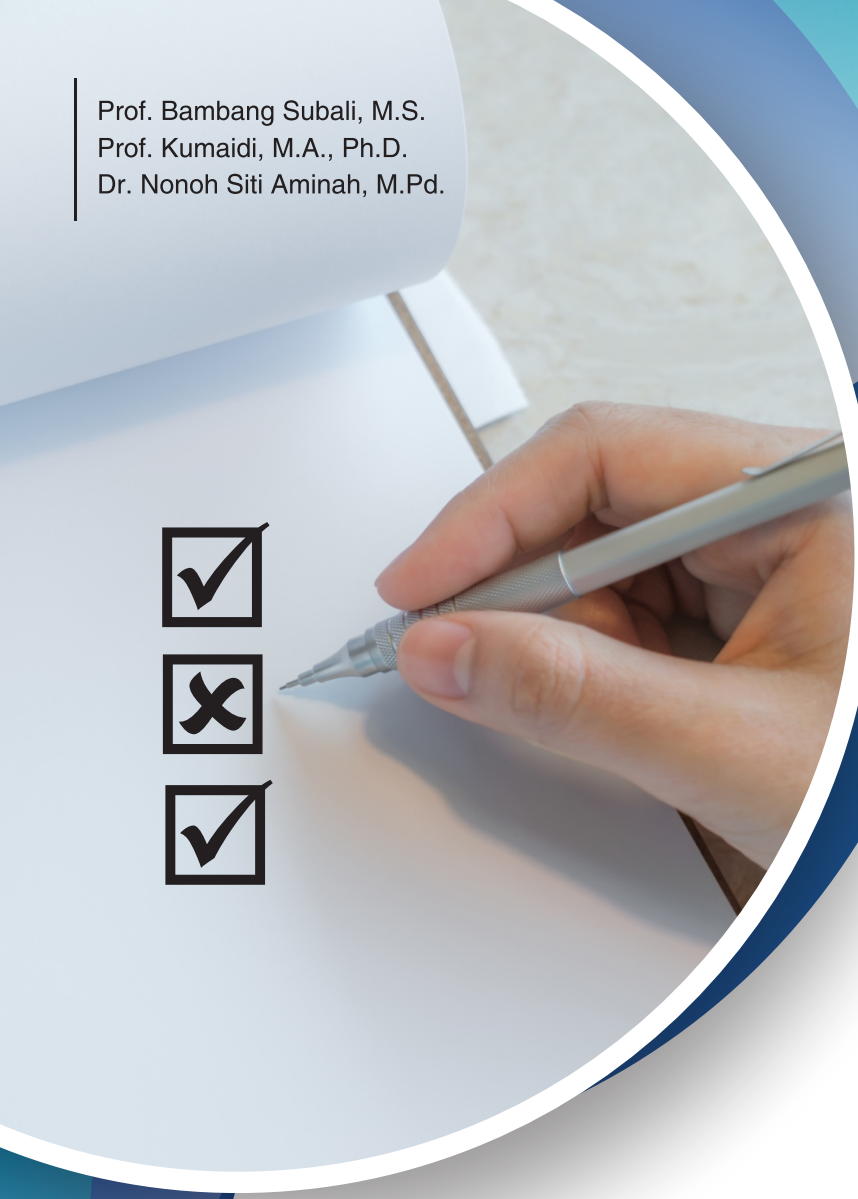
Prof. Bambang Subali, M.S.
Prof. Kumaidi, M.A., Ph.D.
Dr. Nonoh Siti Aminah, M.Pd.



Pengembangan Model Asesmen Konfirmatori Penguasaan Metode Ilmiah Pada Mata Pelajaran IPA di Sekolah Dasar

UNY PRESS
IKAPI IKATAN PENERBIT INDONESIA
Appti

PENGEMBANGAN MODEL ASESMEN KONFIRMATORI PENGUASAAN METODE ILMIAH PADA MATA PELAJARAN IPA DI SEKOLAH DASAR

Prof. Dr. Bambang Subali, M.S.
Prof. Kumaidi, M.A., Ph.D.
Dr. Nonoh Siti Aminah, M.Pd.

UNDANG-UNDANG REPUBLIK INDONESIA
NOMOR 28 TAHUN 2014
TENTANG HAK CIPTA

Pasal 2

Undang-Undang ini berlaku terhadap:

a. semua Ciptaan dan produk Hak Terkait warga negara, penduduk, dan badan hukum Indonesia;

b. semua Ciptaan dan produk Hak Terkait bukan warga negara Indonesia, bukan penduduk Indonesia, dan bukan badan hukum Indonesia yang untuk pertama kali dilakukan Pengumuman di Indonesia;

c. semua Ciptaan dan/atau produk Hak Terkait dan pengguna Ciptaan dan/atau produk Hak Terkait bukan warga negara Indonesia, bukan penduduk Indonesia, dan bukan badan hukum Indonesia dengan ketentuan:
   1. negaranya mempunyai perjanjian bilateral dengan negara Republik Indonesia mengenai pelindungan Hak Cipta dan Hak Terkait; atau
   2. negaranya dan negara Republik Indonesia merupakan pihak atau peserta dalam perjanjian multilateral yang sama mengenai pelindungan Hak Cipta dan Hak Terkait.

BAB XVII
KETENTUAN PIDANA

Pasal 112

Setiap Orang yang dengan tanpa hak melakukan perbuatan sebagaimana dimaksud dalam Pasal 7 ayat (3) dan/atau Pasal 52 untuk Penggunaan Secara Komersial, dipidana dengan pidana penjara paling lama 2 (dua) tahun dan/atau pidana denda paling banyak Rp300.000.000,00 (tiga ratus juta rupiah).

(1) Setiap Orang yang dengan tanpa hak melakukan pelanggaran hak ekonomi sebagaimana dimaksud dalam Pasal 9 ayat (1) huruf i untuk Penggunaan Secara Komersial dipidana dengan pidana penjara paling lama 1 (satu) tahun dan/atau pidana denda paling banyak Rp100.000.000 (seratus juta rupiah).

(2) Setiap Orang yang dengan tanpa hak dan/atau tanpa izin Pencipta atau pemegang Hak Cipta melakukan pelanggaran hak ekonomi Pencipta sebagaimana dimaksud dalam Pasal 9 ayat (1) huruf c, huruf d, huruf f, dan/atau huruf h untuk Penggunaan Secara Komersial dipidana dengan pidana penjara paling lama 3 (tiga) tahun dan/atau pidana denda paling banyak Rp500.000.000,00 (lima ratus juta rupiah).

(3) Setiap Orang yang dengan tanpa hak dan/atau tanpa izin Pencipta atau pemegang Hak Cipta melakukan pelanggaran hak ekonomi Pencipta sebagaimana dimaksud dalam Pasal 9 ayat (1) huruf a, huruf b, huruf e, dan/atau huruf g untuk Penggunaan Secara Komersial dipidana dengan pidana penjara paling lama 4 (empat) tahun dan/atau pidana denda paling banyak Rp1.000.000.000,00 (satu miliar rupiah).

(4) Setiap Orang yang memenuhi unsur sebagaimana dimaksud pada ayat (3) yang dilakukan dalam bentuk pembajakan, dipidana dengan pidana penjara paling lama 10 (sepuluh) tahun dan/atau pidana denda paling banyak Rp4.000.000.000,00 (empat miliar rupiah).

PENGEMBANGAN MODEL ASESMEN KONFIRMATORI PENGUASAAN METODE ILMIAH PADA MATA PELAJARAN IPA DI SEKOLAH DASAR

Prof. Dr. Bambang Subali, M.S.
Prof. Kumaidi, M.A., Ph.D.
Dr. Nonoh Siti Aminah, M.Pd.

UNY PRESS
2018

PENGEMBANGAN MODEL ASESMEN KONFIRMATORI PENGUASAAN METODE ILMIAH PADA MATA PELAJARAN IPA DI SEKOLAH DASAR

Oleh:
Prof. Dr. Bambang Subali, M.S.
Prof. Kumaidi, M.A., Ph.D.
Dr. Nonoh Siti Aminah, M.Pd.

ISBN: 978-602-498-019-1
Edisi Pertama, Desember 2018

Diterbitkan oleh:
UNY Press
Jl. Gejayan, Gg. Alamanda, Komplek Fakultas Teknik UNY
Kampus UNY Karangmalang Yogyakarta 55281
Telp: 0274 – 589346, Mail: unypress.yogyakarta@gmail.com
© 2018 Prof. Bambang Subali, M.S., dkk
Anggota Ikatan Penerbit Indonesia (IKAPI)
Anggota Asosiasi Penerbit Perguruan Tinggi Indonesia (APPTI)

Desain Sampul : Deni Satriya H.
Tata Letak : Pudji Triwibowo

Isi di luar tanggung jawab percetakan

Prof. Bambang Subali, M.S., dkk
PENGEMBANGAN MODEL ASESMEN KONFIRMATORI PENGUASAAN METODE ILMIAH PADA MATA PELAJARAN IPA DI SEKOLAH DASAR
-Ed.1, Cet.1.- Yogyakarta: UNY Press 2018
vi + 153 hlm; 15 x 23 cm
ISBN: 978-602-498-019-1

Kata Pengantar

Buku ini merupakan salah satu luaran kegiatan penelitian Hibah Kompetensi yang dilakukan penulis pada tahun 2016-2018. Salah satu luaran lainnya yang sudah dipublikasi dalam *Journal of TURKISH SCIENCE EDUCATION* Volume 15, Issue 2, June 2018 dengan judul *Developing a Scientific Learning Continuum of Natural Science Subjects at Grades 1 – 4* sementara dua artikel yang lain sedang dalam proses lanjut setelah disubmit awal tahun ini.

Penerbitan hasil penelitian dalam bentuk buku ini selain memenuhi persyaratan pola publikasi buku berbasis hasil penelitian menurut Direktorat Pendidikan Tinggi juga diharapkan dapat menjadi acuan bagi para peneliti dan pendidik yang berkecimpung dalam bidang Pengukuran dan Evaluasi Pendidikan Biologi pada khususnya, bidang Pendidikan Sains serta bidang Pendidikan pada umumnya.

Peneliti juga sedang melakukan penelitian lanjutan bersama mahasiswa Pascasarjana Universitas Negeri Yogyakarta untuk mengembangkan learning continuum untuk memperoleh pendapat para pakar pendidikan. Dalam hal ini dikhususkan untuk aspek biologi.

Kritik dan saran sangat penulis perlukan untuk penyempurnaan buku ini juga untuk penyiapan publikasi hasil penelitian lainnya di masa mendatang yang disajikan dalam bentuk buku.

Yogyakarta, 2018
Penulis

Daftar Isi

HALAMAN JUDUL i
KATA PENGANTAR v
DAFTAR ISI vi

BAB I
PENDAHULUAN 1

BAB II
PERAN GURU DAN PENGAWAS 4

BAB III
HAKEKAT SAINS DAN PERAN METODE ILMIAH 7

BAB IV
TES KETERAMPILAN TERTULIS SEBAGAI INSTRUMEN PENILAIAN KINERJA 13

BAB V
TES KONVERGEN DAN DIVERGEN 17

BAB VI
PENGEMBANGAN ITEM TES 19

BAB VII
ANALISIS KUALITATIF ITEM TES 28

BAB VIII
ANALISIS KUANTATIF ITEM TES 32

BAB IX
ANALISIS ITEM DAN KEGUNAANNYA 51

BAB X
ANALISIS ITEM MENGGUNAKAN PROGRAM QUEST 57

DAFTAR PUSTAKA 72
LAMPIRAN-LAMPIRAN 80

BAB I
PENDAHULUAN

Penguasaan metode ilmiah merupakah hal yang sangat penting dalam belajar sains. Metode ilmiah adalah proses yang ditempuh ilmuwan dalam melakukan penyelidikan untuk memperoleh produk ilmiah (Carin & sund, 1989). Menurut Brum & McKane (LeBoffe & Wisehart, 1989) metode ilmiah merupakan keterampilan proses sians yang tersusun secara sistematis untuk memecahkan masalah. Melalui proses ilmiah yang dilandasi sikap ilmiah para ilmuwan menemukan produk ilmiah berupa fakta, konsep, dan prinsip yang baru dalam sains (Carin & Sund, 1989).

Cara yang ilmiah untuk memecahkan masalah sains dilakukan dengan tahapan-tahapan (a) mengidentifikasi masalah, (b) mengumpulkan fakta dan informasi, (c) merancang penyelidikan/eksperimen, (d) melaporkan semua hasil yang diperoleh, (e) mengulang penyelidikan/eksperimen, (f) memeriksa hasilnya, (g)

menarik simpulan (Belardo, Samia (1999). Membelajarkan sains kepada peserta didi15 tahun harus sudah pada tingkatan melatih peserta didik untuk merancang dan melaksanakan investigasi, membuat data dan menginterpretasikannya (Watts, De Bentley, & Hornsby, 1989)

Menurut Bryze *et. al*. (1995) metode ilmiah terdiri atas sejumlah *basic skills* dan *process skills* yang apabila dirangkai menjadi langkah yang sistematis akan menjadi *investigative skills*. Sementara menurut Rezba (2007: 1) keterampilan proses sains berupa *basic skills* dan apabila dipadukan akan menjadi integrative skills dalam bentuk metode ilmiah.

Aspek basic skills menurut Bryze et al (1995) meliputi subaspek keterampilan (a) mengamati menggunakan pancaindera, (b) merekam data/informasi, (c) mengikuti instruksi, (d) menglasifikasi, (e) mengukur, (f) memanipulasi gerakan, (g) mengimplemetasikan prosedur dan menggunakan peralatan, dan (h) memprediksi. Keterampilan memroses meliputi subaspek keterampilan (a) menginferensi, dan (b) menyeleksi prosedur. Namun hasil penelitian Subali (2009) keterampilan memprediksi lebih tepat dimasukkan ke dalam aspek memroses. Keterampilan melakukan investigasi menurut Subali mencakup subaspek keterampilan (a) merancang investigasi (plan investigation), (b) melaksanakan investigasi (conduct investigation), dan (c) melaporkan hasil investigasi baik secara tertulis maupun secara lisan.

Metode ilmiah dalam pembelajaran sains perlu dikenalkan sedini mungkin kepada peserta didik. Namun demikian, mengingat metode ilmiah merupakan

serangkaian proses ilmiah yang bersifat komplek maka pembelajarannya haruslah sedikit demi sedikit, aspek demi aspek, bahkan subaspek demi subaspek. Dengan demikian, membelajarkan metode ilmiah kepada peserta didik tingkat sekolah dasar. Dalam hal ini, tidak lain adalah bagaimana membelajarkan peserta didik menguasai setiap aspek dan subaspek keterampilan proses sains secara parsial atau terpisah-pisah lebih dahulu. Hal ini sesuai dengan hasil penelitian Subali, Kumaidi, dan Aminah (2016) yang menurut guru perlu diajarkan walau baru pada pengenalan aspek demi aspek.

BAB II

PERAN GURU DAN PENGAWAS

Penguasaan metode ilmiah merupakan bagian dari kompetensi yang harus dikuasai peserta didik. Penilaian berbasis kompetensi berfokus pada hasil (*output*) dan bukan pada masukan atau proses. Artinya bahwa seseorang yang belajar harus menguasai kompetensi/standar yang ditetapkan. Oleh karena itu, penilaian hasil belajar juga diarahkan kepada penguasaan peserta didik atas kompetensi yang harus dikuasainya. Dengan kata lain, fokus penilaian bukan pada bagaimana cara mereka mencapai tingkat penguasaan atas kompetensi yang ditargetkan. Atas dasar itu, sangat perlu dilakukan penilaian awal untuk mengetahui seberapa jauh kompetensi yang akan dipelajari dikuasai oleh peserta didik sebelum pembelajaran. Data dapat dihimpun melalui tes awal. Namun demikian, tidak berarti bahwa dalam penilaian berbasis kompetensi memantau kemajuan hasil belajar selama proses

pembelajaran tidak penting. Proses tetap penting dalam rangka menunjang hasil.

Penilaian berbasis kompetensi dilakukan terhadap setiap peserta didik yang menempuh pembelajaran untuk menentukan penguasaan kompetensi tertentu. Dengan demikian, penilaian berbasis kompetensi tidak ditujukan untuk membandingkan status keberhasilan seseorang terhadap kelompoknya. Kegiatan penilaian dapat dilaksanakan dalam situasi kelompok, misal untuk menilai kompetensi berdiskusi, namun sasaran penilaian tetap pada kemampuan secara individual.

Penetapan status seseorang dalam penilaian berbasis kompetensi adalah menetapkan status seseorang atas keberhasilan yang dicapainya atau status seseorang sebelum ia belajar. Oleh karena itu, penilai harus memiliki rujukan (*reference*) yang jelas dan pasti. Kegiatan yang dilakukan penilai dalam menilai keberhasilan belajar peserta didik adalah membandingkan tingkat kompetensi yang telah dicapai dengan kompetensi yang telah ditetapkan sebagai rujukan. Sekali lagi, bukan membandingkan kompetensi seseorang dengan orang-orang lain di dalam kelompoknya.

Pengawas adalah pihak yang bertanggungjawab untuk memantau dan mengevaluasi apa yang dibelajarkan guru kepada peserta didik. Untuk keperluan tersebut, pengawas harus dapat mengetahui penguasaan peserta didik atas aspek-aspek metode ilmiah dalam sains. Untuk itu, pengawas harus mampu mengembangkan tes untuk memantau penguasaan aspek metode ilmiah oleh peserta didik, agar supaya pengawas dapat memantau penguasaan aspek metode ilmiah yang diajarkan guru kepada siswa

dengan menggunakan tes yang standar berupa tes konfirmatori.

BAB III

HAKEKAT SAINS DAN PERAN METODE ILMIAH

A. Hakekat Pembelajaran Sains

Proses pembelajaran dalam mata pelajaran Biologi yang sesuai dengan hakikat Biologi sebagai bagian/cabang IPA, yakni harus bertumpu pada proses ilmiah yang melibatkan berbagai keterampilan proses sains. Misalnya, keterampilan (a) melakukan pengamatan, (b) mengoleksi data, (c) melakukan pengukuran, dan (d) mengorganisasikan serta mengklasifikasi data (Towle, 1989: 16-31). Keterampilan melakukan pengamatan atau mencandra dilihat dari penjenjangannya menduduki posisi awal dalam melakukan proses sains (Rezba et. al., 1995: 1). Proses yang lebih tinggi berupa keterampilan seperti mengukur, mengklasifikasi. Keterampilan tertinggi berupa yaitu keterampilan bereksperimen.

Collette maupun Gega (Djohar, 1989: 10) membagi keterampilan melakukan proses sains menjadi dua macam, yakni keterampilan dasar dan keterampilan terintegrasi. Keterampilan dasar meliputi kegiatan observasi, klasifikasi,

pengukuran, komunikasi, inferensi, dan prediksi; sedangkan keterampilan terpadu meliputi kegiatan merumuskan hipotesis, mengontrol variabel, merumuskan masalah, dan menginterpretasi data. Sementara Bryce et. al. (1990: 2) membagi keterampilan proses sains mencakup keterampilan proses sains dasar (basic process science skill) yang terdiri dari keterampilan dasar (basic skill) dan keterampilan mengolah/memroses (process skill), sera keterampilan melakukan investigasi (investigation skill).

Sains sebagai proses tidak lain adalah metode ilmiah, yang terdiri atas: (a) pengamatan, (b) formulasi hipotesis yang dapat diuji secara induktif, (c) desain eksperimen secara deduktif lengkap dengan penetapan kelompok kontrol dan kelompok eksperimen, (d) eksperimentasi/pengujian hipotesis, (e) analisis hasil eksperimen, (f) menarik kesimpulan eksperimen, (g) menerima, menolak atau memodifikasi hipotesis, untuk dikembangkan menjadi teori dan hukum, dan (h) sharing hasil penelitian (Brum & McKane, 1989: 10; Hibbard, t.t.: 17-35). IPA atau sains sebagai proses dapat disejajarkan dengan metode ilmiah. Oleh karena itu, belajar IPA atau tidak dapat terlepas dari kemampuan untuk melakukan observasi, memformulasikan hipotesis yang testabel, kemampuan untuk melakukan induksi dan deduksi, merancang dan melaksanakan eksperimen untuk membuktikan hipotesis (Brum & McKane, 1989: 9-10).

B. Hubungan Kemampuan Berpikir Konvergen, Divergen, dan Kreatif

Proses berpikir melibatkan beberapa tahap dan dalam pola yang saling berganti atau saling melengkapi yakni: (a) antara proses deduktif dan proses induktif, (b) antara produk dan asosiasi, dan (c) antara berpikir konvergen dan berpikir

divergen (Garry, 1970: 473-475). Hudson (Atherton, 2005: 1) mendefinisikan kemampuan berpikir divergen sebagai kemampuan berpikir dari satu titik sebagai pusatnya menyebar ke berbagai arah. Berpikir divergen sebagai keterampilan untuk mengelaborasi gagasan secara kreatif. Berpikir divergen akan menghasilkan banyak gagasan yang berbeda tentang suatu topik di dalam suatu periode waktu tertentu. Berpikir divergen dan konvergen idealnya saling melengkapi. Sebaliknya, ciri berpikir konvergen terjadi secara sistematis, di mana seseorang mengorganisasi gagasan-gagasan dan informasi menuju satu pemusatan sehingga tersusunlah gagasan pokoknya. Dalam berpikir analitis kritis terjadi proses berpikir divergen dan konvergen yang terpadu (Conny Semiawan, 1997: 54-58).

Mengacu model struktur intelek menurut Guilford, Gorman (Moh Amien, 1980: 7) menyatakan bahwa kreativitas melibatkan proses berpikir divergen. Struktur intelek dari Guilford yang menggambarkan kemampuan intelektual seseorang, yakni terdiri dari jalinan operasi intelektual yang dilakukan (operation) dengan melibatkan konten/isi (contents), dan akan menghasilkan produk (product).

Wallas, 1926 (Garry, 1970: 475), menyatakan bahwa tahapan proses kreatif ada empat yakni: (a) preparasi, (b) inkubasi, (3) iluminasi, dan (4) verifikasi. Sementara Osborn, 1957 (Garry, 1970: 475) menyatakan ada 7 tahapan dalam proses kreatif, yakni (a) orientasi atau penetapan permasalahan, (b) preparasi dan pengumpulan data, (c) analisis, (d) menggagas atau mengenali alternatif-alternatif, (e) inkubasi dan eluminasi, (f) sintesis, dan (g) evaluasi. Gowan & Treffinger (Conny Semiawan, 1997: 63-64) menyatakan ada tiga tingkatan kreativitas sebagai kemampuan umum peserta

didik berbakat. Ketiga tingkatan tersebut dilihat, baik dari segi kognitif maupun dari aspek afektif.

Kemampuan pada diri manusia dalam taksonomi Bloom dipisahkan menjadi tiga domain, yakni domain (a) kognitif, (b) afektif, dan (c) psikomotor. Sekarang kemampuan manusia dalam taksonomi Bloom yang telah diperbaharui dipisahkan menjadi empat domain (Dettmer, 2006: 71-73). Empat domain tersebut adalah domain (a) kognitif, (b) afektif, (c) sensorimotor (sebagai pengganti psikomotor), dan (d) sosial. Keempat domain tersebut sebagai aktualisasi dalam pembelajaran membentuk satu kesatuan yang disebut dengan unity sehingga memiliki jalinan satu sama lain dalam kaitannya dengan aktivitas pembelajaran dan melakukan sesuatu hal. Kemampuan berkreasi merupakan bagian dari aspek kognitif.

Menurut Torrance (1979: 241-246) proses pemecahan masalah secara kreatif diawali dengan fase peningkatan antisipasi. Fase berikutnya ditandai adanya proses mempertemukan atau menandingkan dan menggali harapan-harapan yang diinginkan dan yang tidak diinginkan. Fase ini ditanda oleh adanya proses diagnostik di dalam otak dalam menghadapi kesulitan, dalam mengintegrasikan berbagai informasi yang tersedia, mengecek kembali informasi, mengelaborasi, dan dalam memilah informasi. Dengan demikian, terjadi proses konvergen dan divergen. Kemudian diakhiri dengan fase yang ditandai adanya kemampuan untuk melampaui hambatan yang ada.

Pembelajaran ideasional sebagaimana direkomendasikan oleh Dettmer (2006: 73) yang dasarnya adalah berbasis gagasan dari masing-masing peserta didik seharusnya dapat dirintis pada seluruh sekolah karena tidak selamanya bahwa anak yang kreatif adalah anak yang cerdas. Hasil penelitian Pollman (1973: 1) menunjukkan tidak ada korelasi yang kuat

antara skor subtes IQ model Lorge Thorndike dan skor tes kreativitas model Torrance yang diperoleh dari 154 peserta didik K-12 di AS. Hasil penelitian Ferrando et al. (2005: 21-50) menunjukkan adanya korelasi yang rendah antara kreativitas dan intelegensi. Peserta didik dengan IQ yang tinggi tidak semuanya lebih kreatif. Menurut Cromie (2007: 1) tidak semua studi menunjukkan adanya korelasi antara tingkatan IQ dan kreativitas. Beberapa studi menunjukkan bahwa peningkatan kreativitas sejalan dengan peningkatan IQ sampai dengan IQ setinggi 120. Kyung Hee Kim (2005: 1) melaporkan bahwa hasil metaanalisis 447 koefisien korelasi menunjukkan banyak skor tes krativitas yang tidak ada hubungannya dengan skor IQ, namun banyak pula yang menunjukkan hubungannya.

Kemampuan berpikir kritis mencakup tiga aspek, yakni: (a) mengidentifikasi hal penting yang sedang dibahas, (b) merekonstruksi argumen, dan (c) mengevaluasi argumen yang sudah direkonstruksi (Bowell & Kemp, 2002 : 6). Dalam belajar sains pada umumnya, keterampilan berpikir kritis menyangkut keterampilan berpikir untuk (a) mengklasifikasi, (b) membuat asumsi, (c) memprediksi dan berhipotesis, (d) menyimpulkan dan menginterpretasikan data serta menarik kesimpulan, (e) mengukur, (f) merancang penyelidikan untuk memecahkan suatu masalah, (g) mengamati atau mencandra, (h) membuat grafik, (i) mereduksi kesalahan eksperimen, (j) mengevaluasi, dan (k) menganalisis (Carin & Sund, 1989, 159-160). Dengan demikian berpikir kritis melibatkan berbagai aspek. Kemampuan berpikir kritis dalam pembelajaran IPA dapat dikembangkan. Pengembangan kemampuan berpikir kritis dapat dilakukan dengan membiasakan memberi pertanyaan divergen dalam setiap aspek yang berupa pertanyaan terbuka dan memiliki order berpikir yang tinggi. Penguasaan kemampuan berpikir divergen pada peserta didik

akan menjadikannya mampu mengambil keputusan sebagai bentuk berpikir konvergen (Collette & Chiappetta, 1994: 142-150).

C. Peran Metode Ilmiah

Metode ilmiah memiliki peran penting untuk menemukan pengetahuan ilmiah. Jika diurai maka menurut Anderson & Krathwohl (2001) pengetahuan iliah terdiri dari pengetahuan faktual, pengetahuan konseptual, dan pengetahuan prosedural. Bahkan berkait dengan aktivitas belajar maka ada pengetahuan metakogbitif yang dikuasai peserta didik sebelum, selama, dan sesudah belajar. Menurut Piaget (Huitt & Hummel, 2003) di dalam otak terdapat banyak sekali penggalan-penggalan pengetahuan yang jika ada stimulus dari luar maka penggalan-penggalan pengetahuan itu akan terangkai membentuk konsep-kosep baru. Dengan diperolehnya fakta baru, konsep baru, dan prosedur baru maka terjadilah proses kreatif pada diri peserta didik.

Metode ilmiah dalam sains dapat dibedakan menjadi metode observasi dan metode eksperimen. Melalui metode observasi para saintis menggali berbagai fenomena yang ada dialam untuk dijadikan fakta baru dan selanjutnya fakta-fakta tersebut digeneralisasi menjadi konsep baru, dan dari konsep-konsep baru dapat disusun menjadi prinsip baru.

Melalui metode eksperimen para saintis memanipulasi variabel atau variabel-variabel bebas untuk diselidiki efeknya terhadap variabel tergayut. Dengan adanya manipulasi maka ada perlakuan yang dikenakan pada objek. Misalnya, peneliti dapat memanipulasi dosis pupuk untuk menyelidiki efek dosis pupuk terhadap pertumbuhan tanaman. Peneliti dapat memanipulasi suhu untuk menyelidiki perubahan wujud benda.

BAB IV

TES KETERAMPILAN TERTULIS SEBAGAI INSTRUMEN PENILAIAN KINERJA

Tes penguasaan metode ilmiah dapat berupa tes tulis keterampilan (skills written test) dapat pula berupa tes performansi dalam bentuk uji petik kerja (work sample test). Penyusunan tes yang baik idealnya mengacu pada learning continuum (NEA, 2003) sehingga dapat identify the child's instructional level. Bagaimana cara menyusun tes tulis telah banyak disajikan oleh para penulis antara lain Millard (2012), Miller (2008), Popham (2005), Gronlund & Linn (1990), Gronlund (1998), Roid & Haladyna (1982). Penelitian berkait dengan bentuk tes antara lain dilakukan oleh Luo & Zhang (2011) dengan judul "Multiple-choice Item and Its Backwash Effect on Language Teaching in China" and Shete, Kausar, Lakhkar, & Khan (2015) on item analysis for evaluating multiple choice questions in Physiology examination.

Kualitas tes selain harus memenuhi validitas teoretik yaitu dari segi validitas konstruk dan validitas isi, juga dapat ditinjau dari segi validitas empiris. Dalam hal ini validitas dapat ditinajau berdasarkan teori tes klasik dan teori tes modern/Item Respons Theory (IRT). Penerapan teori tes modern untuk analisis item telah ditulis antara lain oleh Le (2013).

Menurut Le dengan mengacu pendapat Osteen (2010), dibandingkan dengan teori tes klasik atau *classical test theory* (CTT), teori respon butir atau *item response theory* (IRT) dianggap sebagai standar. Banyak program pengujian masih menerapkan CTT dalam desain dan penilaian terhadap hasil tes. Hal ini disebabkan beberapa keunggulan CTT atas IRT. Sebagai contoh, CTT menggambarkan hubungan antara skor sebenarnya atau skor laten dan skor yang diamati secara linier yang membuat model CTT mudah dipahami dan diterapkan pada banyak peneliti. CTT menawarkan ukuran sampel yang lebih kecil daripada IRT, prosedur matematika CTT jauh lebih sederhana daripada IRT, estimasi parameter pada CTT secara konseptual sangat sederhana dan memerlukan asumsi minimum, membuat model berguna dan dapat diterapkan secara luas, analisis tidak memerlukan penyelidikan/studi *goodness of fit* yang ketat seperti pada IRT. Namun, CTT memiliki beberapa kekurangan utama. Salah satunya adalah bahwa skor tes tergantung testi yang diuji. Artinya, peserta ujian dapat memperoleh nilai yang lebih tinggi pada tes yang lebih mudah dan skor yang lebih rendah pada tes yang lebih sulit, dan dengan demikian tidak ada nilai sebenarnya yang bisa diekstraksikan karena tidak ada informasi tentang kemampuan testi. Ini tidak

memungkinkan dasar untuk mencocokkan item uji dan tingkat kemampuan.

Melalui IRT, tingkat kemampuan anak dan tingkat kesulitan item dapat di plot dalam satu garis menggunakan scala logit. Dengan demikian, tingkat kesulitan item dapat dibandingkan dengan kemampuan testee. Sementara itu, CTT hanya menginformasikan tingkat kesulitan Item sehingga tidak dapat dibandingkan dengan kemampuan siswa (Wright, 1999; Wright & Masters, 1982).

Penelitian terhadap karakteristik item secara empiris telah banyak diteliti antara lain dilakukan oleh Stage (2003) berupa pengalaman dalam menggunakan CTT dan IRT yang dilakukan pada testi swedia. Petrillo, Cano, McLeod, & Coon (2015) tentang penggunaan teori pengukuran CTT, IRT, dan rasch untuk mengevaluasi hasil pengukuran kesabaran. Simedru, Naghiu, Roman, Miclean, & Borgovan (2016) tentang studi perbandingan antara teknik sampling secara klasik dan modern untuk mengidentifikasi kontaminan item tes entomologi. Thorpe & Favia (2012) pada analisis data menggunakan metodologi IRT untuk program dan aplikasi terpilih. Fan (1998) tentang perbandingan empiris statistik item/person testi dengan menggunakan IRT dan CTT. Zoghi & Valipour (2014) pada studi perbandingan CTT dan IRT untuk memperkirakan parameter item tes dalam tes linguistik.

Penelitian terhadap karakteristik item ditinjau berdasarkan karakteristik teori tes klasik dan modern kaitannya dengan metode ilmiah atau aspek keterampilan proses sains telah diteliti oleh *Andi Ulfa Tenri Pada, Badrun Kartowagiran, & Bambang Subali (2016).* A separation index

and fit items of creative thinking skills assessment. Penelitian Subali, Paidi, Mariyam (2016)

Dalam mengembangkan tes penguasaan metode ilmiah, persoalan yang dihadapi adalah sejauh mana tipe tes tulis keterampilan penguasaan metode ilmiah mempengaruhi karakteristik tipe tesnya ditinjau dari CTT dan IRT. Dalam hal ini, (a) apakah bentuk tes pilihan dengan dua pilihan berbeda tingkat kesulitannya dibandingkan dengan tes pilihan dengan tiga pilihan jika ditinjau dari CTT dan IRT, (b) apakah tingkat kesulitan dari hasil analisis menggunakan CTT sama dengan hasil analisis IRT, dan (c) apakah item tipe tes BS dengan yang dibuat berkebalikan kunci jawabannya tidak berbeda karakteristinya jika ditinjau dari CTT dan IRT (jika pada tes BS pola A suatu pernyataan kalimatnya berkategori benar maka pada tes BS pola B kalimat tersebut diubah menjadi kategori yang salah).

BAB V
TES KONVERGEN DAN DIVERGEN

Berkait dengan penilaian pencapaian hasil belajar, Gronlund (1998) mengklasifikasikan menjadi 4 teknik tes, yakni teknik penilaian yang menuntut peserta didik memilih (*selected response test*), menuangkan jawaban (*supply response test*), mendemonstrasikan kinerjanya (*performance*) secara terbatas atau dengan aturan yang tidak boleh dilanggar (*restricted performance assessment*) dan mendemonstrasikannya secara bebas (*extended performance assessment*). Tes dalam bentuk memilih jawaban baik berupa tes pilihan ganda, benar-salah, dan menjodohkan mengukur kemampuan berpikir yang konvergen. Demikian pula tes yang meminta siswa menuliskan jawaban, apabila bentuknya berupa tes isian singkat dan tes uraian objektif, masih mengukur kemampuan berpikir yang konvergen.

Tes dalam bentuk penuangan jawaban (*supply response test*) terutama bentuk uraian non-objektif berpeluang memberi kesempatan kepada peserta didik

untuk secara bebas atau terbuka mengekspresikan pendapatnya. Demikian pula, bentuk tes kinerja/tes praktik yang bebas. Dengan kata lain, peserta didik diminta untuk berpikir tidak menuju ke satu titik melainkan berpikir dari satu titik menyebar ke berbagai arah sehingga mengekspresikan kemampuan berpikir dengan pola divergen. Menurut Gronlund ada perbedaan tingkat kemampuan yang dinilai di antara keempat metode penilaian tersebut.

Tabel 3. Perbandingan metode-metode yang digunakan dalam penilaian

Testing		*Performance assessment*	
Selected response	*Supply response*	*Restricted performance*	*Extended performance*
Realism of tasks			
Low ←			→ *High*
Complexity of tasks			
Low ←			→ *High*
Assessment time needed			
Low ←			→ *High*
Judgment in scoring			
Low ←			→ *High*

BAB VI
PENGEMBANGAN ITEM TES

Tes merupakan salah satu bentuk instrumen dalam penilaian kognitif yang telah dikenal. Penyusunan instrumen tes melalui prosedur yang berurutan. Prosedur penyusunan instrumen tes untuk tes hasil belajar mencakup tahapan (1) menentukan standar/kompetensi, (2) menentukan indikator pencapaian Kompetensi, (3) menyusun kisi-kisi tes, (4) menulis item tes berdasarkan kisi-kisi, (5) melakukan telaah item tes, (6) mengujikan soal pada testi, (7) analisis hasil tes, dan (8) merevisi soal tes.

Pengembangan tes yang berkait dengan implementasi kurikulum harus memperhatikan standar/kompetensi yang ditargetkan juga kompetensi prasyarat yang harus dikuasai jika diperlukan kompetensi prasyarat untuk menguasai kompetensi yang ditargetkan. Untuk pengembangan tes konfirmatori maka pengembang harus menyusun *learning continuum* terlebih dahul sebagai acuan penyusunan kisi-kisi tes.

a. Penyusunan *Learning Continuum* dan Kisi-kisi Tes

Kisi-kisi tes disusun dahulu baru menulis tes, bukan sebaliknya. Pada kisi-kisi ada (1) kompetensi, (2) materi pedagogik spesifik atau bahan ajar, (3) indikator pencapaian kompetensi, (4) teknik dan bentuk penilaian, dan (5) nomor item. Kisi-kisi perlu (1) mewakili isi kurikulum secara tepat, (2) komponennya jelas dan mudah dipahami, (3) dapat ditulis item tesnya. Kisi-kisi tes untuk pengukuran hasil belajar harus mengacu kepada *learning continuum. Learning continuum* tersebut merupakan garis yang berupa *abstract continuum*, karena mengukur berarti meletakkan posisi subjek yang diukur pada garis tersebut (Keeves & Alagumalai, 1999).

Pengembangan *learning continuum* yang pertama harus memperhatikan aspek kemampuan kognitif yang ditargetkan mulai dari kemampuan yang paling sederhana/mudah sampai yang paling sukar. Jika mengacu kepada jenjang/level domain kognitif maka kemampuan mengingat akan lebih mudah/sederhana dibanding kemampuan memahami, dan kemampuan memahami akan lebih mudah daripada kemampuan menerapkan. Kedua, harus memperhatikan kompleksitas bahan ajar (*subject matter skesific pedagogic*). Kompleksitas bahan ajar tergantung kepada keluasan dan kedalaman suatu bagian ilmu yang dijadikan materi sebagai alat untuk menguasai kompetensi yang ditargetkan. Agar peserta didik mengenal bentuk luar tubuh makhluk hidup cukup digunakan bahan ajar berupa "Pengenalan bentuk luar tubuh makhluk hidup". Agar memahami deskripsi bentuk luar tubuh makhluk hidup, peserta didik memerlukan bahan ajar berupa "Deskripsi

bentuk luar tubuh makhluk hidup". Agar peserta didik menemukan ciri-ciri khusus bentuk luar tubuh suatu makhluk hidup maka yang peserta didik memerlukan bahan ajar berupa "Perancangan kegiatan penemuan ciri-ciri khusus suatu makhluk hidup". Dengan demikian, bahan ajar akan menjadi spesifik sejalan dengan jenjang/level kompetensi yang harus dikuasainya.

Bahan ajar juga harus memperhatikan keluasan dan kedalaman keilmuan sesuai dengan karakteristik keilmuannya. Setiap bidang ilmu tentulah ada objek keilmuan yang dipelajari, karakteristik susunan organisasi/spesifikasi tingkatan kompleksitas objek, dan permasalahan/temanya yang dikaji.

Sebagai contoh, Biologi memiliki objek berupa seluruh makhluk hidup yang ada di alam termasuk yang pernah ada yang sekarang sudah musnah. Objek Biologi juga dibedakan dari objek/makhluk hidup yang kasat mata karena berukuran besar seperti tumbuhan pohon, perdu sampai semak; tumbuhan berbiji, tumbuhan paku, sampai tumbuhan lumut; sampai dengan makhluk hidup yang tidak kasat mata yang harus diamati menggunakan mikroskop seperti ganggang, jamur mikroskopis, protista, monera, dan bakteri. Bahkan untuk dapat mengamati tubuh virus diperlukan mikroskop elektron. Oleh karena itu, akan lebih mudah bila mengenal bentuk tubuh luar tumbuhan dan hewan besar daripada mengenal bentuk tubuh luar protista, monera, jamur mikroskopis, dan bakteri yang merupakan makhluk hidup satu sel.

Lokasi keberadaan makhluk hidup juga akan menjadi faktor menentukan tingkat kesulitan dalam

mengenal tubuh luar suatu makhluk hidup. Makhluk hidup yang ada di sekitar peserta didik akan mudah dikenali dibanding makhluk hidup yang di laut dalam atau di kawasan kutub.

Biologi sebagai ilmu juga dibedakan berdasarkan tingkat organisasinya. Tubuh makhluk hidup multiseluler (yang tersusun atas banyak sel) dispesifikasi atas dasar susunan molekulnya, sel-sel penyusunnya, juga jaringan dan organ tubuh yang dimilikinya. Makhluk hidup juga dispesifikasi oleh karakteristik populasinya, karakteristik komunitasnya (sehingga dikenal adanya ekosistem darat dan ekosistem perairan) juga sampai pada tingkatan biomanya. Biologi juga dibedakan atas dasar permasalahan/tema yang dikaji. Ada permasalahan/tema yang berkitan dengan evolusinya, perilaku/etologinya, kesamaan dan perbedaannya, struktur-fungsinya, materi dan energinya, serta lingkungan hidupnya. Bahkan biologi juga dispesifikasi dalam hal penemuan pengetahuan faktual, pengetahuan konseptual, dan pengetahuan proseduralnya, yang secara garis besar atas dasar metode penemuannya dibedakan antara metode observasi/deskriptif dan metode eksperimen.

Memang tidak semua ilmu pengetahuan dapat dipisahkan secara jelas objek, tingkat organisasi objek, dan permasalahan/temanya. Boleh jadi suatu ilmu pengetahuan hanya dapat dicirikan atas dasar objek dan permasalahan/temanya. Namun, ketika suatu ilmu pengetahuan dapat diketahui struktur kelimuannya maka secara otomatis akan dapat dirancang bagaimana tahapan langkah untuk mempelajarinya ketika telah disusun di

dalam *learning continuum*. *Learning continuum* bukan saja diperlukan untuk mengembangkan kurikulum sekolah, namun juga untuk membantu guru dalam merancang pembelajaran di suatu jenjang pendidikan/sekolah untum kelas tertentu. *Learning continuum* akan memandu pengembang kurikulum untuk menata mana kompetensi yang harus dikuasai peserta didik kelas 1, kelas, 2 sampai kelas 6 SD, kelas 7 sampai kelas 9 SMP, dan kelas 10 sampai kelas 12 SMA. Bahkan untuk ilmu yang bernomenklatur yang sama harus dapat dipisahkan mana kompetensi yang akan dikembangkan di S1, S2, dan S3. Hal ini hanya dapat dilakukan maknakala rumusan *leraning continuum* sudah tersedia.

Kurikulum berbasis kompetensi menuntut terkuasainya kompetensi sebagai kemampuan yang didemosntrasikan. Oleh karena itu, kunci utama pengembangan kisi-kisi tes kognitif adalah perumusan indikator. Indikator merupakan ciri spesifik yang dapat diamati sebagai petunjuk bahwa kompetensi dasar telah tercapai. Indikator yang baik adalah : (1) memuat ciri-ciri kompetensi yang hendak diukur, (2) mengandung kata kerja operasional, (3) terkait dengan isi/materi, dan (4) dapat ditulis item tesnya.

b. Penulisan Item Tes

Ada beragam bentuk tes, namun hasil uji coba lapangan dengan menerapkan tes bentuk pilihan ganda dan bentuk benar-salah ternyata bentuk benar-salah memiliki nilai indek reliabilitas yang lerbih rendah. Oleh karena itu dalam tes konfirmatori hanya disarankan menggunakan tes

bentuk pilihan ganda, atau bentuk pilihan ganda disertai alasan.

Item tes ditulis berdasarkan kisi-kisi. Ada rambu-rambu tertentu dalam penulisan tes objektif. Kualitas butir soal sangat ditentukan, antara lain oleh (1) kejelasan rumusan pertanyaan, yakni tidak meragukan dan komunikatif, (2) meyakinkan bahwa hanya ada satu jawaban benar, (3) tidak menggunakan pernyataan/kalimat negatif ganda. (4) pilihan jawaban homogen dan logis, (5) pengecoh (*distracter*) jangan memberikan petunjuk jawaban benar, dan (6) menggunakan bahasa Indonesia yang baik dan benar.

Untuk soal uraian, rambu-rambu penulisan butir soal adalah sebagai berikut.

1) Soal ditulis secara spesifik dan dapat ditangkap jelas oleh siswa. Soal seperti : *"Apa yang kau ketahui tentang bacaan di atas?"* adalah soal yang kurang spesifik. Soal tersebut terbuka dan akan menghasilkan jawaban yang sangat bervariasi, sehingga menyulitkan penskoran.
2) Pertanyaan diawali dengan kata: *bandingkan, berikan alasan, jelaskan, uraikan, mengapa, tafsirkan, hitunglah, simpulkan, buktikan*, dan semacam itu.
3) Menghindari awal pertanyaan dengan kata : siapa, apa, kapan, atau bila, sebab kata-kata tersebut hanya akan memancing jawaban yang berupa reproduksi informasi belaka.
4) Beberapa butir soal dengan jawaban pendek lebih baik daripada satu soal dengan jawaban panjang. Hal ini terkait dengan reliabilitas tes, makin banyak jumlah butir, makin tinggi koefisien reliabilitas tes tersebut.

Berikut disajikan persyaratan baik dari aspek materi/substansi, konstruksi, dan aspek bahasa untuk tes pilihan ganda.

Syarat item pilihan ganda (*multiple choice*)

Item bentuk pilihan yang berupa bentuk pilihan ganda tersusun atas butir soal dan jawaban kunci. Butir soal terdiri atas pernyataan yang belum lengkap disebut pokok soal (*stem*) yang untuk melengkapinya dengan cara memilihnya di antara beberapa alternatif/pilihan yang telah disediakan. Alternatif jawaban yang benar disebut kunci, dan yang salah atau yang kurang benar disebut pengecoh (*distracter*). Selain berupa pernyataan yang belum lengkap, pokok soal dapat berupa suatu pertanyaan. Ada yang hanya berupa pertanyaan dan ada yang diawali dengan uraian kasus, tabel, grafik, atau gambar. Jika item tes bertujuan untuk seleksi, harus ada yang terseleksi masuk kelompok atas dan kelompok bawah. Oleh karena itu, pengecoh harus berfungsi. Setiap pengecoh dipersyaratkan dapat dipilih minimal oleh 5% testi. Jika suatu item memiliki lima alternatif/pilihan jawaban maka yang terkecoh diharapkan minimal 20%. Diharapkan pula yang terkecoh adalah testi/peserta ujian yang berasal dari kelompok bawah. Jika untuk tujuan mengukur prestasi atau pencapaian hasil belajar maka pengecoh tidak akan ada yang memilih jika seluruh testi/peserta ujian benar-benar telah menguasai kompetensinya. Dengan demikian, semuanya akan memilih jawaban kunci. Sebaliknya jika testi/peserta ujian benar-benar belum menguasai kompetensi maka tidak akan ada yang memilih jawaban kunci. Beberapa hal yang

perlu diperhatikan dalam menyusun item bentuk pilihan ganda dari aspek materi/substansi, konstruksi, dan bahasa adalah sebagai berikut.

a) **Aspek materi/substansi**
   (1) item sesuai indikator,
   (2) pernyataan/pertanyaan dan jawaban kunci terumuskan dengan benar,
   (3) materi/substansi yang ditanyakan sesuai dengan tujuan pengukuran (untuk tujuan pengukuran hasil belajar, tujuan pengukuran untuk seleksi, atau tujuan pengukuran untuk konfirmatori/mengukur status), serta
   (4) materi/substansi yang ditanyakan sesuai dengan jenjang, jenis sekolah dan tingkatan kelas.

b) **Aspek konstruksi**
   (1) pokok soal (*stem)* dirumuskan dengan jelas,
   (2) pokok soal dan pilihan dirumuskan dengan tegas,
   (3) pokok soal tidak memberi petunjuk/mengarah kepada pilihan jawaban yang benar,
   (4) pokok soal tidak mengandung pernyataan negatif ganda,
   (5) jika terpaksa menggunakan kata negatif, harus digarisbawahi atau dicetak lain,
   (6) pilihan jawaban homogen,
   (7) hindari adanya alternatif jawaban: "seluruh jawaban di atas benar" atau "tak satu jawaban di atas yang benar" dan yang sejenisnya,
   (8) panjang alternatif /pilihan jawaban relatif sama, jangan ada yang sangat panjang dan ada yang sangat pendek,

(9) pilihan jawaban dalam bentuk angka/waktu diurutkan,
(10) uraian kasus/wacana, gambar, tabel atau grafik benar-benar berfungsi,
(11) hanya ada satu jawaban kunci yang benar, dan
(12) antaritem tidak bergantung satu sama lain.

c) Aspek bahasa

1) rumusan kalimat soal komunikatif,
2) kalimat menggunakan bahasa yang baik dan benar, sesuai dengan jenis bahasanya,
3) rumusan kalimat tidak menimbulkan penafsiran ganda atau salah pengertian,
4) menggunakan bahasa/kata yang umum (bukan bahasa *lokal* atau bahasa serapan baru yang belum dikenal oleh seluruh testi), dan
5) rumusan soal tidak mengandung kata-kata yang dapat menyinggung perasaan testi.

BAB VII

ANALISIS KUALITATIF ITEM TES

A. Pendahuluan

Salah satu tantangan yang dihadapi oleh peneliti adalah bagaimana memenuhi validitas tes. Pemenuhan validitas tes dapat dipenuhi manakala semua item penyusunnya valid. Salah satu cara untuk menyelidiki kevalidan item tes yaitu dengan melakukan telaah item tes atau disebut analisis item secara kualitatif, sering disebut pula analisis secara teoretik.

Penyelidikan validitas dapat saja dilakukan hanya dengan cara melihat tampang apakah item-item yang disusun untuk suatu bidang tertentu. Misalnya, karena soal-soalnya menanyakan hal-hal yang terkait dengan biologi dikatakan item-item tesnya merupakan item tes biologi. Namun, validitas tampang atau validitas muka (*face validity*) yang dimaksudkan dalam buku ini adalah validitas item tes dengan cara memeriksa kesesuaian item dengan kisi-kisi yang diacunya dengan memperhatikan substansi/isi materi,

konstruksi, dan bahasa. Telaah item tes dilakukan oleh (1) bukan oleh penulis item tes, dan (2) dilakukan oleh pakar yang menguasai isi/materi yang diujikan. Dari kegiatan tersebut, penulis item tes dapat mengetahui validitas muka atau validitas tampang (*face validity*) item tes yang diharapkan benar-benar memenuhi persyaratan item tes menurut penilaian pakar.

B. Teknik Analisis Kualitatif / Telaah Item Tes

Dalam melakukan analisis kualitatif, setiap item ditelaah sesuai dengan persyaratan item yang bersangkutan. Berikut disajikan tabel analisis kualitatif sesuai dengan bentuk tes. Cara penggunaannya, penelaah memberi tanda silang (X) jika suatu persyaratan dari suatu item dapat terpenuhi.

Telaah item tes bentuk pilihan ganda (*multiple choice*)

No	Aspek	Nomor item				
a.	**Aspek materi/substansi**	1	2	3	4	dst
1.	Item sesuai dengan indicator					
2.	Pernyataan/pertanyaan dan jawaban kunci terumuskan dengan benar					
3.	Materi/substansi yang ditanyakan sesuai dengan tujuan pengukuran (untuk tujuan pengukuran hasil belajar, tujuan pengukuran untuk seleksi, atau tujuan pengukuran untuk konfirmatori/mengukur status),					
4.	Materi/substansi yang ditanyakan sesuai dengan jenjang, jenis sekolah dan tingkatan kelas					
b.	**Aspek konstruksi**					
1.	Pokok soal (*stem)* dirumuskan dengan jelas					
2.	Pokok soal dan pilihan dirumuskan dengan tegas					
3.	Pokok soal tidak memberi petunjuk/mengarah kepada pilihan jawaban yang benar					
4.	Pokok soal tidak mengandung pernyataan negatif ganda jika terpaksa menggunakan kata negatif, harus digarisbawahi atau dicetak lain					
5.	Pilihan jawaban homogeny					
6.	Hindari adanya alternatif jawaban: "seluruh jawaban di atas benar" atau "tak satu jawaban di atas yang benar" dan yang sejenisnya					

7.	Panjang alternatif /pilihan jawaban relatif sama (jangan ada yang sangat panjang dan ada yang sangat pendek)					
8.	Pilihan jawaban dalam bentuk angka/waktu diurutkan					
9.	Uraian kasus/wacana, gambar, tabel atau grafik benar-benar berfungsi					
10.	Hanya ada satu jawaban kunci yang benar					
11.	Antaritem tidak bergantung satu sama lain					
c.	**Aspek Bahasa**					
1.	Rumusan kalimat soal komunikatif					
2.	Kalimat menggunakan bahasa yang baik dan benar, sesuai dengan jenis bahasanya					
3.	Rumusan kalimat tidak menimbulkan penafsiran ganda atau salah pengertian					
4.	Menggunakan bahasa/kata yang umum (bukan bahasa lokal atau bahasa serapan baru yang belum dikenal oleh seluruh testi)					
5.	Rumusan soal tidak mengandung kata-kata yang dapat menyinggung perasaan testi					

BAB VIII

ANALISIS KUANTITATIF ITEM TES

B. Analisis kuantitatif

Setelah dilakukan telaah item tes, item yang kurang/tidak memenuhi persyaratan direvisi oleh penulis item tes yang bersangkutan. Selanjutnya, item dirangkai menjadi seperangkat tes yang siap diujikan pada siswa. Selama pengujian, perlu diciptakan suasana ujian yang kondusif/optimal agar hasil yang diperoleh benar-benar dapat dipercaya.

Data yang masuk kemudian dianalisis untuk memperoleh kejelasan tentang kualitas tes yang disusun. Analisis dilakukan baik terhadap tes maupun terhadap item tes. Analisis tes untuk mengetahui reliabilitas tes, sedangkan analisis item tes untuk mengetahui kualitas setiap item tes yang bersangkutan. Analisis dapat dilakukan dengan metode yang hanya berupa pengukuran klasik, seperti program komputer ITEMAN, dapat pula menggunakan

kombinasi pengukuran klasik dan modern (*Item Respons Theory*) dengan program komputer Quest.

Agar hasil pengukuran baik, tes yang digunakan harus baik pula. Tes yang berkualitas tinggi apabila tes itu mampu mengukur secara tepat dan hasil pengukurannya dapat diandalkan. Dengan kata lain, tes perlu memenuhi syarat validitas dan reliabilitas. Suatu tes mempunyai validitas tinggi apabila tes tersebut dapat menjalankan fungsi ukurnya. Sisi lain dari konsep validitas yang penting adalah kecermatan pengukuran, yakni kemampuan untuk mendeteksi perbedaan-perbedaan kecil yang ada pada atribut yang diukurnya.

1. Validitas Tes

Suatu alat ukur dinyatakan sahih (valid), jika alat ukur tersebut benar-benar mampu memberikan informasi empirik sesuai dengan apa yang diukur. Suatu neraca pengukur massa dinyatakan sahih jika suatu benda yang memiliki massa 1 kg akan terukur 1 kg pula karena neraca tersebut memiliki akurasi/ keakuratan/ketelitian yang tinggi.

Hakekat kesahihan dalam ilmu sosial selain terkait dengan kelayakan suatu alat ukur juga berkait dengan masalah derajat. Jadi bukan masalah sahih dan tidak sahih. Kesahihan selalu bersifat spesifik kaitannya dengan tujuan dan interpretasi. Kesahihan juga mencerminkan ketunggalan (unidimensionalitas) konsep yang akan diukur. Selain itu alat ukur juga harus memiliki sifat andal (*reliable*). Artinya jika dipakai untuk mengukur secara berulang-ulang selalu tetap/konsisten/stabil hasilnya. Jadi, jika

penimbangan pertama 50 kg, penimbangan kedua juga menunjukkan berat 50 kg, demikian pula pada ulangan yang selanjutnya. Alat ukur yang andal memiliki presisi/ketepatan/kecermatan yang tinggi.

Suatu alat ukur dapat saja memiliki keandalan yang tinggi namun tidak sahih. Misalnya, suatu benda yang memiliki massa 50 kg dan setelah diukur massanya menggunakan suatu alat pengukur massa secara berulang-ulang hasilnya selalu menunjukkan bahwa massa benda tersebut 49 kg. Artinya bahwa alat pengukur massa tersebut andal karena dapat mengukur secara ajeg/konsisten yaitu sebesar 49 kg tetapi tidak sahih/akurat karena tidak menunjukkan besar massa benda tersebut yang sesungguhnya. Kecuali itu, hakekat keandalan suatu alat ukur adalah mengacu kepada hasil yang diperoleh bukan alat ukurnya itu sendiri. Keandalan alat ukur juga tidak berlaku secara umum. Ganti waktu atau ganti sampel mungkin sudah dapat berubah keandalannya. Suatu alat ukur yang andal tidak otomatis sahih dan keandalan suatu alat ukur juga sangat statis.

Menurut Gronlund (1985) hakekat validitas tes dapat ditinjau dari beberapa aspek sebagai berikut.

a. Validitas tes mengacu kepada kepantasan/ketercukupan (*appropriateness*) interpretasi hasil pengukuran yang dikenakan pada kelompok atau individu testi/peserta ujian terhadap instrumen tes yang digunakan. Berbicara tentang validitas adalah berbicara tentang interpretasi yang dapat dibuat dari hasil pengukuran. Semakin valid suatu instrumen tes semakin tepat interpretasi hasil pengukuran yang diperoleh.

b. Validitas adalah sesuatu hal yang memiliki derajat sehingga dapat dibedakan antara yang benar-benar sahih (*valid*) dan yang benar-benar tidak sahih (*invalid*) atau antara yang rendah, sedang, dan tinggi validitasnya.
c. Validitas adalah mengacu kepada suatu tujuan yang spesifik. Misalnya, suatu tes aritmatik dinyatakan memiliki derajat validitas yang tinggi jika dapat menunjukkan keterampilan dalam komputasi, dinyatakan memiliki derajat validitas yang rendah jika hanya untuk menunjukkan kemampuan berpikir aritmatik, dan memiliki derajat validitas yang sedang jika dapat digunakan untuk memprediksi keberhasilan belajar aritmatik ke depan. Akan tetapi, interpretasi tersebut tidak mampu untuk menilai atau menggambarkan validitas guna mempertimbangkan penggunaan hasil tes yang diperoleh.
d. Validitas adalah konsep kesatuan (*unitary*). Hakekat konsepsual validitas tes untuk bidang pendidikan dan psikologi kini didasarkan pada tiga aspek yang ditinjau berdasarkan bukti-bukti empiris yaitu aspek isi (*content*), aspek hubungannya dengan kriteria (*criterion related*) dan aspek konstruk (*construct*).

2. Pendekatan Untuk Menguji Validitas

a. Bukti yang berkaitan dengan isi

Hal yang menjadi pertanyaan pokok yang berkait dengan kesahihan isi (*content validity*) adalah "apakah sampel gugus alat ukur yang digunakan sudah representatif" atau bila berupa tes "seberapa baguskah

sampel tes yang digunakan untuk mengukur gugus tugas merepresentasikan ranah (*domain*) gugus tugas yang akan diukur". Dengan kata lain, secara sederhana dapat dipertanyakan adalah apakah semua item dalam suatu perangkat tes yang diujikan sudah mampu mencerminkan semua komponen ranah kemampuan tertentu dari subjek yang diukur.

Karena tujuan pengukuran untuk mengetahui sejauh mana kemampuan seseorang ditinjau dari segi domain ranah, tinjauan representasi dapat diartikan dari sisi kelengkapan komponen ranah yang akan diukur. Misal, jika yang akan diukur kemampuan kognitif peserta didik dan berdasarkan teori-teori yang ada, peserta didik yang bersangkutan harus sudah memiliki kemampuan kognitif mulai dari mengingat (*to know*) sampai dengan berkreasi (*to create*) pada taksonomi Bloom yang terbaru, item yang digunakan dinyatakan terpenuhi kesahihan isinya alat uji yang digunakan meliputi item jenjang dari mengingat, memahami, mengaplikasikan, menganalisis, mengevaluasi, menyintesis, berimajinasi, dan berkreasi. Oleh karena dalam ranah kognitif antar tingkatan berpikir sudah berbeda dimensinya, pemenuhan validitas dapat dibuat rentang kemampuan sesuai dengan tingkatan berpikir dan cakupan konsep keilmuan yang harus dikuasai. Dengan demikian, jika pengukuran itu dikaitkan dengan hasil belajar, maka akan dapat dibuat rentang kemampuan berdasarkan tingkatan berpikir dan keluasan serta kedalaman aspek keilmuannya dan akan menjadi suatu garis abstrak yang menunjukkan suatu kontinum kemampuan dari yang terendah sampai yang tertinggi yang akan menunjukkan

capaian hasil belajar. Itulah yang kemudian disebut dengan *learning continuum*.

Dalam proses pembelajaran, tujuan yang ditargetkan sudah dirumuskan dalam bentuk kompetensi. Oleh karena itu, persoalan validitas isi bergeser menjadi "Apakah item instrumen yang digunakan sudah mencerminkan seluruh kompetensi yang ditargetkan?". Jika seluruh kompetensi tidak dapat diukur, pertanyaannya adalah "Apakah item instrumen yang digunakan sudah mewakili seluruh kompetensi yang akan diukur?". Jadi, dalam hal ini berbicara tentang representasi butir soal. Karena tinjauannya dikembalikan kepada kompetensi yang ditargetkan, maka kesahihan isi juga disebut dengan kesahihan kurikuler (*Curricular validity*). Pertanyaan yang menarik yaitu apakah kompetensi dalam target pembelajaran hanya spesifik ranah kognitif. Jika hanya ranah kognitif, bagaimana dengan tujuan pendidikan yang diharapkan mencakup seluruh ranah/domain. Bahkan, sebagai hasil pembelajaran idealnya menjadi satu kesatuan/*unity*.

b. Bukti yang berkaitan dengan konstruk atau *construct validity*

Persoalan pokok yang dikaji dalam rangka pemenuhan validitas konstruk adalah pertanyaan mengenai "Apakah hasil pengukuran benar-benar bermakna atau benar-benar dapat diinterpretasikan. Dengan kata lain, apakah alat ukur yang digunakan benar-benar mampu mengukur hal atau variabel yang diukur. Apakah jika setiap butir soal sudah mencerminkan indikator

variabel yang hendak diukur sudah dapat dinyatakan bahwa pengukuran sudah mengukur kompetensi yang dimaksud.

Persoalan yang dihadapi dalam pemenuhan validitas konstruk dalam ranah kognitif bukan hanya terbatas pada kesesuaian item dengan indikator dengan pencapaian kompetensi. Persoalan yang mendasar adalah apakah sejumlah kompetensi yang diukur berada dalam satu dimensi. Jika tidak dalam satu dimensi, akan dipertanyakan bagaimana pemenuhan validitas konstruknya. Oleh karena tidak mudah untuk memenuhi validitas konstruk, dalam pengukuran penguasaan ranah kognitif akan lebih mudah dipahami jika pemenuhan validitas dibatasi dalam pemenuhan validitas isi yang tergambar dalam *learning continuum*.

Dalam bidang psikologi, termasuk di dalamnya psikologi pendidikan, pemenuhan kesahihan konstruk dikaitkan dengan sifat variabel yang akan diukur. Dalam hal ini adalah aspek psikologi meskipun variabel-variabel yang akan diukur bersifat abstrak. Dengan demikian, fenomenanya tidak dapat ditangkap secara langsung melalui panca indera. Oleh karena itu, perlu dicari indikator/indikan yang tepat agar variabel tersebut benar-benar dapat diukur. Hal ini akan mudah ditangkap jika dibandingkan dengan Ilmu Kimia. Fenomena reaksi kimia sulit ditangkap dengan panca indera. Oleh karena itu, cara yang dapat ditempuh adalah membubuhkan zat yang berfungsi sebagai indikator. Zat yang berfungsi sebagai indikator tersebut tidak akan ikut bereaksi, namun akan

berubah warnanya jika reaksi kimia sudah berlangsung. Dengan demikian, dapat dikatakan bahwa zat A sudah bereaksi dengan zat B karena warna zat yang berfungsi sebagai indikator sudah berubah warnanya.

Cara untuk memenuhi validitas konstruk adalah dengan membuat definisi operasional variabel yang akan diukur. Jika akan mengukur minat, maka dibuat terlebih dahulu apa yang dimaksud dengan minat secara operasional. Dari definisi operasional yang sudah dirumuskan selanjutnya dicari indikator-indikatornya. Dengan cara demikian, pemenuhan unidimensionalitas variabel yang akan diukur berpeluang dapat dipenuhi.

Setelah variabel yang akan diukur dijabarkan ke dalam indikator-indikatornya barulah disusun pertanyaan-pertanyaan yang mencerminkan masing-masing indikator tersebut. Dengan demikian, alat uji akan memiliki kesahihan konstruk jika item-itemnya mencerminkan indikator-indikator dari variabel yang akan diukur, juga ketepatan memilih indikator-indikator dari variabel yang diukur.

Menurut Embretson & Gorin (2001) validitas konstruk adalah sentral untuk menetapkan mutu tes. Secara tradisional, spesifikasi item sering hanya samar-samar. Pengembang tes membuat spesifikasi item yang sering kali berisi pertimbangan-pertimbangan dari segi isi secara umum (seperti penetapan ruang lingkup/topik area dan ringkasan dari materi/isi) atau samar-samar menggambarkan pengolahan derajat pemrosesan berpikir (seperti abstrak melawan konkret). Begitu keseluruhan spesifikasi item dihasilkan kemudian dilakukan telaah (*review*) item oleh

penelaah (*reviewer*) untuk meyakinkan ketepatan kunci jawaban dan menguji konten/isi item untuk berbagai isu penyetaraan dan mutu tes. Metode-metode psikometrik diterapkan setelah item dikembangkan. Statistika item dari uji coba empiris menjadi sangat esensial untuk menentukan kualitas item, terutama untuk mengevaluasi bahwa proses-proses "konstruk-relevansi" telah terukur. Secara umum item-item yang tidak saling berkorelasi dengan item-item yang lain tidak dipilih. Prosedur tradisional yang standar tersebut sangat cocok dengan paradigma validitas konstruk menurut teori klasik.

Berdasarkan prinsip psikologi kognitif, bahwa yang pertama yang terpenting dalam pemenuhan validitas konstruk suatu tes kemampuan/abilitas adalah adanya dukungan satu set prinsip teoretis dari item-item tes yang akan disusun. Kemudian, suatu set standar yang baru yang disusun harus didasarkan pada prinsip-prinsip psikologi kognitif. Mengutip pendapat Messick (1995) bahwa item-item yang disusun tidak sekadar cocok/sesuai dengan kriteria/ukuran tradisional, seperti tingkat kesulitan yang sesuai dan daya pembeda yang tinggi, tetapi item-item tersebut juga dibenarkan sebagai bagian juga relevan dengan konstruk dari proses-proses kognitif. Terakhir, bahwa tes yang disusun adalah untuk mengindikasikan kecakapan atau untuk mendiagnosis ketrampilan yang ada sehingga memerlukan bermacam informasi baru tentang item-item yang disusun.

Pemenuhan validitas konstruk secara empiris akan dapat diperoleh jika aspek yang diukur bersifat spesifik. Misalnya, akan lebih mudah memenuhi bukti empiris

mengembangkan tes untuk mengukur kemampuan berkreasi daripada mengembangkan untuk mengukur seluruh aspek kognitif mulai dari mengingat sampai berkreasi. Pemenuhan validitas konstruk secara empiris dapat didekati dengan membandingkan skor yang diperoleh dari kelompok-kelompok yang telah diketahui karakteristiknya. Misalnya, secara umum kemampuan seseorang akan bertambah meningkat sejalan dengan pertambahan umur. Artinya, bahwa suatu tes yang dikenakan pada kelompok yang lebih tua usianya akan menunjukkan skor rata-rata yang lebih tinggi dibanding kelompok yang lebih muda. Skor tes prestasi akan dicapai kelompok yang pernah ikut pelatihan dibandingkan yang belum mengikutinya. Skor rata-rata kemampuan *reading comprehension* pada perempuan lebih tinggi dibandingkan dengan laki-laki. Jadi, jika hasil tes yang dikembangkan setelah diujicobakan pada kelompok yang telah diketahui spesifikasinya menunjukkan hasil yang berkesesuaian berarti terpenuhi bukti empirisnya.

Hasil skor jenis tes tertentu, seperti tes intelegensi, tidak terpengaruh dengan adanya tambahan pengalaman pelatihan. Namun, tes prestasi belajar akan menunjukkan skor yang berbeda antara peserta didik sebelum dan sesudah belajar. Dalam eksperimen pun, akan ada perbedaan skor rata-rata antara kelompok perlakuan dan kelompok pembanding. Bila perlakuan memberikan dampak yang positif, skor rata-rata kelompok perlakuan akan menunjukkan hasil yang lebih tinggi secara nyata (*significant*). Hal ini akan menjadi bukti empiris terhadap tes yang dikembangkan.

Membandingkan skor hasil tes yang dikembangkan dengan skor tes lain yang sama/sebidang juga dapat untuk menunjukkan bukti empiris. Misalnya, hasil suatu tes bakat (*aptitude test*) yang satu akan memiliki korelasi dengan hasil tes bakat yang pernah dikembangkan sebelumnya. Cara ini merupakan cara yang dilakukan dengan menghubungkan hasil tes dengan kriteria (*criterion-related study*).

Cara yang tepat untuk memperoleh bukti empiris adalah menghubungkan dengan penguasaan testi terhadap tes yang ditempuhnya. Di dalam tes beracuan norma sulit untuk mengetahui bagaimana hasil yang dicapai testi orang per orang karena testi yang satu dibandingkan dengan testi lain di dalam kelompoknya. Sementara itu, di dalam tes beracuan kriteria, skor suatu item tes dapat menunjukkan *prefect score* (*zero variability*) jika semua testi benar-benar menguasai aspek yang diujikan melalui item yang bersangkutan. Tes prestasi yang diwujudkan dalam untuk pengujian guna mengetahui penguasaan testi (*mastery testing*) umumnya memiliki variabilitas yang kecil, bahkan variabilitasnya dapat sebesar 0.

c. **Bukti yang berkaitan dengan kriteria** (***criterion-related validity***)

Dalam hal ini untuk melihat kesahihan suatu alat ukur justru dibandingkan dengan alat ukur yang lain, yang standar atau yang baku sebagai patokan (kriteria). Alat ukur yang kita buat dinyatakan memiliki kesahihan yang tinggi jika hasil pengukurannya menunjukkan korelasi yang sangat positif dengan hasil pengukuran yang menggunakan alat ukur baku/standar. Kesahihan yang

demikian disebut kesahihan *concurrent validity*. Pemenuhan validitas berdasarkan bukti empiris yang dikaitkan dengan kriteria dapat dilakukan pula dengan mengorelasikan hasil tes yang telah dikembangkan dengan tes lain yang diterapkan pada waktu mendatang yang berdekatan waktunya. Pemenuhan validitas ini disebut pemenuhan prediktif (*predictive validity*). Kedua hal ini relatif sulit sekali dipenuhi mengingat tidak tersedianya tes standar maupun tes lain yang dapat digunakan untuk menunjukkan daya prediktif dari tes yang telah dikembangkan.

d. Kesahihan yang lain selain yang dilihat dari tiga aspek

1) Kesahihan muka (*face validity*)

Kesahihan ini dibedakan menjadi dua. Pertama adalah kesahihan muka yang hanya dilihat berdasarkan tampak luarnya. Misalnya, pada suatu item mata pelajaran Biologi tentang klasifikasi mikroorganisme, item tersebut dinyatakan memenuhi kesahihan muka jika yang ditanyakan adalah hal-hal yang menyangkut klasifikasi organisme. Kedua, kesahihan ditinjau berdasarkan pandangan orang yang lebih ahli, dapat berarti ahli dalam ilmunya ataupun ahli dari segi evaluasi atau dari segi bahasa. Oleh karena itu, minta pendapat orang yang lebih ahli akan meningkatkan kesahihan item yang telah disusun.

2) Kesahihan antarbudaya (*cultural validity*)

Kesahihan alat ukur juga tercermin dari keseragaman daya tangkap subjek terhadap alat ukur yang

bersangkutan. Suatu butir tes model kasus harus menyajikan kasus yang dipahami secara merata oleh para subjek yang diukur. Seorang yang memahami Biologi akan lebih terpandu mengerjakan soal bahasa Inggris dengan kasus Biologi daripada dengan kasus Sastra misalnya. Jenis-jenis tanaman tumpangsari sudah tidak perlu dihafal oleh anak-anak yang hidup di pedesaan, namun untuk anak kota menjadi pekerjaan rumah tersendiri.

Alat-alat ukur yang diadopsi dari luar negeri harus dikaji benar apakah pertanyaan-pertanyaan yang ada di dalamnya bebas dari faktor budaya. Tidak otomatis bahwa alat ukur yang sudah teruji kesahihannya di luar negeri dipertanyakan kembali jika digunakan di dalam negeri.

d. Hal-Hal yang Diperhatikan Untuk Meningkatkan Validitas

Kesahihan alat ukur juga dapat dipenuhi dengan membuat alat ukur sebaik mungkin antara lain dengan cara berikut ini.

1) Petunjuk mengerjakan harus jelas. Jika alat ukurnya berupa perangkat tes, perintahnya harus jelas agar testi/peserta ujian mengerti apa yang harus dikerjakan. Demikian pula, jika berupa instrumennya berupa angket, usahakan agar responden mengerti cara pengisiannya.
2) Struktur kalimat harus benar, lugas, komunikatif, dan kata-kata yang digunakan sesuai dengan jenjang kemampuan testi/peserta ujian.

3) Konstruksi alat ukur harus memenuhi persyaratan sesuai dengan jenis alat ukurnya. Misalnya alat ukur berupa tes objektif/pilihan ganda harus memenuhi kaidah pembuatan item jenis tersebut. Maksudnya, persyaratan stem, kunci dan pengecoh harus dipenuhi.
4) Pernyataan-pernyataannya jangan mengandung arti ganda karena testi yang lebih cerdas cenderung tidak akan menjawab/mengerjakannya.
5) Cukup waktu untuk mengerjakan.
6) Alat ukur tidak terlalu panjang atau terlalu pendek.
7) Khusus alat ukur berupa alat tes juga memperhatikan hal berikut ini.
   (a) Tingkat kesukaran item hendaknya disesuaikan dengan kemampuan subjek uji.
   (b) Setiap item tes seleksi harus dapat membedakan mana kelompok atas dan mana kelompok bawah. Item tes seleksi yang baik adalah tidak terlalu mudah dan tidak terlalu sukar (indeks kesukaran item idealnya sebesar 0,5) dan memiliki daya pembeda yang tinggi (indeks daya pembeda idealnya 1,0) sehingga mampu membedakan mana yang pandai dan yang tidak.
   (c) Setiap item tes prestasi benar-benar untuk mengukur keberhasilan belajar, jika tesnya untuk mengukur keberhasilan belajar.
   (d) Item tes prestasi hasil belajar yang baik adalah yang berhasil dikerjakan oleh peserta didik yang sungguh-sungguh berhasil dalam belajar dan yang pasti gagal dikerjakan oleh peserta didik yang belum terlibat dalam proses pembelajaran. Hal

tersebut akan ditunjukkan oleh besarnya indeks sensitivitas item sebesar 1,0.

(e) Item diurutkan dari yang termudah ke yang paling sukar.

(f) Jawaban jangan sampai ditemukenali oleh subjek uji, misalnya karena urutan jawabannya terpola (tidak acak).

(g) Jangan sampai terjadi penyimpangan saat dilaksanakan tes dan jangan sampai salah ketika melakukan penskoran hasil pengukuran.

4. Reliabilitas

Pengertian reliabilitas berkaitan dengan konsistensi. Suatu alat ukur dinyatakan *reliable*/andal jika memberikan hasil yang sama pada berkali-kali pengulangan pengukuran. Reliabilitas berlaku pada tingkat suatu perangkat tes. Jadi tidak berlaku untuk masing-masing item tes penyusun suatu perangkat alat ukur.

Berkait dengan reliabilitas tes, Frisbie (2005) menyatakan bahwa reliabilitas tes hasil belajar berbeda dengan reliabilitas tes untuk seleksi karena tes hasil belajar memiliki varians yang rendah jika peserta didik berhasil semua dalam belajarnya. Oleh karena itu secara praktis Frisbie membuat tabel pembandingan ideal antara interpretasi dalam situasi *Norm Reference* (NR) dan *Criterion Reference.*

Tabel 4. *Comparative ideals for NR and CR interpretation situations*

	Norm-Reference	*Criterion-Reference*
Item difficulty	*Moderate*	*Easy-to-hard*
Item discrimination	*High positive*	*Nonnegative*
Score variability	*Maximize*	*Non-issue*
Error estimate	*High reliability coefficient*	*High decision consistency index*

Berdasarkan informasi pada tabel tersebut, tes untuk mengukur keberhasilan belajar menggunakan tes yang skornya diinterpretasikan dengan mengacu kepada kriteria (*criterion-reference*) sehingga item-itemnya memiliki tingkat kesulitan item bervariasi dari mudah sampai sukar (sebagai cerminan tingkat keberhasilan belajar) dan tidak boleh memiliki indeks daya pembeda yang negatif (sebagai cerminan tidak ada testi/peserta ujian yang cerdas menjawab salah). Oleh karena itu, estimasi *error* didasarkan pada tingginya indeks konsistensi (indeks yang tinggi menunjukkan semua testi/peserta ujian yang sudah belajar pasti dapat mengerjakan dengan benar, sementara semua testi/peserta ujian yang belum belajar pasti tidak dapat mengerjakan dengan benar). Sebaliknya, tes untuk tujuan seleksi adalah tes yang dapat memisahkan kelompok yang lolos seleksi dan yang tidak lolos seleksi. Oleh karena itu, interpretasinya dalam dengan mengacu kepada distribusi normal (*norm-reference*) sehingga item-itemnya memiliki indeks kesulitan yang harus moderat/sedang. Artinya, item-itemnya tidak terlalu sukar dan tidak terlalu mudah (sebagai cerminan bahwa kelompok atas yang pasti dapat

mengerjakan) dan indeks daya pembeda harus tinggi (sebagai cerminan dapat membedakan kelompok atas dan bawah dengan meyakinkan). Oleh karena itu, estimasi *error* didasarkan pada tingkat tingginya keandalan/reliabilitas tes (indeks yang tes tinggi mencerminkan bahwa semakin cerdas testi/peserta ujian di dalam kelompoknya semakin tinggi pula skor yang diperolehnya).

Hasil analisis item untuk tujuan seleksi yang diterapkan pada item untuk tujuan pengukuran prestasi akan menunjukkan hasil yang tidak sesuai harapan. Berikut disajikan ilustrasinya.

Tabel 5. Analisis Item dengan mencari Daya Beda dan Indeks Kesukaran Item untuk 15 Item

Peserta Ujian	Nomor item															Skor	Nilai (0 -100)	Capaian KKM 75
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15			
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	15	100	Berhasil
2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	15	100	Berhasil
3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	15	100	Berhasil
4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	15	100	Berhasil
5	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	15	100	Berhasil
6	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	15	100	Berhasil
7	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	15	100	Berhasil
8	1	1	1	0	1	1	0	0	1	1	1	1	1	1	1	12	80	Berhasil
9	1	0	0	1	1	1	1	1	0	0	1	1	1	1	1	11	73	Gagal
10	1	0	0	1	0	1	1	1	1	1	0	1	1	1	1	11	73	Gagal
11	0	0	1	0	1	1	1	1	1	1	0	1	1	1	1	11	73	Gagal
12	0	1	1	1	0	1	0	1	0	1	1	0	1	0	1	9	60	Gagal
13	1	1	1	1	1	1	0	0	0	0	1	0	0	0	0	7	47	Gagal
14	1	1	1	0	1	1	0	0	0	0	1	0	0	0	0	6	40	Gagal
15	0	0	0	0	0	1	1	1	0	0	1	1	1	0	0	6	40	Gagal
16	1	1	0	0	0	0	0	0	0	0	1	0	0	0	0	3	20	Gagal
17	1	1	0	0	0	0	1	1	1	0	0	1	1	1	0	8	53	Gagal
18	1	0	1	0	0	0	0	1	1	1	0	0	1	1	1	8	53	Gagal

19	0	1	1	0	0	0	1	0	0	1	0	1	0	0	1	6	40	Gagal
20	0	0	0	0	1	0	1	0	1	0	0	1	0	1	0	5	33	Gagal
21	0	0	1	1	1	0	0	0	1	0	0	0	0	1	0	5	33	Gagal
22	0	0	1	0	1	0	1	0	0	0	0	1	1	0	0	5	33	Gagal
23	0	0	0	1	1	0	0	0	0	1	0	0	1	0	1	5	33	Gagal
24	0	0	0	1	0	0	0	1	0	0	1	0	0	0	0	3	20	Gagal
25	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	2	13	Gagal
26	0	0	0	0	0	0	0	1	0	0	0	1	1	0	0	3	20	Gagal
27	0	0	0	0	1	0	1	0	0	0	0	1	0	0	0	3	20	Gagal
28	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	2	13	Gagal
29	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	2	13	Gagal
30	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	7	Gagal
Kel atas	12	11	12	11	12	15	11	12	10	11	13	12	13	11	12			
Kel bawah	3	3	5	4	6	0	5	4	4	3	6	6	5	4	3			
Daya beda	0.6	0.5	0.5	0.5	0.4	1.0	0.4	0.5	0.4	0.5	0.5	0.4	0.5	0.5	0.6			
Tingkat kesukaran	0.5	0.5	0.6	0.5	0.6	0.5	0.5	0.5	0.5	0.5	0.6	0.6	0.6	0.5	0.5			

Jika dilihat daya beda minimal 0,4 berarti seluruh item memenuhi syarat tes standar. Tingkat kesukaran berkisar antara 0,5 sampai 0,6 juga hampir ideal. Kenyataan hanya 8 dari 30 peserta didik yang mencapai KKM. Belum lagi jika memperhatikan syarat item NRT di mana agar upaya muncul kelompok bawah, masih ada syarat bahwa minimal setiap pilihan pengecoh pada setiap soal pilihan ganda harus ada 5% peserta yang terkecoh. Jika ada 4 pengecoh/distraktor dan satu kunci, dan yang terkecoh peserta yang sama, dalam keadaan ekstrim akan ada 20% peserta didik yang memperoleh skor 0 dan ada 80% peserta didik yang memperoleh skor 100. Jika dikombinasi dengan syarat daya beda dan indeks kesukaran, maka tidak mungkin ada testi yang memperoleh skor mutlak karena indeks tingkat kesukaran minimal 0,3. Dengan kata lain, harus ada peserta didik sebanyak 30% yang masuk dalam

kategori gagal. Karena itulah, penilaian prestasi harus menggunakan instrumen yang memiliki indek sensitivitas yang tinggi.

Indek sensitivitas suatu item dapat diperoleh, jika tersedia hasil pretes dan postes. Indeks sensitivitas suatu item diperoleh dengan mencari banyaknya testi yang berhasil mengerjakan item yang bersangkutan dalam postes dikurangi banyaknya testi yang berhasil mengerjakan item yang bersangkutan dalam pretes, kemudian hasilnya dibagi dengan banyaknya peserta.

Jika mengikuti pendapat Frisbie maka untuk mengetahui apakah item tersebut memiliki konsistensi yang tinggi bukan dihitung berdasarkan indeks reliabilitas Alpha Cronbach atau Kuder-Richardson 20 atau 21 namun didasarkan pada *agrrement index atau Kappa index.* Pada lampiran disajikan cara menghitungnya hanya dengan melakukan sekali tes.

BAB IX
ANALISIS ITEM DAN KEGUNAANNYA

A. Analisis Item untuk Tujuan Pemenuhan Reliabilitas

Analisis item dapat dilakukan dengan pendekatan teori tes klasik (*Classical Test Theory* atau CTT) dan teori tes modern yang dikenal dengan teori respons item (*Item Respons Theory* atau IRT). Masing-masing pendekatan sudah disajikan program analisis dalam bentuk *software* dengan spesifikasi, kelebihan, dan kelemahan berkaitan dengan penyajian informasi yang dapat diperoleh.

Pemilihan item tes dalam prosedur pengembangan tes menggunakan CTT didasarkan pada: (a) nilai/indeks kesukaran item, dan (b) korelasi skor item dan skor total atau disingkat korelasi item-total yang menunjukkan reliabilitas. Item yang memiliki korelasi item-total paling tinggi dipakai sebagai elemen suatu tes untuk membentuk suatu skala dengan konsistensi internal tinggi. Tujuannya adalah untuk memperkecil sumbangan *error* acak skor-skor tes sehingga reliabilitas tes dapat ditingkatkan. Distribusi

skor-skor tes total yang diperoleh dari lapangan dibandingkan dengan distribusi yang diinginkan oleh pengembang tes. Sejumlah item mungkin perlu diganti untuk memperoleh sedekat/semirip mungkin antara distribusi skor total yang diinginkan dan distribusi skor total yang diperoleh dari lapangan. Tes dengan format-format paralel untuk pencarian bukti empiris pada umumnya diciptakan untuk memperoleh distribusi-distribusi skor tes yang identik. Kesamaan nilai rata-rata, varians, dan *error* skor ditafsirkan sebagai bukti empiris tes dengan format paralel (Stark *et al.*, 2001).

Langkah pertama dalam CTT adalah sebelum penulisan item mulai, pengembang tes harus mempunyai suatu pemahaman yang baik terhadap konstruk variabel (kemampuan) yang akan diukur. Menurut Stark *et al.* dengan mengacu pendapat Nunnally *et al.*, berdasarkan "*rule of thumb*" lazimnya disepakati bahwa banyaknya item tes yang harus dibuat sedikitnya dua kali dari banyaknya item tes yang akan dikemas sebagai perangkat tes yang siap digunakan.

Jika bentuk itemnya berupa pilihan ganda maka sejumlah besar item harus disusun untuk menghasilkan format-format ganda yang dikembangkan. Item-item tersebut harus diujikan terlebih dahulu menggunakan suatu sampel yang serupa dengan populasi yang dituju. Sampel ini, yang selanjutnya diacu sebagai sampel untuk pedoman saat dilakukan kalibrasi, haruslah besar. Tujuannya adalah agar cukup untuk menyediakan statistika item CTT yang stabil. Item-item dengan korelasi item-total tinggi harus ada di dalam tes karena item-item tersebut meningkatkan

konsistensi skala internal (reliabilitas). Hal seperti itu juga akan mereduksi/memperkecil kesalahan baku pengukuran (*standard error of measurement*).

Kesukaran item (dinyatakan sebagai nilai p) juga harus dipertimbangkan untuk membuat suatu tes dengan distribusi skor total yang diinginkan (Stark *et al.*, 2001). Selain itu, untuk memperoleh distribusi skor skala yang diinginkan maka harus dilakukan penggantian item. Agar skalanya meningkat maka item dengan nilai p yang rendah harus diganti dengan nilai p yang tinggi. Untuk memperkecil dampak penggantian item terhadap reliabilitas maka item-item yang memiliki korelasi item-total yang rendah diganti dengan item-item yang memiliki korelasi item-total yang tinggi. Di dalam praktik, beberapa penyeimbangan konten/isi juga diperlukan. Oleh karena itu, setelah dilakukan penggantian kemudian item-item tersebut dianalisis lagi.

Ada keterbatasan penggunaan pendekatan CTT untuk analisis item (Stark *et al.*, 2001). Pertama, statistika CTT bergantung kepada subpopulasi penempuh tes. Berbeda grup penempuh tes berbeda pula nilai rata-rata skor dari atribut variabel yang diukur. Dengan demikian, para pengembang tes harus hati-hati ketika memilih sampel untuk kalibrasi item. Jika sampel-sampel kalibrasi berbeda karakteristik/sifat dengan sampel operasional (sampel populasi yang sesungguhnya sebagai target), properti-properti psikometri hasil pengukuran akan berubah secara dramatis. Kedua, di dalam CTT, ketepatan pengukuran suatu tes (galat baku pengukuran atau *standard error of measurement*) secara implisit dirata-ratakan ke semua level

kemampuan yang diukur karena tidak ada informasi tentang kemampuan testi (*ability*). Dengan demikian, ketepatan pengukuran pada level-level skor yang tertentu tidak dikenal/tidak diketahui. Oleh karena itu, dikembangkan analisis item menggunakan teori respons item atau *item response theory* (IRT). Program analisis item yang mendasarkan pada IRT akan menghitung *standard error of measurement* item demi item. Dengan demikian, kumulatif *standard error of measurement* yang diujudkan dalam bentuk *reliability case of estimate* dalam program Quest dapat dinyatakan sebagai angka yang menunjukkan reliabilitas tes.

Langkah kegiatan mengonstruksi tes menggunakan pendekatan IRT, seperti halnya pada penggunaan pendekatan CTT. Penulis harus membuat dua sampai tiga kali banyaknya item seperti yang diinginkan di dalam perangkat tes yang final akan digunakan. Skor mentah akan dikalibrasi menurut IRT menjadi skor dalam bentuk skala *logit*. Untuk itu diperlukan sampel kalibrasi heterogen yang besar, idealnya akan mencapai angka kalibrasi yang benar-benar stabil jika sampel mencapai 250 untuk item pilihan ganda. Model IRT yang lebih kompleks, seperti model IRT untuk skala politomus (soal bentuk uraian dengan lebih dari dua kategori skor), memerlukan sampel lebih besar untuk mengestimasi parameter.

Jika menggunakan program tertentu dengan pendekatan berdasarkan teori kemungkinan yang maksimum (*maximum likelihood*) seperti program Quest, maka sebelum mengestimasi parameter item, perlu untuk melakukan suatu analisis item menurut teori tes klasik untuk menghapuskan item-item yang mempunyai skor

mendekati nihil (tidak atau sedikit sekali yang dapat mengerjakan), tentu saja item yang demikian akan memiliki korelasi-korelasi item-total negatif. Item ini akan menyebabkan permasalahan konvergensi/pemusatan (Stark *et al.* (2001). Demikian pula item yang mempunyai skor prefect, di mana untuk tes pilihan ganda skor *prefect* adalah 1 untuk setiap testi atau *person/case*.

Program analisis atau disebut program kalibrasi menggunakan IRT adalah mendasarkan pada distribusi logistik, yakni distribusi yang menyerupai distribusi normal dengan nilai logistik D sebesar 1,7. Analisis item menggunakan IRT ada yang melakukan kalibrasi berdasarkan satu parameter yakni hanya didasarkan pada tingkat kesukaran (diberi simbol β atau b) sehingga disebut model satu parameter logistik tau model 1-PL atau disebut Model Rasch (Rasch *Model*). Ada yang mendasarkan pada dua parameter, yakni daya beda (diberi simbol a) dan tingkat kesukaran (b) sehingga disebut model 2-PL. Ada pula yang mendasarkan pada tiga parameter, yakni daya pembeda, tingkat kesukaran, dan *guessing* (diberi simbol c), sehingga disebut Model 3-PL.

Berdasarkan skala yang digunakan ada yang hanya dibuat skala dikotomus, yakni hanya dibedakan menjadi dua kategori yakni kategori 1 untuk skor 0 dan kategori 2 untuk skor 1, seperti skor tes pilihan yang pada akhirnya hanya ditetapkan benar atau salah. Ada pula yang dapat dibuat dalam bentuk skala politomus, yakni hanya dibedakan lebih dari dua kategori. Misalnya, testi diminta mengemukakan dua faktor yang menjadi penyebab terjadinya suatu peristiwa. Dengan demikian, jawaban testi

ada tiga kategori, yakni kategori 1 untuk skor 0 (testi tidak menjawab), kategori 2 untuk skor 1 (testi hanya mengemukakan dengan benar satu faktor penyebab), dan kategori 3 untuk skor 2 (testi dapat menyebutkan dengan benar dua faktor penyebab yang dimaksudkan).

3. Kegunaan Hasil Analisis Item untuk Pemenuhan Validitas

Istilah valid digunakan dalam pengukuran berdasarkan teori tes klasik. Pemenuhan validitas tes tidak mudah diperoleh. Validitas tes hanya dapat diperoleh ketika peneliti memiliki tes yang sudah baku sebagai pembanding jika menggunakan pendekatan *concurrent validity*. Atau menggunakan hasil tes lain yang memiliki daya prediksi tinggi jika mewnggunakan pendekatan *predictive validity*. Menurut Wright (1999) juga Wright & Masters (1982) jika suatu item sesuai (*fit*) dengan model yang digunakan (maksudnya seperti model logistik satu parameter atau dua parameter) maka dapat diartikan bahwa item tersebut dapat mengukur kemampuan sehingga item tersebut boleh dinyatakan "valid" sebagai alat ukur. Oleh karena itu, jika semua item sesuai (*fit*) terhadap model yang digunakan maka tes yang tersusun atas item-item yang bersangkutan juga dapat dinyatakan "valid" sebagai alat ukur menurut teori tes klasik.

BAB X

ANALISIS ITEM MENGGUNAKAN PROGRAM *QUEST*

A. Ukuran sampel

Semakin bertambah banyak parameter di dalam model politomus sebagai lawan model dikotomus, akan semakin bertambah pula informasi di dalam data. Namun, diperlukan estimasi yang stabil di dalam ukuran sampel yang sama. Ukuran sampel untuk data politomus menggunakan *Graded Model* (GM) yang merupakan model 2-PL sekitar 250 dapat diterima untuk aplikasi dalam penelitian, sedangkan 500 sampai 1000 untuk penggunaan operasional (Muraki & Bock, 1998: 35). Ahli lain ada yang menyatakan bahwa untuk keperluan kalibrasi dalam IRT ukuran sampel antara 200 sampai 1000 tergantung model yang dipilih. Penelitian disertasi dapat menggunakan sampel yang kecil (Crocker & Algina, 1986: 322). Sebagian ahli menyatakan bahwa ukuran sampel khusus untuk model

1-PL berupa Rasch *Model* (RM) antara 30 sampai 300 dengan batas INFIT t sebesar -2 sampai +2 (Bond & Fox, 2007: 43). Jadi dalam hal ini menggunakan batas kesalahan 5%, sehingga besarnya nilai INFIT t ±1,96 atau dibulatkan menjadi ±2,0. Dengan demikian, suatu item menjadi tidak *fit* menurut Model Rasch bila memiliki nilai <-2,0 atau > +2,0 sebagai pembulatan harga 1,96 untuk nilai t dengan alpha 5%.

B. Prinsip Analisis Menggunakan Program Quest

Hasil tes uraian dapat dianalisis menurut model kredit parsial (*Partial Credit Model* atau PCM) sedangkan hasil tes pilihan dianalisis menurut Model Rasch (*Rasch Model* atau RM).

Beberapa pertimbangan dalam pemakaian PCM sebagai perluasan RM yang merupakan model 1-PL, dapat menggunakan sampel yang tidak sebesar kalau melakukan kalibrasi data politomus menggunakan model 2-PL atau 3-PL (Keeves & Masters, 1999: 12-13). Kedua, bahwa karakteristik respons terhadap setiap item mengikuti PCM yakni bahwa tingkat kesukaran dari suatu tahapan kategori di bawahnya ke kategori di atasnya tidak sama antaritem satu dan yang lain, sehingga besarnya *delta* untuk suatu tahapan kategori di bawahnya dan *delta* untuk tahapan kategori di atasnya tidak sama antaritem satu dengan item lainnya.

PCM adalah perluasan dari model 1-PL/RM. Persamaan RM menurut Han & Hambleton (2007: 15) dituliskan sebagai berikut.

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

e adalah konstanta eksponensial yang nilainya 2,718, dan *D* adalah faktor penskalaan yang nilainya 1,7. Nilai kemungkinan setiap testi berhasil mengerjakan item i dipahami sebagai fungsi logistik perbedaan dua parameter, yakni parameter kemampuan/kecakapan sebesar θ dan parameter tingkat kesukaran item sebesar b_i. Persamaan nomor satu dapat ditulis kembali sebagai berikut.

$$P_i(\theta) = \frac{1}{1+e^{-D(\theta-b_i)}} = \frac{\exp(D(\theta-b_i))}{1+\exp(D(\theta-b_i))} = \frac{P_{i1}(\theta)}{P_{i0}(\theta)+P_{i1}(\theta)} \qquad (2)$$

$P_{i1}(\theta)$ adalah peluang testi yang dipilih acak, yang memiliki tingkat kemampuan/kecakapan (*proficiency level*) sebesar θ untuk meraih skor 1 pada item *i*, sedangkan $P_{i0}(\theta)$ adalah peluang testi yang dipilih acak, yang memiliki tingkat kemampuan sebesar θ untuk memperoleh skor 0 pada item *i*.

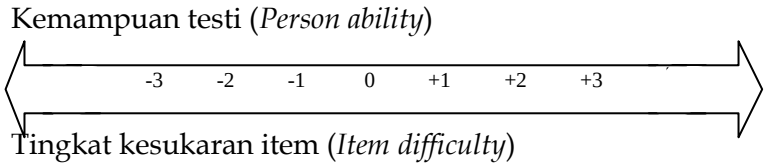
Persamaan RM dalam bentuk persamaan untuk testi (*case/person*) n dan item i dengan skor x sebesar 0 atau 1 dengan kemampuan sebesar β dan tingkat kesukaran item sebesar δ dituliskan sebagai berikut (Masters, 1999: 101; Wright & Masters, 1982: 39-40).

$$P_{nix} = \frac{1}{1+exp(\beta_n - \delta_{i1})} \quad \text{untuk x = 0} \qquad (3)$$

dan

$$P_{nix} = \frac{exp(\beta_n - \delta_{i1})}{1 + exp(\beta_n - \delta_{i1})} \quad \text{untuk x = 1} \qquad (4)$$

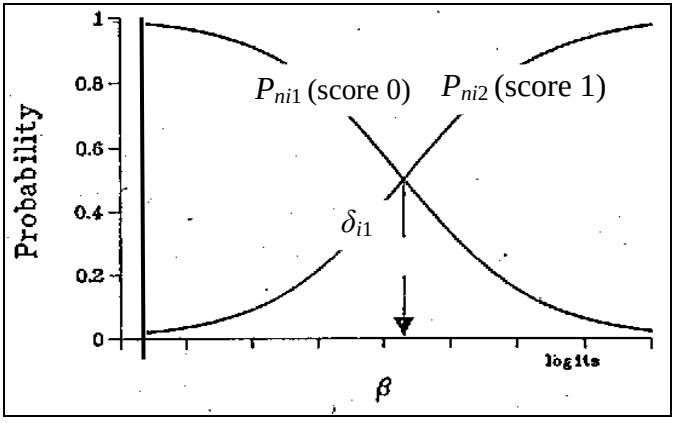
Tingkat kemampuan testi maupun tingkat kesukaran item dalam Rasch *Model* (RM) diekspresikan pada satu garis berupa absis pada grafik dengan satuan berupa *logit* (*logg-odd unit*). Garis tersebut terbentang dari -∞ sampai dengan +∞, bila digambar akan tersaji pada Gambar 1 (Keeves & Alagumalai, 1999: 27).



Gambar 1. Skala Rasch
(Sumber: Keeves & Alagumalai, 1999: 27).

Karena diekspresikan dalam satu garis maka dapat dibandingkan besarnya kemampuan testi/siswa dengan tingkat atau indeks kesukaran item yang diujikan. Testi yang memiliki skor *ability* di bawah atau lebih rendah dari indeks kesukaran item, maka testi yang bersangkutan tidak akan memiliki kemampuan untuk mengerjakannya dengan benar.

Lokasi tingkat kesukaran sebesar δ_{i1} pada grafik merupakan perpotongan antara curve respons P_{ni1} memiliki skor 0 dengan kurva respons P_{ni2}. memiliki skor 1, dan pada grafik akan terlihat sebagaimana tersaji pada Gambar 2.



Catatan: Titik potong antara kurva peluang skor 0 (kategori-1) dan skor 1 (kategori-2) menunjukkan lokasi *Delta*-1 sebesar δ_{i1}

Gambar 2. Kurva Peluang pada Skala Dikotomus Menuru *Rasch Model*
(Sumber: Wright & Masters, 1982: 40)

Skala politomus memiliki skor x sebesar 0, 1, 2, 3 ....., m_i. Peluang seorang testi (*case/person*) pada tingkat kemampuan θ meraih skor sebesar x di atas x-1 dapat dihitung dengan persamaan sebagai berikut (Han & Hambleton, 2007: 15).

$$\frac{P_{ix}(\theta)}{P_{ix-1}(\theta)+P_{ix}(\theta)} = \frac{\exp(D(\theta - b_{ix}))}{1+\exp(D(\theta - b_{ix}))} \text{ untuk x = 0, 1, 2, ..., } m_i \qquad (5)$$

$P_{ix}(\theta)$ dan $P_{ix\text{-}1}(\theta)$ mengacu pada peluang seorang testi (*case/person*) sebesar θ, meraih skor x dan x-1. Hal yang perlu diperhatikan bahwa pada persamaan nomor lima, jumlah parameter kesukaran item kini menjadi m_i (jumlah kategori respons dikurangi satu). Peluang seorang testi (*case/person*) yang dipilih acak, dengan tingkat kemampuan sebesar θ, untuk memperoleh skor x pada item i dapat dituliskan dengan persamaan sebagai berikut.

$$P_{ix}(\theta) = \frac{\exp\sum_{k=0}^{x}(D(\theta - b_{ik}))}{\sum_{h}^{m_i}\exp\sum_{k=0}^{h}(D(\theta - b_{ik}))} \quad \text{untuk x = 1, 2, 3, ….., } m_i \qquad (6)$$

Fungsi Persamaan nomor sembilan sering disebut fungsi respons kategori skor (*score category response function* atau SCRF).

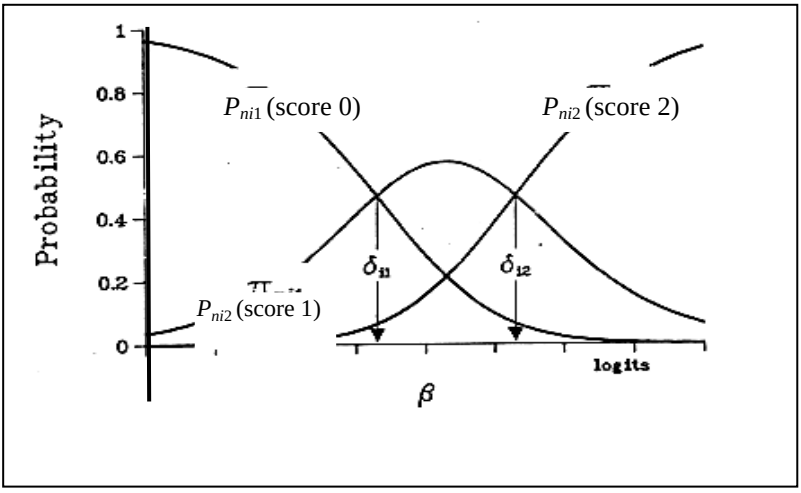
Masters (1999: 101) dan Wright & Masters, (1982: 39) menuliskan persamaan nomor untuk testi (*case/person*) n dan item i dengan skor x sebesar 0, 1, 2, …., m_i dengan kemampuan sebesar β dan tingkat kesukaran item sebesar δ yang dituliskan dalam PCM dengan rumus sebagai berikut.

$$P_{nix} = \frac{1}{1 + exp(\beta_n - \delta_{ij})} \quad \text{untuk x = 0} \qquad (7)$$

dan

$$P_{nix} = \frac{exp(\beta_n - \delta_{ij})}{1 + exp(\beta_n - \delta_{ij})} \quad \text{untuk x = 1, 2, 3, ..., } m_i \qquad (8)$$

Dengan demikian, tingkat kesukaran item (*difficulty*) untuk item i sebesar δ akan terurai menjadi nilai *delta* sebesar δ_{ij} untuk x = 1, 2, 3, m_i. Item nomor 1 yang memiliki tiga kategori atau diskor secara politomus tiga kategori, memiliki δ_{11} dan δ_{12}, item nomor 2 memiliki δ_{21} dan δ_{22}. Besarnya nilai *delta*-1 menunjukkan nilai yang diperlukan testi (*case/person*) untuk berpindah dari kategori-1 (skor 0) ke kategori-2 (skor 1) dan nilai *delta*-2 menunjukkan nilai yang diperlukan untuk berpindah dari kategori-2 (skor 1) ke kategori-3 (skor 2). Besarnya *delta*-1 dapat lebih kecil, sama, atau lebih besar dari *delta*-2. *Delta*-1 dalam grafik menunjukkan perpotongan antara kurva respons P_{ni1} (skor 0) dengan kurva respons P_{ni2} (skor 1) dan *delta*-2 menunjukkan perpotongan kurva respons kurva respons P_{ni2} (skor 2) dengan kurva respons P_{ni3} (skor 3), dan pada grafik akan terlihat pada Gambar 3.



Catatan: Titik potong antara kurva peluang skor 0 (kategori-1) dan skor 1 (kategori-2) menunjukkan lokasi *Delta*-1 sebesar δ_{i1}, titik potong kurva peluang skor (kategori-2) dan skor 2 (kategori-3) menunjukkan lokasi *Delta*-2 sebesar δ_{i2}

Gambar 3. Kurva Peluang Skala Politomus Tiga Kategori Menurut *Partial Credit Model* (PCM) (Sumber: Wright & Masters, 1982: 44).

Kategori pada *Rating Scale Model* atau RSM, merupakan kategori yang berjenjang/berperingkat (*ordered category*). Sebagai contoh, respons item 1, 2, 3 dengan kategori-1 "tidak setuju", kategori-2 "setuju", dan kategori-3 "sangat setuju merupakan kategori yang berjenjang. Oleh karena itu, nilai *delta* sebesar δ_{ij} (yang menunjukkan karakteristik spesifik tingkat kesukaran item i pada tahapan/step/ kategori j) dipecah menjadi tingkat kesukaran (*difficulty*) sebesar δ_i ditambah nilai *tau* sebesar τ_{ij} (yang menunjukkan karakteristik spesifik tingkat kesukaran suatu tahapan/step/kategori j dari item i) mengikuti pada formula yang diajukan Andrich (1978). Oleh karena itu, persamaannya dapat ditulis sebagai berikut (Masters, 1999: 101 dan Wright & Masters, 1982: 39).

$$P_{nix} = \frac{exp(\beta_n - \delta_i - \tau_{ij})}{1 + exp(\beta_n - \delta_i - \tau_{ij})} \quad \text{untuk } x = 1, 2, 3, ..., m_i \qquad (9)$$

Elemen sentral dari program Quest adalah IRT mengikuti Rasch *Model* (RM). Dalam hal ini, dapat pula

digunakan pada data respons yang diskor secara politomus. Program Quest dalam melakukan estimasi parameter, baik untuk item maupun untuk testi (*case/person*) menggunakan *unconditional* (UCON) atau *joint maximum likelihood* (Adam & Khoo, 1996: 89).

Skor mentah seorang testi dalam penskalaan sebesar *r* dikonversi menjadi skala *logit* yang menunjukkan n kemampuan sebesar b = log[(*r*/(*L*-*r*)], di mana *L* adalah banyaknya *activities* (item). Sementara, nilai r dapat dikonversi menjadi skala *logit* yang menunjukkan tingkat kesukaran sebesar d = log[(*N*-*S*)/*S*)], di mana *N* adalah banyaknya testi (*case/person*) dan *S* adalah skor suatu item (Wright & Masters, 1982: 28-31). Besarnya *S* untuk data pengukuran yang diskor secara politomus dalam program Quest diubah menjadi w_{ij} dan tingkat kesukaran sebesar d akan diubah menjadi nilai δ_{ij}. Persamaan untuk RSM dalam program Quest dituliskan sebagai berikut.

$$P(X_{ni} = x_{ni}) = \frac{\exp \sum_{j=0}^{x_n} w_{ij}(\beta_n - \delta_i - \tau_{ij})}{\sum_{k=0}^{ni} \exp \sum_{j=0}^{k} w_{ij}(\beta_n - \delta_i - \tau_{ij})} \qquad (10)$$

β_n adalah komponen tingkat kemampuan (*ability*) dari testi (*case/person*) n, w_{ij} adalah skor yang ditetapkan untuk step j dalam suatu item i, sedangkan *difficulty* sebesar δ_i serta *tau* sebesar τ_{ij} adalah karakteristik spesifik tingkat kesukaran item dan tingkat kesukaran kategori j dari item i mengikuti formula yang diajukan Andrich, 1978 (Wright & Masters, 1982: 28-31; Swaminathan, 1999: 50). Item i yang diskor

secara politomus tiga kategori menurut RSM memiliki satu nilai kesukaran item atau *item difficulty* sebesar δ_{ij} dan dua buah nilai parameter *tau* berupa *tau*-1 dan *tau*-2 sebesar τ_{i1} dan τ_{i2}.

Selain menyajikan nilai *tau*, program Quest menyajikan tingkat kesukaran dalam bentuk nilai *threshold* (ambang batas) untuk RSM. Nilai *threshold* yang dihitung berdasarkan nilai *tau* sebagaimana yang diperkenalkan oleh Masters (1988) mengikuti kaidah Thurstone (Adam & Khoo, 1996: 90). Dengan demikian, suatu *threshold* (yang juga diberi simbol τ_{ij} oleh Wright & Master, 1982) merepresentasikan atau keberfungsiannya identik dengan nilai *tau*. Nilai *threshold* untuk suatu tahapan dari item i adalah tingkat kemampuan (*ability level*) yang dibutuhkan oleh testi (*case/person*) untuk melewatinya dengan peluang 0,50 (peluang tertinggi). Nilai *threshold* mengikuti kaidah Thorstone yang diasumsikan bahwa setiap respons tunggal berada pada lokasi μ (baik respons kelompok maupun individual). Proses respons yang melalui teknik penskalaan diubah ke dalam fungsi logistik (Andrich, 1999: 113-114).

Besarnya menjadi w_{ij} sebagai skor yang ditetapkan untuk step j suatu item i sesuai dengan banyaknya kategori, dan besarnya menjadi w_{ij} adalah 0, 1, 2, ... m. Bila besarnya menjadi w_{ij} = 0, maka persamaannya dapat dituliskan dalam ekspresi tunggal dalam formula untuk menjadi *w*:

$$\exp \sum_{j=0}^{0} w_{ij}(\beta_n - \delta_i - \tau_{ij}) \equiv 1 \qquad (11)$$

Untuk kepastian identifikasi digunakan dua *constraint*, yakni:

$$\sum_{j=0}^{ni} \tau_{ij} \equiv 0 \text{ dan } \sum_{i=1}^{l} \delta_i \equiv 0.$$

Besarnya δ_i ditambah τ_{ij} pada RSM sama dengan δ_{ij} pada *Partial Credit Model* (PCM) (Wright & Masters, 1982: 56; Swaminathan, 1999: 51). Dengan demikian, persamaan untuk RSM dapat dituliskan menjadi persamaan untuk PCM sebagai berikut.

$$P(X_{ni} = x_{ni}) = \frac{\exp \sum_{j=0}^{x_n} w_{ij}(\beta_n - \delta_{ij})}{\sum_{k=0}^{ni} \exp \sum_{j=0}^{k} w_{ij}(\beta_n - \delta_{ij})} \quad (12)$$

Penerapan pada data pengukuran yang skor secara dikotomus akan direduksi sehingga formula Rasch *Model* (RM) atau disebut model parameter logistik dalam program Quest dituliskan dengan persamaan:

$$P(X_{ni} = x_{ni}) = \frac{\exp(x_{ni} w_{ij}(\beta_n - \delta_i))}{1 + \exp(x_{ni} w_{ij}(\beta_n - \delta_i))} \quad (13)$$

1. Perhitungan Estimasi untuk Item

Penetapan *fit* item secara keseluruhan dengan model dalam program Quest (Adam & Kho, 1996) didasarkan pada besarnya nilai rata-rata INFIT *Mean of Square* (INFIT MNSQ) beserta simpangan bakunya atau nilai rata-rata INFIT *Mean of* INFIT *t*. Penetapan *fit* tiap item dengan model dalam

program Quest didasarkan pada besarnya nilai INFIT MNSQ atau nilai INFIT *t* item yang bersangkutan.

Besarnya kuadrat tengah yang tertimbang (*Wighted Mean Square*)—dalam program Quest disingkat INFIT MNSQ)—adalah dengan nilai harapan/ekspektasi sebesar 1,0 dan varians sebesar 0,0. Sementara besarnya kuadrat tengah tertimbang terstandar (*Standardized Weighted Mean Square*) atau t_i dengan nilai harapan ekspektasi sebesar 0,0 dan varians sebesar 1,0.

2. Perhitungan Estimasi untuk Testi

Penetapan *fit* testi (*case/person*) secara keseluruhan dengan model dalam program Quest (Adam & Kho, 1996) juga didasarkan pada besarnya nilai rata-rata INFIT *Mean of Square* (INFIT MNSQ) beserta simpangan bakunya. Dapat pula didasarkan pada besarnya nilai rata-rata INFIT *Mean of* INFIT *t*. Penetapan *fit* tiap testi (*case/person*) dengan model dalam program Quest didasarkan pada besarnya nilai INFIT MNSQ atau nilai INFIT *t* item yang bersangkutan. Langkah perhitungannya sebagai berikut (Wright & Masters, 1982: 108-109).

Besarnya jumlah kuadrat tertimbang (*Weighted Sum of Square*) untuk setiap testi (*case/person*) dengan nilai harapan ekspektasi INFIT MNSQ sebesar 1,0 dan varians sebesar 0,0 atau menggunakan kriteria menurut besarnya kuadrat tengah tertimbang terstandar (*Standardized Weighted Mean Square*) atau t_i dengan nilai harapan ekspektasi sebesar 0,0 dan varians sebesar 1,0.

3. Pengujian Validitas untuk Mengetahui *Fit* Item dan Testi terhadap Model

Item characteritic curve (ICC) akan membentuk kurva yang mendatar (*flat*) bila besarnya INFIT MNSQ untuk item atau *e* lebih besar dari satuan logit 1,30 atau lebih kecil dari satuan logit 0,77 dengan rata-rata 1,0. Bila nilainya >1,30 akibatnya membentuk kurva platikurtis (kurva yang terlalu tumpul) dan bila <0,77 akan terlalu kurva leptokurtis (kurva yang terlalu runcing) (Keeves & Alagumalai 1999: 36). Oleh karena itu, dalam program Quest ditetapkan bahwa suatu item atau testi/*case*/*person* dinyatakan *fit* dengan model dengan batas kisaran INFIT MNSQ dari 0,77 sampai 1,30 (Adam & Khoo, 1996:30 & 90). Ada pula peneliti yang menggunakan batas yang lebih ketat, yakni dengan kisaran 0,83 sampai dengan 1,20 dan ada yang menggunakan pengujian berdasarkan besarnya nilai INFIT t. Dalam hal ini menggunakan kisaran nilai t adalah ± 2,0 (pembulatan nilai t sebesar ± 1,96 jika taraf kesalahan atau *alpha* sebesar 5% (Keeves & Alagumalai 1999: 34-36; Bond & Fox, 2007: 43). Dengan demikian, suatu item menjadi tidak *fit* menurut Model Rasch bila memiliki nilai <-2,0 atau > +2,0 (*probability* atau peluang <0,05).

4. Estimasi Reliabilitas

Wright & Masters (1982) menyatakan "The traditional formulation of test "reliability" can be derived from a "true score" model which assumes that the observed test score of each person can be resolved into two components: an unknownable true score and a random error. Test reliability is defined as the proportion of a

sample's observed score variance SD^2 which is due to the sample's true score variance ST^2. $R = ST^2/SD^2 = 1 - (SE^2/SD^2)$ where the observed variance is partitioned into two components $SD^2 = ST^2 + SE^2$, and SE^2 is the error variance of the test, averaged over that sample". The magnitude of the traditional reliability coefficient depends not only on the average test error variance SE^2 which is intended to describe how precisely the test measures (i.e. for a given ST^2, the greater the precision of measurement, the smaller SE^2 and the larger R), but also on the sample true score variance.ST^2 which is intended to describe the "true" ability dispersion of the 'true" sample (i.e-, for a given SE^2, the greater the sample true score variance ST^2, the larger R).'

Estimasi reliabilitas tes menurut IRT dihitung berdasarkan item dan disebut indeks sparasi item. Jika dihitung berdasarkan testi (*case/person*) dan disebut dengan indeks sparasi person. Semakin tinggi estimasi indeks sparasi item semakin tepat/*fit* keseluruhan item dianalisis menurut model yang digunakan (apakah menurut RM, PCM, atau RSM). Semakin tinggi indeks sparasi *person* semakin konsisten setiap item pengukur digunakan untuk mengukur testi yang bersangkutan. Estimasi reliabilitas berdasarkan testi (*case/person*) sama kedudukannya dengan reliabilitas menurut CTT—yakni reliabilitas menurut alpha Cronbach untuk data politomus dan reliabilitas menurut Kuder-Richardson-20 untuk data dikotomus. Indeks separasi item (*item separation index* atau R_I) oleh Wright & Master (1999: 96) disebut dengan istilah "reliabilitas sampel", sedangkan indeks sparasi person disebut dengan "reliabilitas tes".

Program Quest juga menyajikan hasil reliabilitas tes menurut CTT, yakni berupa indeks konsistensi internal, yang untuk penskoran politomus merupakan indeks alpha Cronbach dan untuk penskoran dikotomus merupakan indeks KR-20 dan indeks KR-21 (Adam & Khoo, 1996: 93). Dalam hal ini, reliabilitas yang berlaku adalah untuk tes yang berfungsi seleksi, bukan untuk pengukuran prestasi. Untuk pengukuran prestasi maka perlu dikonversi menjadi indeks Kappa atau indeks persetujuan (*agreement index*).

Selain program Quest juga tersedia program Q yang dikeluarkan oleh A.C.E.R. Hasil analisis menggunakan program Q dapat digunakan sebagai pembanding untuk melihat sajian hasil analisis antara program QUSET dan program Q.

DAFTAR PUSTAKA

Adams, R.J. & Kho, Seik-Tom. (1996). *Acer quest version 2.1*. Camberwell, Victoria: The Australian Council for Educational Research.

Anderson, L. W. & Krathwohl, D. R., et al (Eds..) (2001) *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*. Allyn & Bacon. Boston, MA (Pearson Education Group)

Andrich, D. (1999). Rating scale analysis. In: Masters, G.N. & Keeves, J.P. (1999). *Advances in measurement in educational research and assessment.* Amsterdam: Pergamon, An imprint of Elsevier Science.

Atherton, (2005) diakses 03 Desember 2006 dari http://www.learningandteaching.info/learning/converge.htm.

Atkinson, R.L., Atkinson, R.C., & Hilgard, E.R. (1983). *Introduction to psychology.* Terjemahan Nurdjannah Taufik dan Rukmini Barhana. 1997. Jakarta: Erlangga.

Badan Standar Nasional Pendidikan. (2007). *Panduan penilaian kelompok mata pelajaran IPTEK*. Jakarta: Badan Standar Nasional Pendidikan, Depdiknas.

Bond, T.G. & Fox, Ch.M. (2007). *Applying the rasch model*: *Fundamental measurement in the human sciences*. 2-nd ed. Mahwah, New Jersey: Lawrence Erlbaum Associates, Publishers.

Bowell, T. & Kemp, G. (2002). *Critical thinking*: *A concise guide*. London: Routledge.

Brown, D.H. (2004). *Language assessment: Principle and classroom practice*. San Fransisco: Pearson Education, Inc.

Brum & McKane. (1989). *Study guide biology: Exploring life*. New York: John Wiley & Sons.

Bryce, T.G.K., McCall, J., MacGregor, J., Robertson, I.J., dan Weston, R.A.J. 1990. *Techniques for assessing process skills in practical science: Teacher's guide*. Oxford: Heinemann Educational Books.

Carin, A.A. dan Sund, R.B. 1989. *Teaching Science Through Discovery*. Columbus: Merrill Publishing Company.

Collete, A.T. & Chiappetta, EL. 1994. *Science Instruction in the Middle and Secondary Scholls* (Edisi ke-3). New York: Macmillan Publishing Company.

Conny R. Semiawan (1997). *Perpektif Pendidikan Anak Berbakat*. Jakarta: Gramedia Widiasarana Indonesia.

Cromie, W.J. 2007. *Creativity Tied to Mental Illness*: *Irrelevance Can Make You Mad*, (Online), (http://www.news.harvard.edu/gazette/...reativity.html, diakses 29 Januari 2009).

Dettmer, P. (2006). New Blooms in Established Fields: Four Domains of Learning and Doing [Versi elektronik]. *Roeper Review*, 28, 2, 70-78.

Djemari Mardapi. (2007). *Teknik penyusunan instrumen tes dan non tes*. Yogyakarta: Mitra Cendekia Press.

Djohar. (1989). *Dimensi pendidikan sains menyongsong tahun 2000*. Pidato pengukuhan sebagai guru besar dalam bidang pendidikan sains pada FPMIPA IKIP YOGYAKARTA. Diucapkan pada rapat senat terbuka IKIP YOGYAKARTA 11-03-1989.

Ferrando, M., Prieto, M.D., Ferrandiz, C. & Sanchesz, C. 2005. Intelligence and Creativity. *Electronic Journal of Reseacrch in Education*, ISSN: 1696-2095, 7, 3(3): 21-50, (Online, diakses 29 Januari 2009).

Frisbie, D.A. (2005). Measurement 101: Some fundamentals revisited. *Educational Measurement: Issues and Practice*. Fall 2005. Vol. 24. No.3. pp.21 28.

Garry, R. (1970). *The nature and conditions of learning*. 3-rd ed. Englewood Cliffs, New Jersey: Prentice-Hall, Inc.

Gronlund, N.E. (1998). *Assessment of student achievement*, 9-th ed. Boston: Allyn and Bacon.

Hambleton, R.K. & Swaminathan, H. (1985). *Item response theory*. Boston: Kluwer Nijjhoff Publiser.

Hambleton, R.K., Swaminathan, H., & Rogers, H.J. (1991). *Foundamentals of item responses theory*. Newbury Park: Sage Publications.

Han, Kyung T. & Hambleton, R.K. (2007). *User's manual for WinGen2*: *Windows software that generates IRT model parameters and item response*. (Media elektronik]. Massachusetts: Center for Educational Assessment.

Han, Kyung T. & Hambleton, R.K. (2007). *User's manual for WinGen2*: *Windows software that generates IRT model parameters and item response*. (Media elektronik]. Massachusetts: Center for Educational Assessment.

Hibbard, K.M. (t.t.). *Performance assessment in the science classroom*. New York: McGraw-Hill Companies.

Hopkins, K.D & Stanley, J.C. 1981. *Educational and psychological measurement and evaluastion*. New Jersey: Prentice Hall, Inc.

Huitt, W., & Hummel, J. (2003). Piaget's theory of cognitive development. Educational Psychology Interactive. Valdosta, GA: Valdosta State University. Diakses dari http://www.edpsycinteractive.org/topics/cognition/piaget.html

Johnson. E.B. (2002). *Contextual teaching and learning: What it is and why it's here to stay*. Thousand Oaks, California: Corwin Press, Inc.

Keeves, J.P. & Masters, G.N. (1999). Introduction. In: Masters, G.N. & Keeves, J.P. (1999). *Advances in measurement in educational research and assessment.* Amasterdam: Pergamon, An imprint of Elsevier Science.

Kolen, M.J. & Brennan, R.L. (1995). *Test equating: Methods and practices*. New York: Springer-Verlag New York Inc.

Kolen, M.J. (1999). Equating of test. In: Masters, G.N. & Keeves, J.P. (1999). *Advances in measurement in educational research and assessment.* Amsterdam: Pergamon, An imprint of Elsevier Science.

Kunnan, A.J. 1998. *Validation in language assessment*. New Jersey: Lawrence Elbaum Associates, Publisher.

Masters, G.N. (1999). Partial credit model. In: Masters, G.N. & Keeves, J.P. (1999). *Advances in measurement in educational research and assessment.* Amsterdam: Pergamon, An imprint of Elsevier Science.

Messick, S. (1988). The one and future issues of validity: Assessing the meanng and consequences of measurement. In: Waine, H. & Braun, H.I. (1988). *Test validity*. Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.

Moh Amien (1980). *Peranan kreativitas dalam pendidikan*. Pidato Dies IKIP Yogyakarta, diucapkan di muka sidang senat terbuka IKIP Yogyakarta, tanggal 21 Mei 1980.

Muijs, D. & Reynolds, D. (2008). *Effective teaching: Teori dan aplikasi*. (Terjemahan Helly Prajitno Soetjipto & Sri Mulyantini Soecipta). London: Sage Publications Ltd. (Buku asli diterbitkan tahun 2008).

Muraki, E. & Bock, R.D. (1998) Parscale: *IRT item analysis and test scoring for rating scale data.* Chicago: Scientific Software Internatinal, Inc.

Pollman, J., Uprichard, E., Malone, U., & Coop, R. 1973. *Multivariate Analysis of The Relationship Between Creativity and Intelllegence*. Paper presented at annual meeting of American Educational Reserach Association, New Orleans, Lousiana, February 25-March 1, 1973.

Rezba, R.J., Sparague, C.S., Fiel, R.L., Funk, H.J., Okey, J.R., & Jaus, H.H. (1995). *Learning and assessing science process skills*. 3rd ed. Iowa: Kendall/Hunt Publishing Company.

Smith, J.K. (2003). Reconsidering reliability in classroom assessment and grading [Versi elektronik]. *Educational Measurement, Issues and Practice*, 22, 4, 26-33.

Stark, S., Chernyshenko, S., Chuah, D.,Wayne Lee, & Wilington, P. (2001a). *IRT modeling lab: IRT tutorial* [Versi elektronik]. Urbana: University of Illinois.

Stark, S., Chernyshenko, S., Chuah, D.,Wayne Lee, & Wilington, P. (2001b). *IRT modeling lab: Test Development Using Classical Test Theory* [Versi elektronik]. Urbana: University of Illinois.

Torrance, E.P. (1979). Three stage model for teaching for creative thinking. Dalam: Lawson, A.E. *The psychology of teaching for thinking and creativity*. Columbus: ERIC.

Towle, A. 1989. *Modern biology*. Austin: Holt, Rinehart and Winston.

Wright, B.D. & Masters, G.N. (1982). *Rating scale analsis*. Chicago: Mesa Press.

Wright, B.D. (1999). Rasch measurement model. In: Masters, G.N. & Keeves, J.P. (1999). *Advances in measurement in educational research and assessment.* Amasterdam: Pergamon, An imprint of Elsevier Science.

LAMPIRAN-LAMPIRAN

Lampiran 1

Lampiran 1. Rumusan *Learning Continuum* Keterampilan Proses Sains Oleh Peneliti sebagai *Abstract Continuum* Pengukuran Keterampilan Proses Sains untuk Pengembangan Instrumen Tes

I. Keterampilan dasar (*basic skills*) terdiri atas: (a) ketrampilan mengamati, (b) keterampilan merekam data/informasi (c) keterampilan mengikuti instruksi, (d) keterampilan melakukan pengukuran, (e) keterampilan memanipulasi gerakan, dan (f) keterampilan mengimplemetasikan prosedur/teknik/ penggunaan peralatan

A. Ketrampilan melakukan pengamatan

1. Mengidentifikasi substansi berwarna berdasarkan deskripsinya atau mencocokkan dengan warna rujukan
2. Mengamati dan menggambarkan perubahan warna, bentuk, atau tingkatan

3. Mencocokkan objek yang serupa atau yang representatif dan mengidentifikasi perbedaan-perbedaan secara lebih terperinci
4. Mencocokkan suatu objek dengan beragam representasi visualnya
5. Mengidentifikasi situasi-situasi berpotensi penuh risiko di sekolah (kerja laboratorium) dengan yang ada di dalam kehidupan sehari-hari
6. Mengidentifikasi peninggalan kuno yang memiliki nilai ilmiah atau dampak teknologi di alam atau di suatu areal atau di dalam foto
7. Mendemonstrasikan kegiatan mendengarkan secara hati-hati (mengidentifikasi suara dengan cermat)

B. Keterampilan merekam data/informasi

1. Merekam informasi sederhana dengan mempresentasikannya dalam bermacam bentuk sesuai klasifikasinya
2. Membuat suatu ringkasan dalam bentuk tertulis ataupun secara lisan
3. Memberi label suatu bagan/diagram secara benar
4. Menggambar bentuk-bentuk sederhana dengan akurat
5. Menyusun informasi dalam bentuk tabel disertai dengan judul tabel
6. Menghasilkan tema yang tepat beserta informasi yang ada di dalamnya
7. Melengkapi suatu bagan/carta, grafik, atau histogram
8. Mengonstruksi suatu bagan/carta, grafik, atau histogram sesuai data

9. Menyampaikan data dalam bentuk yang sederhana dan menyajikan dalam bentuk suatu bagan/carta, grafik, atau histogram

C. Keterampilan mengikuti instruksi

1. Menyiapkan peralatan atau menyelesaikan suatu prosedur setelah guru mendemonstrasikannya
2. Menyiapkan peralatan atau menyelesaikan suatu prosedur menurut perintah yang disampaikan secara lisan
3. Menyiapkan peralatan atau menyelesaikan suatu prosedur menggunakan kartu petunjuk kerja yang berisi informasi dalam bentuk yang sebagian besar berupa gambar.
4. Menyiapkan peralatan atau menyelesaikan suatu prosedur menggunakan kartu petunjuk kerja yang berisi informasi berupa tulisan

D. Keterampilan melakukan pengukuran

1. Menemukenali atau menyediakan unit-unit/alat pengukur yang benar untuk pengukuran-pengukuran pada umumnya
2. Membaca skala meteran/pita ukur dan melakukan pengukuran dengan tepat
3. Membaca skala termometer laboratorium dan termometer klinis dan melakukan pengukuran dengan tepat
4. Membaca skala alat penimbang atau bejana ukur dan melakukan pengukuran dengan tepat

5. Membaca bermacam skala pada posisi naik-turun dan melakukan pengukuran dengan tepat
6. Membaca berbagai skala yang dilakukan dengan cara ditekan (*dial scale*) dan melakukan pengukuran dengan tepat
7. Mengestimasi secara kasar berat relatif, volume, dan interval waktu jangka pendek
8. Menggunakan kisi-kisi/jaring (*grid*) untuk mengestimasi luas suatu area/permukaan benda

E. Keterampilan melakukan manipulasi gerakan

1. Mendemonstrasikan kendali motorik kasar dalam berbagai cara
2. Mendemonstrasikan kendali motorik halus dalam berbagai cara (seperti menyetel/mengatur, melengkapi/mewarnai bagan, memotong/menggunting, merakit bagian yang terpisah)

F. Keterampilan mengimplemetasikan prosedur / teknik / penggunaan peralatan

1. Mengidentifikasi peralatan laboratorium yang biasanya digunakan
2. Memilih peralatan laboratorium yang sesuai dengan tugas yang dikerjakan
3. Mengadopsi prosedur laboratorium dengan memperkecil/meminimalkan risiko/kesalahan
4. Memindahkan suatu material padat ataupun cairan dengan menggunakan tempat yang spesifik
5. Menggunakan sendok ukur untuk menghasilkan suatu volume tertentu yang diperlukan

6. Menggunakan pipet tetes atau alat penyemprot untuk memindahkan suatu cairan
7. Mengombinasikan/menyampur sejumlah kecil bahan untuk menghasilkan campuran/ ramuan sesuai dengan standar yang ditetapkan/diberikan
8. Menghancurkan material padat menggunakan prosedur/metode yang tepat/sesuai
9. Melakukan teknik penyaringan dengan benar
10. Menggunakan kertas uji atau bahan kimia cair uji dengan benar
11. Menjaga keselamatan kerja jika bekerja:
    a. menggunakan lampu sepiritus (lampu Bunsen) atau menggunakan api sesuai keperluan
    b. memanaskan suatu material padat menggunakan teknik yang sesuai
    c. memanaskan suatu material cair menggunakan tabung reaksi atau wadah yang sesuai
    d. memanaskan bahan yang mudah terbakar menggunakan gelas beker berisi air
12. Menggunakan termometer laboratorium atau termometer klinis
13. Menggunakan lensa/kaca pembesar
14. menggunakan mikroskop
15. menggunakan *stop-clock* atau *stop-watch*
    a. untuk mengukur suatu selang waktu tertentu secara teliti
    b. untuk membaca pada interval waktu tertentu yang telah ditetapkan.
16. Mengeset suatu rangkaian elektrik sederhana sesuai bagan rangkaian yang ditetapkan

17. Menyiapkan preparat untuk diamati di bawah mikroskop
18. Menggunakan berbagai instrumen untuk kerja lapangan (seperti pengukur cahaya, pH meter)
19. Menggunakan berbagai teknik-teknik kerja di lapangan (seperti bekerja menggunakan kuadrat, kerja menggunakan jaring)

II. Keterampilan mengolah/memroses (*process skills*) terdiri atas: (a) keterampilan membuat prediksi, (b) keterampilan membuat inferensi, dan (c) keterampilan menyeleksi prosedur

A. Keterampilan membuat prediksi

1. Membuat dugaan kondisi objek ke depan berdasarkan fakta yang dihadapi
2. Membuat prakiraan perubahan ciri morfologi makhluk hidup berdasarkan kondisi yang dikenakan
3. Membuat prakiraan perubahan ciri anatomi makhluk hidup berdasarkan kondisi yang dikenakan
4. Membuat prakiraan perubahan ciri fisiologi makhluk hidup berdasarkan kondisi yang dikenakan

B. Keterampilan membuat inferensi

1. Membedakan antara hasil observasi dari rujukannya
2. Membuat generalisasi yang masuk akal berdasarkan hasil observasi
3. Menggunakan generalisasi dan hasil observasi untuk menarik kesimpulan yang sahih

4. Mengombinasikan observasi dan informasi yang diberikan untuk merumuskan hipotesis
5. Membuat deduksi dari hipotesis
6. Menggunakan hasil observasi untuk mengonfirmasikan atau membuktikan kesalahan/menyangkal hipotesis yang ada
7. Memodifikasi hipotesis untuk mengakomodasi observasi-observasi baru

C. Keterampilan menyeleksi prosedur

1. Mengantisipasi risiko-risiko dan mengadopsi tindakan pencegahan dan prosedur yang tepat/sesuai dalam melakukan setiap percobaan laboratorium atau investigasi praktis.
2. Mengidentifikasi suatu prosedur yang sesuai (seperti. pengamatan, pencacahan, pengukuran—mengukur panjang, berat, suhu, volume, dan waktu—pengambilan sampel, pemisahan, penghancuran, pemanasan, menggunakan indikator kimia) dan memilih peralatan/aparatus yang tepat sesuai dengan permasalahan praktis yang dihadapi
3. Memilih komponen peralatan yang sesuai untuk menghasilkan pengukuran yang teliti/akurat
4. Memilih variabel-variabel yang sesuai, mengumpulkan data yang relevan, dan memilih suatu bentuk sajian hasil yang sesuai untuk suatu prosedur investigasi yang telah dipilih,

III. Keterampilan menginvestigasi (*investigative skills*) terdiri atas: (a) keterampilan merancang investigasi, (b)

keterampilan melaksanakan investigasi, dan (c) keterampilan melaporkan hasil investigasi

A. Keterampilan merancang investigasi

1. Menetapkan variabel bebas sebagai dalam investigasi termasuk merancang faktor perlakuan dalam eksperimen
2. Mengidentifikasi hubungan antara variabel bebas dan variabel tergayut dalam suatu investigasi
3. Mengemukakan latar belakang pentingnya suatu investigasi harus dilakukan
4. Merumuskan tujuan investigasi
5. Merumuskan manfaat suatu investigasi
6. Menetapkan kelompok perlakuan dan kelompok kontrol dalam status eksperimen
7. Merancang pengendalian variabel penekan/pengganggu (*suppressed variable*) menjadi variabel kendali dalam suatu eksperimen
8. Mengidentifikasi variabel random dalam suatu investigasi
9. Menetapkan hipotesis penelitian dalam suatu investigasi
10. Merancang sampel yang diperlukan dalam investigasi
11. Menetapkan prosedur pengumpulan data investigasi
12. Merancang jenis data yang harus dihimpun dalam investigasi
13. Merancang penyajian data hasil investigasi
14. Merancang teknik analisis data hasil investigasi

B. Keterampilan melaksanakan investigasi, secara simultan: (1)menyusun rangkaian peralatan investigasi

/eksperimen sesuai rancangan, (2) mengatur perlakuan dalam eksperimen sesuai rancangan, (3) menangani variabel penekan/pengganggu atau variabel asing sesuai rancangan, (4) melakukan pengukuran untuk menghimpun data sesuai rancangan, (5) mencatat seluruh data eskperimen sesuai rancangan, (6) menjaga kestabilan kondisi lingkungan jika melakukan eksperimen, (7) mengamati dan mencatat data yang menunjukkan penyimpangan bila melakukan eksperimen.

C. Keterampilan melaporkan hasil investigasi

1. Menyajikan hasil penelitian dalam bentuk diagram, tabel ataupun grafik
2. Menyajikan kesimpulan data hasil penelitian
3. Menyajikan pembahasan suatu hasil investigasi
4. Menetapkan kebenaran hasil investigasi
5. Menarik kesimpulan umum hasil investigasi berdasarkan hasil observasi dan generalisasi data (termasuk hasil analisis statistika)
6. Menuangkan hasil penelitian, pembahasan hasil penelitian, dan kesimpulan penelitian dalam suatu laporan secara tertulis
7. Memaparkan hasil penelitian, pembahasan hasil penelitian, dan kesimpulan penelitian dalam suatu laporan secara lisan dalam suatu seminar/presentasi kelas

Contoh kisi-kisi dan Item Instrumen Tes Penguasaan Metode Ilmiah Tahun 2017

Tabel 1. Kisi-kisi dan Soal beserta Kunci untuk Kelompok Indikator I-III

Kode	Indikator	Pilihan dua alternatif	Pilihan tiga alternatif
1.1.1	Memilih dan mencocokkan sendiri objek berupa makhluk hidup dengan gambarnya.	1. Jika kamu melihat dirimu dan fotomu maka yang sama persis antara lain .... a. warna rambutmu* b. besar tubuhmu	1. Jika kamu melihat ayam dan fotonya maka yang sama persis antara lain .... a. warnanya* b. besarnya c. panjangnya
1.1.2	Mengidentifik asi/mengenali sendiri dampak teknologi di alam,di suatu areal, atau didalam gambar foto.	2. Selain rasanya, perbedaan antara semangka unggul dengan semangka biasa adalah... a. banyaknya biji* b. warna biji	2. Apa perbedaan yang terdapat pada jambu biji unggul dengan jambu biji biasa? a. ukuran buah* b. warna daun c. warna biji
1.1.3	Memilih/mengid entifikasi/ Mengenali sendiri nama/jenis hewan berdasarkan suara hewan yang didengar.	3. Suara "kukuruyuuuuk" yang terdengar dipagi hari adalah suara .... a. ayam* b. itik	3. Suara "guk guk guk guuk" yang terdengan di malam hari adalah suara .... a. kambing b. anjing* c. angsa

Kode	Indikator	Pilihan dua alternatif	Pilihan tiga alternatif
1.2.1	Melengkapi sendiri suatu bagan/carta, grafik atau histogram tentang fenomena kehidupan makhluk hidup.	4. Ikan ----- sungai Sapi ----- darat Mana yang sesuai dengan bagan di atas? a. Itik manila ----- kolam b. Kera ----- pepohonan*	4. Kucing-----pemakan daging kambing ----- pemakan rumput yang sesuai dengan bagan di atas yaitu.... a. burung merpati ----- pemakan biji-bijian* b. ular ----- pemakan dedaunan c. kerbau ----- pemakan buah
1.3.1	Menyiapkan sendiri peralatan atau melakukan/melanjutkan sendiri langkah yang harus ditempuh dalam melakukan pengamatan gejala kehidupan makhluk hidup setelah guru menjelaskan secara lisan.	5. Setelah guru menjelaskan cara-cara melakukan pengamatan terhadap tinggi tanaman secara lisan maka apa langkah yang harus kamu lakukan selanjutnya? a. Memilih alat yang harus digunakan untuk mengamati* b. Segera melakukan pengukuran terhadap tinggi tanaman	5. Setelah guru menjelaskan cara-cara melakukan pengamatan terhadap tinggi tanaman secara lisan maka apa langkah yang harus kamu lakukan selanjutnya? a. melakukan pengukuran terhadap tinggi tanaman b. memilih alat yang harus digunakan untuk mengamati* c. mencatat hasil pengamatan tinggi tanaman

Kode	Indikator	Pilihan dua alternatif	Pilihan tiga alternatif
1.4.1	Menentukan sendiri dasar untuk memisahkan makhluk hidup (hewan dan tumbuhan) atau bagian tubuhnya yang memiliki ciri yang berbeda dan menyatukannya berdasarkan ciri yang sama menurut hasil pengamatan.	6. Selain sama-sama berguna untuk berjalan, salah satu perbedaan yang tedapat pada kaki kambing dengan kaki sapi adalah .... a. jumlah kaki b. besar kaki*	6. Perbedaan yang menyolok antara daun singkong dan daun jarak adalah pada ..... a. bentuknya* b. pertulanganya c. warnanya
1.5.1	Menemukenali atau menyediakan unit-unit/alat pengukur yang benar untuk pengukuran-pengukuran pada umumnya	7. Jika kamu hendak mengukur suhu tubuh temanmu, alat yang kamu gunakan adalah..... a. termometer* b. tensimeter	7. Di bawah ini yang bukan merupakan kegunaan penggaris adalah dapat untuk alat pengukur .... a. tinggi badan b. panjang lengan c. berat badan*
1.5.2	Menemukan sendiri penyebab ketidaktepatan dalam membaca skala meteran atau pita ukur ketika melakukan pengukuran ciri tubuh makhluk hidup.	8. Selain kelainan pada mata, penyebab ketidaktepatan membaca skala pada pita ukur ketika kamu mengukur panjang lengan tangan temanmu adalah... a. cahaya lampu di dalam ruangan yang terang b. mata tidak sejajar dengan skala pada pita ukur*	8. Dibawah ini, apa penyebab utama terjadinya <u>ketidaktepatan</u> hasil mengukur lebar daun suatu tanaman? a. angka skala pada pita ukur terlalu kecil b. cepat dalam membaca angka skala pita ukur c. kepala bergerak saat melihatangka skala pita ukur*

Kode	Indikator	Pilihan dua alternatif	Pilihan tiga alternatif
1.5.3	Menemukan sendiri penyebab ketidaktepatan dalam melakukan pengukuran suhu tubuh makhluk hidup menggunakan termometer digital.	9. Ketika kamu mengukur suhu tubuh temanmu yang sedang sakit menggunakan termometer digital, ternyata hasil yang didapat salah. Hal ini dapatdisebabkan oleh…. a. Keadaan temanmu yang demam akibat sakit b. Tidak cermat melihat angka skala termometer*	9. Ketika kamu diminta untuk mengukur suhu tubuh adikmu, kamu tidak dapat melihat angka pada termometer digital yang digunakan. Hal ini disebabkan oleh…. a. baterai termometer digital habis* b. temanmu yang ada mengganggum c. tubuh adikmu dalam keadaan demam
1.6.1	Menemukan sendiri cara menggunakan keterampilan tangan untuk membuat karya yang berhubungan dengan gejala kehidupan makhluk hidup.	10. Untuk dapat membuat herbarium atau awetan daun yang dikeringkan maka kamu perlu …. a. Menjaga agar helaian daun tidak terlipat* b. Memisahkan daun yang telah kering	10. Untuk dapat membuat herbarium atau awetan daun yang dikeringkan maka kamu perlu …. a. menjaga agar helaian daun tidak terlipat* b. memisahkan daun yang telah kering c. memetik daun yang di ujung ranting

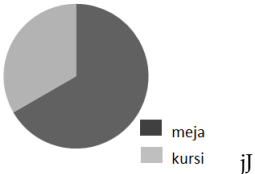
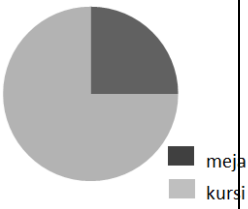
Kode	Indikator	Pilihan dua alternatif	Pilihan tiga alternatif
1.6.2	Mencari sendiri cara memindahkan benda padat atau cair dalam suatu percobaan yang berhubungan dengan gejala kehidupan makhluk hidup.	11. Jika kamu diminta menangkap ikan dari kolam yang akan dipindahkan ke dalam aquarium maka kamu dapat melakukan dengan menggunakan… a. jaring* b. panic	11. Jika kamu diminta memindahkan tanaman dari pot ke tanah yang sudah disiapkan lubangnya maka yang tepat adalah …. a. membiarkan akar bersama tanahnya dan menanamnya* b. mencabut tanamannya kemudian menanamnya c. menanam tanamannya bersama-sama dengan potnya
1.6.3	Mencari sendiri hal yang menyebabkan kesalahan dalam menghaluskan bahan yang diperoleh dari makhluk hidup.	12. Jika kamu diminta memotong rumput menjadi potongan yang kecil-kecil untuk pakan ternak maka dapat terjadi kegagalan akibat …. a. pisau kurang tajam* b. rumput terlalu hijau	12. Jika kamu diminta memotong rumput menjadi potongan yang kecil-kecil untuk pakan ternak maka dapat terjadi kegagalan akibat …. a. rumput cukup banyak b. pisau kurang tajam* c. rumput terlalu hijau

Kode	Indikator	Pilihan dua alternatif	Pilihan tiga alternatif
1.6.4	Menemukan sendiri hal yang menyebabkan kesalahan dalam menggunakan termometer untuk mengukur dan mengamati gejala kehidupan makhluk hidup.	13. Ketika kamu menggunakan tensimeter mengukur tekanan darah temanmu yang sedang sakit ternyata hasilnya berbeda jika dipasang pada pergelangan dibandingkan pada lengan tangan, hal ini menunjukkan .... a. kurang lama memasang tensimeter b. salah tempat memasang tensimeter*	13. Ketika kamu menggunakan termometer untuk mengukur suhu tubuh temanmu ternyata hasilnya berbeda dengan hasil pengukuran yang dilakukan guru. Hal ini disebabkan .... a. saat melihat angka diulang b. termometer menjadi rusak c. salah memasang termometer*
2.1.1	Menemukan sendiri perbedaan antara bentuk organ tubuh makhluk hidup yang ada digambar	14. Jika kamu mengamati gambar ayam jantan dan betina maka perbedaan keduanya terletak pada panjang bulu .... a. ekor* b. sayap	14. Jika kamu mengamati foto bunga dan foto buah suatu tanaman maka kemungkinan adanya perbedaannya terletak pada .... a. bentuk tangkainya b. panjang tangkainya* c. warna tangkainya
2.1.2	Merumuskan sendiri kesimpulan berdasarkan data pengamatan terhadap makhluk hidup.	15. Ketika kamu mengamati berbagai jenis hewan seperti kadal, cicak dan buaya, ciri manakah yang menyebabkan ketiga hewan tersebut termasuk dalam kelompok yang sama? a. <u>tidak</u> berdaun telinga* b. hidup di air	15. Ketika kamu mengamati berbagai jenis hewan seperti sapi, kambing, dan kuda, ciri manakah yang dimiliki ketiganya sehingga termasuk kelompok yang sama? a. Memiliki gigi taring b. memakan rumput* c. memiliki kuku yang tajam

Kode	Indikator	Pilihan dua alternatif	Pilihan tiga alternatif
2.1.3	Merumuskan sendiri dugaan/ prakiraan sementara (hipotesis) berdasarkan hasil pengamatan/ penyelidikan dan informasi yang ada	16. Bila kamu meletakkan pot tanaman di dekat jendela maka tanaman tersebut batangnya akan membengkok .... a. mendekati ke arah jendela* b. menjauh dari arah jendela	16. Bila kamu akan memelihara burung sejak kecil maka .... a. menjadi lebih cepat jinak* b. sulit menjadi jinak c. tetap menjadi burung liar
2.2.1	Memperkirakan sendiri kemungkinan yang terjadi ketika suatu mahluk hidup saat sekarang sedang berada pada kondisi yang tidak mendukung berdasar fakta yang dihadapi. (*anchor*)	17. Apa yang akan terjadi apabila sungai hampir mengering saat musim kemarau? a. Banyak ikan yang akan mati. * b. Banyak ikan yang bertelur.	17. Pada musim kemarau yang panjang, apa yang akan terjadi pada tanaman tomat yang ditanam petani? a. Tanaman tomat tumbuh subur b. daun tanaman tomat akan layu dan menguning* c. daun tanaman tomat akan bertambah banyak
2.2.2	Memperkirakan perubahan proses fisiologis yang terjadi di dalam makhluk hidup bila diberikan kondisi tertentu. (*anchor*)	18. Setelah kamu selesai olahraga apa yang kamu rasakan pada tubuhmu? a. Jantung berdegub lebih cepat.* b. Jantung berdegub lebih lambat.	18. Pada saat kamu pagi hari apa yang kamu rasakan dibanding pada saat siang hari? a. Tubuh mengeluarkan banyak keringat b. tubuh mengeluarkan sedikit keringat a. tubuh <u>tidak</u> berkeringat*

Kode	Indikator	Pilihan dua alternatif	Pilihan tiga alternatif
2.3.1	Menentukan sendiri langkah kerja yang sesuai dalam suatu pengamatan, pengukuran, pemanasan, penghancuran dan menggunakan indikator kimia serta memilih sendiri peralatan yang sesuai dengan permasalahan yang dihadapi dalam kehidupan sehari-hari. (*anchor*)	19. Jika kamu akan membandingkan keliling dua batang pohon tetapi kamu tidak memiliki pita ukur maka akan lebih tepat jika menggunakan…. a. kain yang diberi angka dengan jarak seperti pada pita ukur* b. kayu yang diberi angka dengan jarak seperti pada pita ukur	19. Jika kamu akan menghancurkan butiran jagung untuk memberi pakan anak ayam maka lebih baik kamu menggunakan …. a. pisau b. penumbuk* c. gunting
2.3.2	Memilih sendiri komponen/jenis peralatan yang sesuai untuk mengukur sesuatu dengan hasil yang teliti *(anchor).*	20. Bila kamu akan mengukur detak jantung maka kamu dapat menggunakan .... a. tensimeter b. stetoskop*	20. Bila kamu akan mengukur tekanan darah maka kamu dapat menggunakan .... a. termometer b. tensimeter* c. stetoskop
1.1.1	Memilih sendiri gejala yang akan dibandingkan bila dihadapkan pada dua benda tak hidup untuk mengidentifikasi perbedaan secara terperinci.	21. Jika kamu mengamati sepeda motor dan mobil, apa perbedaan dari keduanya? a. B Bahan bakarnya. b. JJ umlah rodanya.*	21. Jika kamu mengamati bis dan truk, apa yang membedakan keduanya? a. suaranya * b. bentuk rodanya c. bentuknya pintunya

1.2.2	Membuat sendiri ringkasan suatu paragraf/bab/ buku yang mengulas gejala benda tak hidup	22. Jika kamu diperintah untuk meringkas bab mengenai penyebab banjir maka kalimat yang tidak tepat adalah a. Bencana alam berupa banjir dapat disebabkan karena banyak manusia yang membuang sampah disungai b. Membuang sampah di sungai dilakukan untuk mengurangi sampah yang menumpuk di halaman*	22. Jika kamu diperintah untuk meringkas bab mengenai peran air bagi manusia kalimat yang perlu ditulis adalah a. Air tidak akan dapat bercampur dengan minyak b. Air mudah menguap dan mudah pula membeku c. Air dapat dimanfaatkan untuk berbagai keperluan*
1.2.3	Membuat histogram tentang gejala benda tak hidup lengkap dengan labelnya.	23. Jika kamu diminta membuat histogram rumah berdinding tembok dan yang bukan berdinding tembok di suatu kampong maka yang perlu diperhatikan adalah…. a. Banyaknya rumah yang dindingnya bertembok dan banyaknya rumah yang bukan berdinding bertembok.* b. Banyaknya kepemilikan rumah berdinding bertembok dan banyaknya kepemilikan rumah yang berdinding bertembok.	23. Jika kamu diminta untuk membuat histogram jumlah jendela dan pintu yang ada pada suatu gedung sekolah maka yang harus dicantumkan dalam histogram adalah .... a. bentuk jendela dan pintu yang ada b. jumlah jendela dan pintu yang ada* c. lokasi ditemukanya jendela dan pintu

Kode	Indikator	Pilihan dua alternatif	Pilihan tiga alternatif
1.2.4	Menyampaikan sendiri informasi tentang suatu benda tak hidup yang tersaji dalam bentuk bagan/carta, grafik atau histogram.	24. Informasi yang tepat berdasarkan sajian data berikut yaitu ....  a. Jumlah meja dua kali dari jumlah kursi* b. Jumlah meja setengah kali dari jumlah kursi	24. Informasi yang tepat berdasarkan sajian data berikut yaitu ....  a. Jumlah meja tiga kali dari jumlah kursi b. Jumlah meja lebih sedikit dari jumlah kursi c. Jumlah meja sepertiga kali dari jumlah kursi*
1.5.1	Menemukenali atau menyediakan unit-unit/alat pengukur yang benar untuk pengukuran-pengukuran pada umumnya terhadap gejala benda tak hidup	25. Jika kamu diminta untuk mengukur tinggi meja, apa yang sebaiknya kamu gunakan? a. Penggaris berskala.* b. Jengkal tangan.	25. Kamu akan mengukur volume air, untuk memperoleh ukuran yang tepat sebaiknya menggunakan .... a. gelas ukur * b. pita ukur c. pipet kaca

Kode	Indikator	Pilihan dua alternatif	Pilihan tiga alternatif
1.5.4	Menemukan sendiri penyebab ketidaktepatan dalam melakukan pengukuran ciri benda tak hidup menggunakan alat penimbang atau alat pengukur volume. *(anchor)*	26.Jika kamu menimbang berat buku sebanyak dua kali tetapi hasilnya berbeda maka kemungkinan yang terjadi adalah .... a. kKesalahan saat melihat angka pada timbangan* b. Buku yang ditimbang termasuk terlalu ringan	26. Jika kamu dan kemudian temanmu diminta mengukur volume sebuah kerikil menggunakan gelas ukur dan ternyata didapatkan hasil yang berbeda maka kemungkinannya akibat .... a. kaca gelas ukur yang digunakan tergolong tebal b. bentuk batunya <u>tidak bersegi</u> atau <u>tidak beraturan</u> c. salah saat melihat permukaan air pada gelas ukur*

Kode	Indikator	Pilihan dua alternatif	Pilihan tiga alternatif
1.5.8	Menemukan sendiri kesalahan dalam penggunaan kisi- kisi/jaring (*grid*) untuk mengestimasi luas suatu area/permukaan benda/bagian benda tak hidup. (*anchor*)	27. Jika kamu diminta melakukan pengukuran luas gambar rumah menggunakan kertas bergaris kisi-kisi berjarak satu milimeter (*milimeter block*) maka kamu harus membuat blad gambar rumah yang diukur pada kertas tersebut. Jika hasilnya lebih sempit dari luas gambar rumah yang sesungguhnya maka kemungkinannya .... a. kamu kurang cermat menghitung banyaknya kotak milimeter pada kertas kisi-kisi yang tertutup daun* b. ada bagian gambar rumah yang menutup kertas bergaris kisi-kisi berjarak satu millimeter	27. Jika kamu diminta melakukan pengukuran luas gambar mobil menggunakan kertas bergaris kisi-kisi (*milimeter block*) maka kamu harus membuat blad gambar mobil yang diukur pada kertas tersebut dengan tepat. Jika terjadi hasilnya berbeda dengan luas gambar mobil yang diukur maka kemungkinannya .... a. kotak milimeter pada kertas bergaris kisi-kisi (*milimeter block*) terlihat kecil b. ada bagian gambar mobil yang tampak berwarna gelap dan ada yang terang c. kurang cermat dalam menghitung banyaknya kotak milimeter pada blad yang dibuat*

Kode	Indikator	Pilihan dua alternatif	Pilihan tiga alternatif
1.6.2	Menemukan sendiri cara menggunakan keterampilan tangan untuk membuat karya yang berhubungan dengan gejala benda tank hidup.	28. Jika kamu diminta untuk membuat vas bunga dari tanah liat maka cara yang dapat dilakukan adalah .... a. tanah liat dikeringkan di bawah sinar matahari, kemudian dibentuk menjadi vas b. tanah liat dibasahi dengan air, dibentuk menjadi vas baru kemudian dikeringkan*	28. Jika kamu diminta untuk membuat kotak pensil dari kardus bekas maka langkah awal yang perlu dilakukan adalah .... a. membuat pola atau bangun kotak pensil yang akan dibuat* b. memotong dan melipat kardus lalu mengelemnya c. menggunting kardus dan membentuknya menjadi kotak pensil
1.7.2	Mencari sendiri jenis peralatan yang akan digunakan untuk mengukur gejala benda tak hidup sesuai dengan tugas yang diberikan oleh Bapak/Ibu Guru.	29. Jika kamu diminta mengukur lingkar tiang bendera, maka alat ukur yang kamu gunakan adalah .... a. penggaris berskala b. *roll meter* dari plastik*	29. Jika kamu diminta untuk mengukur volume sebuah kerikil maka peralatan yang dibutuhkan adalah ... a. gelas ukur yang muat dimasuki kerikil* b. bejana besar untuk menampung air c. mistar ukur yang akan digunakan

Kode	Indikator	Pilihan dua alternatif	Pilihan tiga alternatif
1.7.3	Mencari sendiri hal yang menyebabkan kurang berfungsinya pipet tetes untuk memindahkan larutan pada percobaan yang berhubungan dengan gejala benda tak hidup.	30. Seandainya kamu diminta memindahkan cairan menggunakan pipet akan terjadi kegagalan karena .... a. karet pipet sudah tidak bisa kembali ketika ditekan* b. ujung pipet terlalu masuk ke dalam botol tempat cairan vitamin	30. Seandainya kamu diminta meneteskan cairan pada suatu benda menggunakan pipet mungkin akan terjadi kegagalan karena .... a. karet pipet terpasang kuat pada bagian pangkal pipet b. ujung pipet terlalu masuk ke dalam botol tempat cairan vitamin c. karet pipet sudah tidak bisa mengembang kembali ketika ditekan*
1.7.4	Mencari sendiri hal yang menyebabkan kurang berfungsinya kertas uji bahan kimia cair pada pengujian bahan yang diperoleh dari benda tak hidup.	31. Ketika kamu hendak mengukur pH cairan sabun menggunakan kertas pH yang sudah basah maka kemungkinan yang terjadi adalah ... a. pH air sabun <u>tidak</u> terukur dengan tepat* b. kertas pH berubah-ubah warnanya	31. Jika kamu diminta mengukur pH air akuarium menggunakan kertas pH dengan dua kali ulangan maka akan terjadi perbedaan hasil apabila .... a. kertas pH belum pernah digunakan b. salah satu kertas pH sudah rusak* c. kertas pH yang digunakan masih baru

Kode	Indikator	Pilihan dua alternatif	Pilihan tiga alternatif
1.7.5	Menemukan sendiri hal yang menyebabkan kesalahan dalam menggunakan *stopclock* atau *stopwatch.*untuk mengukur gejala benda tak hidup.	32. Kesalahan yang terjadi jika kamu menggunakan *stopwatch* untuk mengukur kecepatan arus air sungai dapat terjadi karena …. a. g Gerak jarum pada *stopwatch* terlalu lambat* b. s *Stopwatch* sudah beberapa kali digunakan	32. Ketika kamu mengukur ulang kecepatan gerak bola yang dilepas dari papan setinggi dua meter menggunakan *stopwatch* ternyata ada selisih waktu yang tercatat. Ada beberapa kemungkinan yang menjadi penyebabnya<u>kecu ali</u> .... a. *stopwatch*masi h baru karena baru satu minggu dibeli* b. salah membaca angka yang tertunjuk jarum pada *stopwatch* c. kurang memahami cara menggunakan *stopwatch*

Kode	Indikator	Pilihan dua alternatif	Pilihan tiga alternatif
2.1.1	Merumuskan sendiri kesimpulan yang sesuai dengan data hasil pengamatan/pe nyelidikan yang ada terhadap benda tak hidup dan informasi yang ada dalam buku pelajaran. *(anchor)*	33. Kesamaan yang dimiliki oleh televisi, radio, tape recorder, dan telepon genggam yaitu …. a. mMemiliki kesamaan berat bila ditimbang a. Dapat mengeluarkan suara bila dihidupkan *	33. Televisi, radio, *tape recorder*, dan telepon genggam merupakan hasil teknologi yang digunakan oleh manusia. Kesamaan yang dimiliki oleh yaitu …. a. memiliki kesamaan berat bila ditimbang b. dapat mengeluarkan suara bila dihidupkan * c. memiliki kesamaan ukuran volume
2.1.2	Mencari sendiri untuk membuktikan dugaan/ prakiraan sementara (hipotesis) dalam suatu percobaan berdasarkan hasil percobaan/ penyelidikan tentang gejala benda tak hidup.	34. Jika kamu memiliki dua kelompok botol plastik, kelompok botol pertama diisi air mendidih sedangkan kelompok botol kedua diisi air dingin maka yang akan terjadi yaitu kelompok botol yang diisi …. a. air dingin akan penyok air dingin akan penyok b. air mendidi akan penyok*	34. Jika kamu memiliki dua kelompok botol plastik dan kelompok botol pertama diisi air mendidih sedang-kan kelompok botol kedua diisi air dingin maka yang akan terjadi yaitu …. a. kelompok botol yang diisi air dingin akan penyok b. kelompok botol yang diisi air mendidih akan penyok* c. kedua kelompok botol tidak mengalami perubahan

Kode	Indikator	Pilihan dua alternatif	Pilihan tiga alternatif
2.2.4	Memperkirakan perubahan yang terjadi pada benda tak hidup bila diberikan proses pada kondisi tertentu. (*anchor*)	35. Perbedaan yang tampak antara mobil yang terus menerus untuk menempuh jarak jauh dan mobil yang untuk menempuh jarak dekat antara lain bahwa mobil yang terus menerus dipakai untuk menempuh jarak jauh semakin lama semakin .... a. tipis bannya* b. kuat mesinnya	35. Perbedaan antara sepeda motor yang terus menerus untuk menempuh jarak jauh dan sepeda motor yang hanya untuk menempuh jarak dekat yaitu sepeda motor yang dipakai terus menerus untuk menempuh jarak jauh akan lebih cepat .... a. tipis bannya* b. kuat mesinnya c. kuat rantainya

Tabel 2. Kisi-kisi dan Soal beserta Kunci untuk Kelompok Indikator II-IV

Kode	Indikator	Pilihan dua alternatif	Pilihan tiga alternatif
1.1.1	Memilih/menentukan Sendiri jenis makhluk hidup yang akan diamati perubahan bagian tubuhnya berdasarkan warna, bentuk, dan tingkatan.	1. Jika kamu memiliki mangga muda dan mangga yang sudah matang maka yang berbeda adalah..... a. bentuknya b. warna kulitnya*	1. Jika kamu melihat anak sapi dan sapi dewasa maka yang berbeda adalah..... a. bentuk kepalanya b. tinggi tubuhnya* c. jumlah kakinya
1.1.2	Mengidentifikasi /menemukenali sendiri dampak teknologi di alam, di suatu areal, atau di dalam gambar foto. (*anchor*)	2. Selain rasanya, perbedaan antara semangka unggul dengan semangka biasa adalah... a. banyaknya biji* b. warna biji	2. Apa perbedaan yang terdapat pada jambu biji unggul dengan jambu biji biasa? a. ukuran buah* b. warna daun c. warna biji

Kode	Indikator	Pilihan dua alternatif	Pilihan tiga alternatif
1.2.1	Menyajikan sendiri data dalam bentuk tabel lengkap dengan labelnya.	3. Jika kamu diminta membuat tabel yang menyajikan berat setiap buah mangga dari 10 buah mangga Madu dan 10 buah mangga Manalagi maka kolom pertama pada tabel berisi nomor urut untuk menunjukkan nomor buah mangga kemudian pada kolom kedua dan ketiga pada masing-masing kepala kolom diisi ..... a. nama mangga* b. berat mangga	3. Jika kamu diminta membuat tabel yang menyajikan panjang daun setiap tanaman mangga dari 10 tanaman mangga Madu dan 10 tanaman mangga Manalagi maka kolom pertama pada tabel berisi nomor urut untuk menunjukkan nomor daun tanaman mangga kemudian pada kolom kedua dan ketiga pada masing-masing kepala kolom diisi ..... a. nama tanaman mangga* b. panjang daun mangga c. keterangan asal tanaman mangga
1.2.3	Menentukan sendiri tubuh atau bagian tubuh makhluk hidup yang akan digambar dan digambar dengan akurat.	4. Kamu diminta untuk menggambar buah manggis, tetapi hasil gambaranmu berbeda dengan aslinya. Mengapa itu terjadi? a. Belum pernah melihat buah manggis* b. Manggis termasuk buah yang langka	3. Jika kamu ingin menggambar seekor ayam sesuai dengan aslinya, yang dapat kamu lakukan adalah .... a. difoto kemudian menggambarnya dengan melihat foto b. menggambarnya dengan teliti sambil melihat ayamnya* c. melihat banyak ayam lebih dahulu kemudian menggambarnya

Kode	Indikator	Pilihan dua alternatif	Pilihan tiga alternatif
1.2.4	Membuat sendiri suatu bagan/carta, grafik atau histogram tentang kehidupan makhluk hidup.	5. Agar kamu dapat membuat bagan/carta atau grafik/diagram banyaknya buah 3 jenis mangga maka kamu perlu .... a. mengamati terlebih dahulu bentuk buah masing-masing jenis b. menghitung terlebih dahulu banyaknya buah masing-masing jenis*	5. Agar kamu dapat membuat bagan/carta atau grafik/diagram banyaknya buah 3 jenis mangga maka kamu perlu .... a. mengamati terlebih dahulu bentuk buah masing-masing jenis b. menghitung terlebih dahulu banyaknya buah masing-masing jenis* c. menghitung presentase jumlah buah mangga dari masing-masing jenis

Kode	Indikator	Pilihan dua alternatif	Pilihan tiga alternatif
1.3.3	Menyiapkan sendiri peralatan, melakukan atau melanjutkan sendiri langkah yang harus ditempuh dalam melakukan pengamatan gejala kehidupan makhluk hidup setelah membaca petunjuk yang tertulis di Lembar Kerja.	6. Perhatikan petunjuk kegiatan mengukur panjang daun bayam di bawah ini! 1) menyiapkan daun bayam yang diukur 2) menyiapkan penggaris berskala untuk mengukur 3) meletakkan penggaris berskala dari pangkal sampai ujung helaian daun dan mulai dari skala 0 4) membaca skala pada penggaris Berdasarkan petunjuk kegiatan di atas kita dapat melatakkan kertas penggaris berskala dengan titik 0 pada .... a. ujung daun atau pangkal daun* b. pangkal daun dan ujung daun	6. Perhatikan petunjuk kegiatan mengamati pertumbuhn tanaman bayam di bawah ini! 1) menyiapkan tanaman bayam dipot 2) menyirami tanaman bayam setiap pagi dan sore hari 3) mengukur tinggi tanaman dan menghitung jumlah daun setiap hari 4) melakukan pengukuran tersebut selama 5 hari Alat apa yang salah satunya perlu kamu siapkan untuk melakukan kegiatan tersebut? a. busur derajat b. kaca pembesar c. penggaris berskala*

Kode	Indikator	Pilihan dua alternatif	Pilihan tiga alternatif
1.4.3	Menentukan sendiri dasar untuk memisahkan makhluk hidup (hewan dan tumbuhan) atau bagian tubuhnya yang memiliki ciri yang berbeda dan menyatukannya berdasarkan ciri yang sama menurut hasil pengamatan. (*anchor*)	7. Selain sama-sama berguna untuk berjalan, salah satu persamaan yang tedapat pada kaki kambing dengan kaki sapi adalah .... a. ukuran kaki b. jumlah kaki*	4. Perbedaan yang terdapat pada daun singkong dan daun jeruk adalah..... a. pertulangannya* b. warnanya c. bentuknya
1.5.1	Menemukenali atau menyediakan unit-unit/alat pengukur yang benar untuk pengukuran-pengukuran pada umumnya (*anchor*)	8. Jika kamu hendak mengukur suhu tubuh temanmu maka alat yang kamu gunakan adalah..... a. termometer* b. tensimeter	5. Di bawah ini yang <u>bukan</u> merupakan fungsi dari penggaris berskala adalah dapat untuk alat pengukur .... a. tinggi badan b. panjang lengan tangan c. berat badan*

Kode	Indikator	Pilihan dua alternatif	Pilihan tiga alternatif
1.5.3	Menemukan/ mencari sendiri penyebab ketidaktepatan dalam membaca skala pengukuran suhu tubuh makhluk hidup menggunakan termometer laboratorium atau termometer klinis.	9. Bila temanmu mengukur suhu tubuhnya menggunakan termometer dan menunjukkan angka 34 °C padahal temanmu tidak sakit maka terjadi kesalahan pengukuran. Apa kemungkinan penyebabnya? a. Ruangan luas b. Alat rusak*	6. Bila temanmu mengukur suhu tubuhnya menggunakan termometer dan menunjukkan angka 34 °C padahal temanmu tidak sakit maka dipastikan salah mengukur. Apa yang menjadi penyebab kesalahannya? a. Karena termometer yang digunakan hanya satu b. Karena kurang lama menempelkannya pada ketiak* c. Angka skala yang tertera pada termometer kecil
1.5.7	Mengestimasi/ memperkirakan sendiri secara kasar kesamaan ukuran dua tubuh atau bagian tubuh makhluk hidup.	10. Saat diberi tugas untuk mengamati ayam di sekitar rumah, temanmu menemukan dua ekor ayam jantan tinggi dan besarnya hampir sama sehingga temanmu memperkirakaan kedua ayam tersebut sama .... a. umurnya* b. tidurnya	7. Saat diberi tugas untuk mengamati itik manila di sekitar rumah, temanmu menemukan sejumlah itik betina yang terlihat sama dalam hal .... a. bentuk paruhnya b. berat tubuhnya c. besar tubuhnya

Kode	Indikator	Pilihan dua alternatif	Pilihan tiga alternatif
1.6.3	Menemukan sendiri cara menggunakan tubuh/organ tubuh makhluk hidup sebagai alat pengukur.	11. Apabila kita ingin mengukur panjang salah satu lengan tangan kita dengan teliti tetapi tidak ada penggaris berskala maka kita dapat menggunakan .... a. jengkal tangan* b. depa tangan	8. Apabila kita ingin mengukur panjang ubi ketela pohon yang terletak dilantai dengan teliti tetapi tidak tersedia pita ukur maka dapat menggunakan... a. depa tangan b. telapak kaki* c. tungkai kak
1.7.2	Memilih peralatan laboratorium yang sesuai dengan tugas yang dikerjakan untuk mengamati gejala kehidupan makhluk hidup.	12. Saat peragaan memindahkan air untuk mencairkan pupuk daun menggunakan sendok takar, air tumpah. Hal tersebut karena .... a. ukuran sendok takar kurang besar b. sendok takar dalam posisi tidak datar*	9. Ketika temanmu diminta mengukur suhu tubuhnya, ia mengambil termometer pengukur suhu ruangan dan disisipkan ke ketiaknya. Teman-teman satu kelas memandangnya heran karena merupakan tindakan yang .... a. lucu b. jenaka c. salah*
1.7.4	Memindahkan suatu material padat ataupun cairan dengan menggunakan tempat yang spesifik yang berhubungan dengan kegiatan yang berkaitan dengan makhluk hidup	13. Santan kelapa yang kita peroleh dapat saja menjadi sangat encer karena terlalu .... a. lama saat memarut kelapanya sehingga parutan menjadi terlalu halus b. banyak menambahkan air saat parutan kelapa akan diperas*	10. Saat peragaan memindahkan air untuk mencairkan pupuk daun menggunakan sendok, air tumpah. Hal tersebut karena .... a. ukuran sendok kurang besar b. sendok dalam posisi tidak datar* c. sendok yang digunakan sendok tepung

Kode	Indikator	Pilihan dua alternatif	Pilihan tiga alternatif
1.7.9	Mencari sendiri hal yang menyebabkan kesalahan dalam melakukan penyaringan untuk memperoeh ekstrak yang diperoleh dari makhluk hidup.	14. Kendala yang kamu hadapi ketika diminta menghitung jumlah kaki kutu dengan bantuan kaca pembesar yaitu ... a. kutu kepala terus bergerak dengan cepat* b. kutu kepala ukurannya kurang besar	11. Saat kamu ingin memperoleh cairan buah mangga ternyata cairan yang diperoleh terlalu sedikit dibandingkan cairan mangga yang diperoleh temanmu pada waktu yang sama karena .... a. lubang-lubang pada alat penyaringanmu ada yang tersumbat* b. tangkai alat penyaringmu berbeda dengan milik temanmu c. diameter alat penyaringmu lebih besar dibanding milik temanmu
1.7.13	Menemukan sendiri hal yang menyebabkan kesalahan dalam menggunakan lensa pembesar/lup untuk mengamati gejala kehidupan makhluk hidup.	15. Nila, lele, dan hiu termasuk satu kelompok berdasarkan organ yang dimiliki yaitu kelompok ikan .... a. air tawar b. berinsang*	12. Kesalahan yang kamu lakukan ketika menggunakan kaca pembesar untuk mengamati bentuk telur kutu .... a. mata terlalu dekat atau terlalu jauh dari kaca pembesar * b. kaca pembesar yang kamu gunakan terlalu bersih c. kaca pembesarnya bukan yang biasa kamu pakai

Kode	Indikator	Pilihan dua alternatif	Pilihan tiga alternatif
2.1.1	Merumuskan sendiri kesimpulan berdasarkan data pengamatan terhadap makhluk hidup.	16. Tinggi 5 tanaman ketela pohon umur 1 bulan yang tanpa pupuk berturut-turut 20 cm, 21 cm, 22 cm, 20 cm, 19 cm sedangkan tinggi 5 tanaman ketela pohon yang dipupuk urea 10 gram tiap pohon berturut-turut 25 cm, 26 cm, 24 cm, 23 cm, 25 cm. Hal itu mendukung dugaan bahwa pemberian pupuk urea pada ketela pohon ..... a. kurang mempercepat pertumbuhan meningginya b. lebih mempercepat pertumbuhan meningginya*	13. Tanaman Mawar, Melati dan Anggrek Bulan termasuk satu kelompok berdasarkan organ yang dimiliki yaitu kelompok tanaman .... a. berbunga* b. darat c. hias

Kode	Indikator	Pilihan dua alternatif	Pilihan tiga alternatif
2.2.1		17. Jika tanaman mawar dipupuk dan disiram secara teratur maka... a. jumlah daunnya berkurang b. jumlah daunnya bertambah *	14. Tinggi 5 tanaman bayam umur 1 bulan yang tanpa pupuk berturut-turut 21 cm, 20 cm, 22 cm, 20 cm, 19 cm sedangkan tinggi 5 tanaman bayam umur 1 bulan yang dipupuk urea 10 gram tiap pohon berturut-turut 25 cm, 24 cm, 24 cm, 25 cm, 24 cm. Hal itu mendukung dugaan bahwa tanpa pupuk urea maka tanaman bayam ..... a. terhambat laju pertumbuhan meningginya* b. menjadi cepat laju pertumbuhan meningginya c. tidak terpengaruh laju pertumbuhan meningginya

Kode	Indikator	Pilihan dua alternatif	Pilihan tiga alternatif
2.1.2	Mencari sendiri fakta yang dapat mendukung dugaan/ prakiraan sementara (hipotesis) dalam suatu percobaan/ penyelidika n tentang gejala kehidupan makhluk hidup.	18. Pada saat kamu pagi hari apa yang kamu rasakan dibanding pada saat siang hari? a. Tubuh mengeluarkan banyak keringat b. Tubuh mengeluarkan sedikit keringat *	15. Tanaman jeruk yang jarang disiram dan dipupuk secara teratur akan berakibat hasil panen .... a. menjadi sedikit* b. menjadi tambah banyak c. tambah bagus
2.2.4	Memperkirakan perubahan proses fisiologis yang terjadi di dalam makhluk hidup bila diberikan kondisi tertentu. (*anchor*)	19. Jika kamu akan membandingkan keliling dua batang pohon tetapi kamu tidak memiliki pita ukur maka akan lebih tepat jika menggunakan …. a. kain yang diberi angka dengan jarak seperti pada pita ukur* b. kayu yang diberi angka dengan jarak seperti pada pita ukur	16. Setelah kamu selesai olahraga apa yang kamu rasakan pada tubuhmu? a. Tubuh tidak berkeringat b. Jantung berdegub lebih lambat c. Jantung berdegub lebih cepat*

Kode	Indikator	Pilihan dua alternatif	Pilihan tiga alternatif
2.3.1	Menentukan sendiri langkah kerja yang sesuai dalam suatu pengamatan, pengukuran, pemanasan, penghancuran dan penggunaan indikator kimia serta memilih sendiri peralatan yang sesuai dengan permasalahan yang dihadapi dalam kehidupan sehari - hari. (*anchor*)	20. Jika kamu akan memelihara anak ayam tetapi tidak memiliki kandang maka agar ayam tidak berkeliaran lebih baik kamu memasukkannya ke dalam .... a. bak plastik tertutup rapat b. kardus yang diberi lubang kecil*	17. Jika kamu akan memelihara anak ayam tetapi tidak memiliki kandang maka agar ayam tidak berkeliaran lebih baik kamu memasukkannya ke dalam .... a. bak plastik tertutup b. kardus yang diberi lubang kecil* c. lemari pendingin yang telah rusak
1.1.5	Mengidentifikasi / mengenali sendiri keadaan yang berpotensi penuh resiko ketika melakukan pengamatan/ percobaan di sekolah yang berkaitan dengan benda tak hidup dengan yang sama dengan keadaan sehari-hari di rumah.	21. Jika tidak berhati-hati saat menyalakan korek api untuk menyalakan lilin, yang dapat terjadi adalah.... a. kebakaran* b. tangan tergores	18. Seandainya diberi tugas oleh guru untuk memotong gambar daun pada kertas yang diamati maka bahaya yang mungkin terjadi padamu jika tidak berhati-hati adalah.... a. terkena gunting* b. terkena jarum c. tersengat listrik

Kode	Indikator	Pilihan dua alternatif	Pilihan tiga alternatif
1.1.2	Mencocokkan sendiri gambar suatu benda tak hidup dengan yang sesungguhnya atau sebaliknya untuk mengetahui keragaman penampakannya	22. Selain warnanya, jika kamu melihat kursi dan fotonya maka yang sama persis adalah.... a. besarnya b. jumlah kakinya*	19. Jika kamu melihat almari dan fotonya maka kesamaannya antara lain…. a. Bentuknya* b. besarnya c. panjangnya
1.2.1	Memberi sendiri bagan/diagram suatu gejala benda tak hidup secara benar lengkap dengan labelnya.	23. Bagan urutan terjadinya peristiwa hujan yang benar adalah…. a. air menguap – membentuk awan – mengembun – terjadi hujan* b. air menguap – mengembun – membentuk awan – terjadi hujan	20. Bagan urutan terjadinya peristiwa hujan yang benar adalah…. a. Air menguap – membentuk awan – mengembun – terjadi hujan* b. Air mengembun – menguap – membentuk awan – terjadi hujan c. Air menguap – mengembun – membentuk awan – terjadi hujan
1.2.2	Membuat suatu tulisan yang berisi informasi tentang hasil pengamatan gejala benda tak hidup lengkap dengan judulnya.	24. Jika kamu diminta membuat suatu tulisan berupa hasil pengamatanmu tentang batu bata maka dapat kamu tulis antara lain yaitu .... a. umurnya b. warnanya*	21. Jika kamu diminta membuat suatu tulisan berupa hasil pengamatanmu tentang batu bata maka dapat kamu tulis antara lain yaitu .... a. umurnya b. warnanya* c. pembuatnya

Kode	Indikator	Pilihan dua alternatif	Pilihan tiga alternatif
1.3.1	Menyiapkan sendiri peralatan atau melakukan/ melanjutkan sendiri langkah yang harus ditempuh dalam melakukan pengamatan gejala benda tak hidup setelah guru memperagakan/ mendemosntrasikan.	25. Salah satu hal yang dapat menyebabkan gagalnya pengamatan terhadap bagian-bagian sepeda motor yang memerlukan contoh dari guru sebelum mengerjakannya yaitu pada saat diberi penjelasan kamu terus …. a. diam diri saja b. mengantuk*	22. Salah satu hal yang dapat menyebabkan gagalnya pengamatan terhadap bagian-bagian sepeda listrik yang memerlukan contoh dari guru sebelum mengerjakannya karena pada saat diberi penjelasan kamu terus …. a. diam diri saja b. mengantuk* c. memandang guru

Kode	Indikator	Pilihan dua alternatif	Pilihan tiga alternatif
1.4.2	Menentukan sendiri dasar untuk memisahkan benda tak hidup atau bagiannya yang memiliki ciri berbeda berdasarkan data yang tersedia dalam buku/diberikan oleh guru.	26. Hasil pengamatan menunjukkan sepeda memiliki dua roda, tiap roda memiliki ban yang di dalamnya diisi udara dengan cara memompanya, tanpa mesin untuk menggerakkannya. Mobil memiliki empat roda, tiap roda juga memiliki ban yang di dalamnya diisi udara dengan cara memompanya, ada mesin untuk menggerakkannya. Berdasarkan informasi tersebut maka mobil dan motor dibedakan karena …. a. keberadaan bannya b. jumlah rodanya*	26. Hasil pengamatan menunjukkan sepeda memiliki dua roda bulat, tiap roda memiliki ban yang di dalamnya diisi udara dengan cara memompanya, dan tanpa mesin tanpa bahan bakar. Mobil memiliki empat roda, tiap roda juga memiliki ban yang di dalamnya diisi udara dengan cara memompanya, ada mesin untuk menggerakkannya. Berdasarkan informasi tersebut maka mobil dan sepeda dibedakan karena …. a. bentuk ban yang terpasang b. jumlah roda yang terpasang* c. keberadaan udara di dalam ban

Kode	Indikator	Pilihan dua alternatif	Pilihan tiga alternatif
1.5.5	Menemukan sendiri penyebab ketidaktepatan dalam melakukan pengukuran benda tak hidup menggunakan alat yang berskala naik turun.	27. Perhatikan gambar timbangan di bawah ini! Jika timbangan pada gambar saat digunakan untuk menimbang sekantung gula pasir berada dalam keadaan terlalu miring maka yang terjadi adalah.... a. hasil timbangannya akan salah* b. hasil timbangannya akan tetap benar	27. Jika kamu menimbang berat sekantung gula pasir menggunakan timbangan bayi seperti pada kegiatan di posyandu dan kamu <u>tidak teliti</u> saat membaca angka skala pada timbangan maka .... a. kantung gula pasir tersebut akan berubah posisinya dan timbangan akan menjadi rusak karenanya b. sekantung gula pasir tersebut menjadi berubah beratnya dapat menjadi lebih berat atau lebih ringan c. hasil penimbangan sekantung gula pasir tersebut lebih berat atau lebih ringan dari berat sesungguhnya*

Kode	Indikator	Pilihan dua alternatif	han tiga alternatif
1.6.1	Menemukan sendiri cara menggunakan tubuh/organ tubuh makhluk hidup sebagai alat pengukur.	28. Jika kita tidak memiliki penggaris berskala maka untuk mengukur panjang meja kita dapat menggunakan .... a. jengkal tangan* b. jari tangan	28. Jika tidak ada mistar berskala maka langkah kaki dapat digunakan untuk mengukur .... a. berat badan pemain sepak bola b. tinggi tiang gawang sepak bola c. jarak antartiang gawang sepak bola*
1.7.1	Memilih sendiri cara kerja untuk mengurangi kesalahan dalam melakukan percobaan yang berhubungan dengan benda tak hidup.	29. Agar hasil percobaan dapat dipercaya maka untuk mengetahui akibat perbedaan tingkat pemanasan terhadap pemuaian panjang besi maka …. a.disediakan potongan-potongan besi yang panjangnya sama* b. digunakan potongan besi yang besar dan kecil	29. Agar hasil percobaan dapat dipercaya untuk mengetahui akibat perbedaan tingkat pemanasan terhadap pemuaian panjang besi maka …. a. disediakan potongan-potongan besi yang panjangnya sama* b. digunakan potongan besi yang besar dan yang kecil ukurannya c. tiap batang besi dilakukan pemanasan sekali dan dua kali

Kode	Indikator	Pilihan dua alternatif	han tiga alternatif
1.7.2	Mencari sendiri hal yang menyebabkan kesalahan dalam membuat campuran atau larutan bahan yang digunakan dalam percobaan yang berhubungan dengan benda tak hidup.	30. Jika kamu akan membuat air teh manis maka tidak boleh …. a. terlalu lama kamu aduk-aduk b. terlalu sedikit menambahkan gula*	30. Jika kamu akan membuat air teh manis maka tidak boleh …. a. terlalu lama kamu aduk-aduk b. terlalu sedikit menambahkan gula* c. menggunakan gelas yang besar
1.7.3	Menemukan sendiri langkah-langkah kerja yang aman dalam menggunakan peralatan laboratorium yang mudah terbakar.	31. Jika kita harus praktik berkelompok menggunakan alkohol sebaiknya …. a. menjauhkannya dari api seperti nyala lilin atau korek* b. jangan terlalu dekat dengan kelompok lain	31.Jika kita harus praktik berkelompok menggunakan alkohol sebaiknya …. a. menjauhkannya dari api seperti nyala lilin atau korek* b. jangan terlalu dekat dengan kelompok lain c. menggunakan baju khusus untuk praktik

Kode	Indikator	Pilihan dua alternatif	Pilihan tiga alternatif
2.1.1	Merumuskan sendiri kesimpulan yang sesuai dengan data hasil pengamatan/ penyelidikan yang ada dan informasi yang ada dalam buku pelajaran yang berkaitan dengan benda tak hidup.	32. Hasil pengamatan menunjukkan dibandingkan batu sungai, batu bata lebih lunak, warnanya juga lebih cerah, dan bangunnya teratur yaitu berbentuk balok. Hasil dari membaca buku ternyata batu bata dibuat untuk membangun bangunan tembok rumah. Di kawasan tertentu yang tidak tersedia batu bata juga digunakan batu kali untuk membuat tembok rumah. Kesimpulan yang dapat ditarik yaitu batu bata digunakan sebagai batu buatan .... a. pengganti batu sungai* b. tandingan batu sungai	32. Hasil pengamatan menunjukkan dibandingkan batu sungai, batu bata lebih lunak, warnanya juga lebih cerah, dan bangunnya teratur yaitu berbentuk balok. Hasil dari membaca buku ternyata batu bata dibuat untuk membangun tembok rumah. Di kawasan tertentu yang tidak tersediaa batu bata juga digunakan batu kali untuk membuat tembok rumah. Kesimpulan yang dapat ditarik yaitu batu bata digunakan sebagai batu buatan .... a. tandingan batu sungai b. yang berbeda dengan batu sungai c. pengganti batu sungai*

Kode	Indikator	Pilihan dua alternatif	Pilihan tiga alternatif
2.1.2	Merumuskan sendiri dugaan/ prakiraan sementara dalam suatu percobaan berdasarkan hasil percobaan/ penyelidikan tentang gejala benda tak hidup.	33. Jika kamu menjemur pakaian basah di bawah sinar matahari pakaian tersebut akan …. a. tetap basah b. menjadi kering*	33. Jika kamu menjemur pakaian basah di bawah sinar matahari pakaian tersebut akan …. a. tetap basah b. menjadi kering* c. menjadi basah
2.3.1	Menemukan sendiri cara mengantisipasi risiko serta mengambil tindakan pencegahan dan cara kerja yang sesuai dalam melakukan percobaan/ penyelidikan tentang gejala pada benda tak hidup.	34. Agar tangan tidak terkena minyak ketika kamu praktik IPA maka yang terbaik kamu mengenakan sarung tangan yang terbuat daeri …. a. kain b. karet*	34. Agar tangan tidak terkena minyak ketika kamu praktik IPA maka yang terbaik kamu …. a. mengenakan sarung tangan yang terbuat dari kain b. membungkus tanganmu menggunakan kain bekas c. mengenakan sarung tangan yang terbuat dari karet*

Kode	Indikator	Pilihan dua alternatif	Pilihan tiga alternatif
2.3.3	Memilih sendiri hal-hal yang dapat mempengaruhi hasil suatu percobaan/ penyelidikan, menentukan sendiri data yang diperoleh sesuai hasil percobaan/ penyelidikan dan memilih sendiri cara menyajikan data hasil percobaan tentang gejala benda tak hidup.	35. Saya akan melakukan percobaan untuk mengetahui akibat lama penjemuran terhadap tingkat kekeringan kain maka ada banyak hal yang harus saya siapkan, diantaranya yaitu akan saya pilih …. a. lokasi penjemuran yang berbeda-beda, kemudian setelah data saya peroleh akan saya sajikan dalam bentuk tabel b. satu lokasi penjemuran, kemudian setelah data saya peroleh akan saya sajikan dalam bentuk diagram*	35. Jika saya akan melakukan percobaan untuk mengetahui akibat lama penjemuran terhadap tingkat kekeringan kain maka ada banyak hal yang harus saya siapkan, diantaranya yaitu …. a. akan saya pilih lokasi penjemuran yang berbeda-beda, kemudian setelah data saya peroleh akan saya sajikan dalam bentuk table b. akan saya pilih satu lokasi penjemuran, kemudian setelah data saya peroleh akan saya sajikan dalam bentuk histogram* c. akan saya pilih lokasi yang terkena matahari dan yang tidak, kemudian setelah data saya peroleh akan saya sajikan dalam bentuk diagram

Tabel 3. Kisi-kisi dan Soal beserta Kunci untuk Kelompok Indikator I-III yang diujikan tahun 2018

Kode	Kelompok Indikator I-III	Soal dengan dua pilihan jawaban
1.1.1	Memilih dan mencocokkan sendiri objek berupa makhluk hidup dengan gambarnya.	1. Jika kamu melihat dirimu dan fotomu maka yang sama persis antara lain .... a. warna rambutmu* b. besar tubuhmu
1.1.2	Mengidentifikasi /mengenali sendiri dampak teknologi di alam,di suatu areal, atau didalam gambar foto. (*anchor*)	2. Selain rasanya, perbedaan antara jambu biji unggul dengan jambu biji biasa adalah... a. banyaknya biji* b. warna biji
1.2.1	Melengkapi sendiri suatu bagan/carta, grafik atau histogram tentang fenomena kehidupan makhluk hidup.	3. Ikan ----- sungai Sapi ----- darat Mana yang sesuai dengan bagan di atas? a. Itik manila ----- kolam b. Kera ----- pepohonan*
1.2.3	Menentukan sendiri tubuh atau bagian tubuh makhluk hidup yang akan digambar dan digambar dengan akurat.	4. Jika kamu ingin menggambar seekor ayam sesuai dengan aslinya, yang dapat kamu lakukan adalah …. a. difoto kemudian menggambarnya dengan melihat foto b. menggambarnya dengan teliti sambil melihat ayamnya*

Kode	Kelompok Indikator I-III	Soal dengan dua pilihan jawaban
1.3.1	Menyiapkan sendiri peralatan atau melakukan/melanjutkan sendiri langkah yang harus ditempuh dalam melakukan pengamatan gejala kehidupan makhluk hidup setelah guru menjelaskan secara lisan.	5. Setelah guru menjelaskan cara-cara melakukan pengamatan terhadap tinggi batang secara lisan maka apa langkah yang harus kamu lakukan selanjutnya? a. Memilih alat yang harus digunakan untuk mengamati* b. Segera melakukan pengukuran terhadap tinggi tanaman
1.4.3	Menentukan sendiri dasar untuk memisahkan makhluk hidup (hewan dan tumbuhan) atau bagian tubuhnya yang memiliki ciri yang berbeda dan menyatukannya berdasarkan ciri yang sama menurut hasil pengamatan. (*anchor*)	6. Perbedaan mencolok antara daun singkong dan daun jeruk adalah pada ..... a. pertulangannya* b. warnanya
1.5.1	Menemukenali atau menyediakan unit-unit/alat pengukur yang benar untuk pengukuran-pengukuran pada umumnya	7. Jika kamu hendak mengukur suhu tubuh temanmu, alat yang kamu gunakan adalah..... a. termometer* b. tensimeter

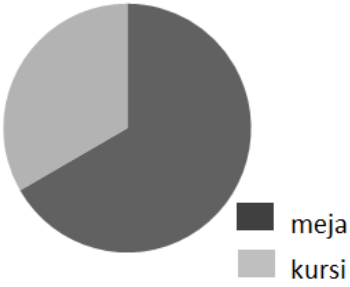
Kode	Kelompok Indikator I-III	Soal dengan dua pilihan jawaban
1.5.2	Menemukan sendiri penyebab ketidaktepatan dalam membaca skala meteran atau pita ukur ketika melakukan pengukuran ciri tubuh makhluk hidup.	8. Selain kelainan pada mata, penyebab ketidaktepatan membaca skala pada pita ukur ketika kamu mengukur panjang lengan tangan temanmu adalah... a. cahaya lampu di dalam ruangan yang terang b. mata tidak sejajar dengan skala pada pita ukur*
1.5.3	Menemukan/ mencari sendiri penyebab ketidaktepatan dalam membaca skala pengukuran suhu tubuh makhluk hidup menggunakan termometer laboratorium atau termometer klinis.	9. Bila temanmu mengukur suhu tubuhnya menggunakan termometer dan menunjukkan angka 34 ^{o}C padahal temanmu <u>tidak</u> sakit maka kemungkinan penyebabnya adalah ... a. ruangannya luas b. alatnya rusak*
1.6.1	Menemukan sendiri cara menggunakan keterampilan tangan untuk membuat karya yang berhubungan dengan gejala kehidupan makhluk hidup.	10. Untuk dapat membuat herbarium atau awetan daun yang dikeringkan maka kamu perlu …. c. menjaga agar helaian daun <u>tidak terlipat</u>* d. memisahkan daun yang telah kering

Kode	Kelompok Indikator I-III	Soal dengan dua pilihan jawaban
1.6.2	Mencari sendiri cara memindahkan benda padat atau cair dalam suatu percobaan yang berhubungan dengan gejala kehidupan makhluk hidup.	11. Jika kamu diminta menangkap ikan dari kolam yang akan dipindahkan ke dalam aquarium maka kamu dapat melakukan dengan menggunakan… a. jaring* b. panic
1.6.3	Mencari sendiri hal yang menyebabkan kesalahan dalam menghaluskan bahan yang diperoleh dari makhluk hidup.	12. Jika kamu diminta memotong rumput menjadi potongan yang kecil-kecil untuk pakan ternak maka dapat terjadi kegagalan akibat …. a. pisau kurang tajam* b. rumput terlalu hijau
1.7.4	Memindahkan suatu material padat ataupun cairan dengan menggunakan tempat yang spesifik yang berhubungan dengan kegiatan yang berkaitan dengan makhluk hidup anchor	13. Saat memindahkan air untuk mencairkan pupuk daun menggunakan sendok, air tumpah. Hal tersebut karena .... a. ukuran sendok takar kecil b. sendok takar dalam posisi miring*

Kode	Kelompok Indikator I-III	Soal dengan dua pilihan jawaban
2.1.1	Menemukan sendiri perbedaan antara bentuk organ tubuh makhluk hidup yang ada digambar	14. Jika kamu mengamati gambar ayam jantan dan betina maka perbedaan keduanya terletak pada panjang bulu …. a. ekor* b. sayap
2.1.2	Merumuskan sendiri kesimpulan berdasarkan data pengamatan terhadap makhluk hidup.	15. Ketika kamu mengamati berbagai jenis hewan seperti kadal, cicak dan buaya, ciri manakah yang menyebabkan ketiga hewan tersebut termasuk dalam kelompok yang sama? a. tidak berdaun telinga* b. hidup di air
2.1.3	Merumuskan sendiri dugaan/ prakiraan sementara (hipotesis) berdasarkan hasil pengamatan/ penyelidikan dan informasi yang ada	16. Bila kamu meletakkan pot tanaman di dekat jendela maka tanaman tersebut batangnya akan membengkok .... a. mendekati ke arah jendela* b. menjauh dari arah jendela
2.2.1	Memperkirakan sendiri kemungkinan yang terjadi ketika suatu mahluk hidup saat sekarang sedang berada pada kondisi yang tidak mendukung berdasar fakta yang dihadapi.	17. Apa yang akan terjadi apabila sungai hampir mengering saat musim kemarau? a. Banyak ikan yang akan mati. * b. Banyak ikan yang bertelur.

Kode	Kelompok Indikator I-III	Soal dengan dua pilihan jawaban
2.2.2	Memperkirakan perubahan proses fisiologis yang terjadi di dalam makhluk hidup bila diberikan kondisi tertentu.	18. Setelah kamu selesai olahraga apa yang kamu rasakan pada tubuhmu? a. Jantung berdegub lebih cepat.* b. Jantung berdegub lebih lambat.
2.3.1	Menentukan sendiri langkah kerja yang sesuai dalam suatu pengamatan, pengukuran, pemanasan, penghancuran dan menggunakan indikator kimia serta memilih sendiri peralatan yang sesuai dengan permasalahan yang dihadapi dalam kehidupan sehari-hari.	19. Jika kamu akan membandingkan keliling dua batang pohon tetapi kamu tidak memiliki pita ukur maka akan lebih tepat jika menggunakan…. a. kain yang diberi angka dengan jarak seperti pada pita ukur* b. kayu yang diberi angka dengan jarak seperti pada pita ukur
2.3.1	Menentukan sendiri langkah kerja yang sesuai dalam suatu pengamatan, pengukuran, pemanasan, penghancuran dan penggunaan indikator kimia serta memilih sendiri peralatan yang sesuai dengan permasalahan yang dihadapi dalam kehidupan sehari - hari. (*anchor*)	20. Jika kamu akan memelihara anak ayam tetapi tidak memiliki kandang maka agar ayam tidak berkeliaran lebih baik kamu memasukkannya ke dalam …. a. bak plastik tertutup rapat b. kardus yang diberi lubang kecil*

Kode	Kelompok Indikator I-III	Soal dengan dua pilihan jawaban
1.1.1	Memilih sendiri gejala yang akan dibandingkan bila dihadapkan pada dua benda tak hidup untuk mengidentifikasi perbedaan secara terperinci.	21. Jika kamu mengamati sepeda motor dan mobil, apa perbedaan dari keduanya? c. Bahan bakarnya. d. Jumlah rodanya.*
1.2.2	Membuat sendiri ringkasan suatu paragraf/bab/ buku yang mengulas gejala benda tak hidup	22. Jika kamu diperintah untuk meringkas bab mengenai penyebab banjir maka kalimat yang <u>tidak</u> tepat adalah a. Bencana alam berupa banjir dapat disebabkan karena banyak manusia yang membuang sampah disungai b. Membuang sampah di sungai dilakukan untuk mengurangi sampah yang menumpuk di halaman*
1.2.3	Membuat histogram tentang gejala benda tak hidup lengkap dengan labelnya.	23. Jika kamu diminta untuk membuat histogram jumlah jendela dan pintu yang ada pada suatu gedung sekolah maka yang harus dicantumkan dalam histogram adalah .... a. bentuk jendela dan pintu yang ada b. jumlah jendela dan pintu yang ada*

Kode	Kelompok Indikator I-III	Soal dengan dua pilihan jawaban
1.2.4	Menyampaikan sendiri informasi tentang suatu benda tak hidup yang tersaji dalam bentuk bagan/carta, grafik atau histogram.	24. Informasi yang tepat berdasarkan sajian data berikut yaitu ....  c. jumlah meja dua kali dari jumlah kursi* d. jumlah meja setengah kali dari jumlah kursi
1.5.1	Menemukenali atau menyediakan unit-unit/alat pengukur yang benar untuk pengukuran-pengukuran pada umumnya terhadap gejala benda tak hidup	25.Jika kamu diminta untuk mengukur tinggi meja, apa yang sebaiknya kamu gunakan? a. Penggaris berskala.* b. Jengkal tangan.
1.5.4	Menemukan sendiri penyebab ketidaktepatan dalam melakukan pengukuran ciri benda tak hidup menggunakan alat penimbang atau alat pengukur volume. *(anchor)*	26.Jika kamu menimbang berat buku sebanyak dua kali tetapi hasilnya berbeda maka kemungkinan yang terjadi adalah .... c. kesalahan saat melihat angka pada timbangan* d. buku yang ditimbang termasuk terlalu ringan

Kode	Kelompok Indikator I-III	Soal dengan dua pilihan jawaban
1.5.8	Menemukan sendiri kesalahan dalam penggunaan kisi-kisi/jaring (*grid*) untuk mengestimasi luas suatu area/permukaan benda/bagian benda tak hidup. (*anchor*)	27. Jika kamu diminta melakukan pengukuran luas gambar rumah menggunakan kertas bergaris kisi-kisi berjarak satu milimeter (*milimeter block*) maka kamu harus membuat blad gambar rumah yang diukur pada kertas tersebut. Jika hasilnya lebih sempit dari luas gambar rumah yang sesungguhnya maka kemungkinannya .... a. kamu kurang cermat menghitung banyaknya kotak milimeter pada kertas kisi-kisi yang tertutup daun* b. ada bagian gambar rumah yang menutup kertas bergaris kisi-kisi berjarak satu milimeter
1.6.2	Menemukan sendiri cara menggunakan keterampilan tangan untuk membuat karya yang berhubungan dengan gejala benda tank hidup.	28. Jika kamu diminta untuk membuat vas bunga dari tanah liat maka cara yang dapat dilakukan adalah .... a. tanah liat dikeringkan di bawah sinar matahari, kemudian dibentuk menjadi vas b. tanah liat dibasahi dengan air, dibentuk menjadi vas baru kemudian dikeringkan*

Kode	Kelompok Indikator I-III	Soal dengan dua pilihan jawaban
1.7.2	Mencari sendiri jenis peralatan yang akan digunakan untuk mengukur gejala benda tak hidup sesuai dengan tugas yang diberikan oleh Bapak/Ibu Guru.	29. Jika kamu diminta mengukur lingkar tiang bendera, maka alat ukur yang kamu gunakan adalah .... a. penggaris berskala b. *roll meter* dari plastik*
1.7.3	Mencari sendiri hal yang menyebabkan kurang berfungsinya pipet tetes untuk memindahkan larutan pada percobaan yang berhubungan dengan gejala benda tak hidup.	30. Seandainya kamu diminta memindahkan cairan menggunakan pipet akan terjadi kegagalan karena .... a. karet pipet sudah tidak bisa kembali ketika ditekan* b. ujung pipet terlalu masuk ke dalam botol tempat cairan vitamin
1.7.4	Mencari sendiri hal yang menyebabkan kurang berfungsinya kertas uji bahan kimia cair pada pengujian bahan yang diperoleh dari benda tak hidup.	31. Ketika kamu hendak mengukur pH cairan sabun menggunakan kertas pH yang sudah basah maka kemungkinan yang terjadi adalah ... a. pH air sabun <u>tidak</u> terukur dengan tepat* b. kertas pH berubah-ubah warnanya

1.7.5	Menemukan sendiri hal yang menyebabkan kesalahan dalam menggunakan *stopclock* atau *stopwatch*.untuk mengukur gejala benda tak hidup.	32. Kesalahan yang terjadi jika kamu menggunakan *stopwatch* untuk mengukur kecepatan arus air sungai dapat terjadi karena …. a. gerak jarum pada *stopwatch* terlalu lambat* b. *stopwatch* sudah beberapa kali digunakan
2.1.1	Merumuskan sendiri kesimpulan yang sesuai dengan data hasil pengamatan/penyelidikan yang ada terhadap benda tak hidup dan informasi yang ada dalam buku pelajaran. *(anchor)*	33. Kesamaan yang dimiliki oleh televisi, radio, *tape recorder*, dan telepon genggam yaitu …. b. memiliki kesamaan berat bila ditimbang b. dapat mengeluarkan suara bila dihidupkan *
2.1.2	Mencari sendiri untuk membuktikan dugaan/ prakiraan sementara (hipotesis) dalam suatu percobaan berdasarkan hasil percobaan/ penyelidikan tentang gejala benda tak hidup.	34. Jika kamu memiliki dua kelompok botol plastik, kelompok botol pertama diisi air mendidih sedangkan kelompok botol kedua diisi air dingin maka yang akan terjadi yaitu kelompok botol yang diisi …. a. air dingin akan penyok b. air mendidih akan penyok*

Kode	Kelompok Indikator I-III	Soal dengan dua pilihan jawaban
2.2.4	Memperkirakan perubahan yang terjadi pada benda tak hidup bila diberikan proses pada kondisi tertentu. (*anchor*)	35. Perbedaan yang tampak antara mobil yang terus menerus untuk menempuh jarak jauh dan mobil yang untuk menempuh jarak dekat antara lain bahwa mobil yang terus menerus dipakai untuk menempuh jarak jauh semakin lama semakin …. a. tipis bannya* b. kuat mesinnya

Tabel 4. Kisi-kisi dan Soal beserta Kunci untuk Kelompok Indikator II-IV yang diujikan tahun 2018

Kode	Kelompok Indikator II-IV	Soal dengan dua pilihan jawaban
1.1.1	Memilih/menentukan Sendiri jenis makhluk hidup yang akan diamati perubahan bagian tubuhnya berdasarkan warna, bentuk, dan tingkatan.	1. Jika kamu memiliki mangga muda dan mangga yang sudah matang maka yang berbeda adalah..... c. bentuknya d. warna kulitnya*
1.1.2	Mengidentifikasi/menemukenali sendiri dampak teknologi di alam, di suatu areal, atau di dalam gambar foto. (*anchor*)	2. Apa perbedaan yang terdapat pada jambu biji unggul dengan jambu biji biasa? a. ukuran buah* b. warna daun
1.2.1	Melengkapi suatu bagan/carta, grafik atau histogram tentang fenomena/ciri/gejalamakhluk hidup	3. Jika kamu diminta membuat tabel yang menyajikan berat setiap buah mangga dari 10 buah mangga Madu dan 10 buah mangga Manalagi maka kolom pertama pada tabel berisi nomor urut untuk menunjukkan nomor buah mangga kemudian pada kolom kedua dan ketiga pada masing-masing kepala kolom diisi ..... a. nama mangga* b. berat mangga

Kode	Kelompok Indikator II-IV	Soal dengan dua pilihan jawaban
1.2.3	Menentukan sendiri tubuh atau bagian tubuh makhluk hidup yang akan digambar dan digambar dengan akurat.	4. Jika kamu ingin menggambar seekor ayam sesuai dengan aslinya, yang dapat kamu lakukan adalah …. a. difoto kemudian menggambarnya dengan melihat foto b. menggambarnya dengan teliti sambil melihat ayamnya*
1.3.1	Membuat sendiri suatu bagan/carta, grafik atau histogram tentang kehidupan makhluk hidup.	5. Agar kamu dapat membuat bagan/carta atau grafik/diagram banyaknya buah 3 jenis mangga maka kamu perlu .... a. mengamati terlebih dahulu bentuk buah masing-masing jenis b. menghitung terlebih dahulu banyaknya buah masing-masing jenis*

Kode	Kelompok Indikator II-IV	Soal dengan dua pilihan jawaban
1.4.3	Menyiapkan sendiri peralatan, melakukan atau melanjutkan sendiri langkah yang harus ditempuh dalam melakukan pengamatan gejala kehidupan makhluk hidup setelah membaca petunjuk yang tertulis di Lembar Kerja.	6. Perhatikan petunjuk kegiatan mengamati pertumbuhn tanaman bayam di bawah ini! 1) menyiapkan tanaman bayam di pot 2) menyiraminya setiap pagi dan sore hari 3) mengukur tinggi tanaman dan menghitung jumlah daun setiap hari 4) melakukan pengukuran tersebut selama 5 hari Alat apa yang salah satunya perlu kamu siapkan untuk melakukan kegiatan tersebut? a. busur derajat b. penggaris berskala*
1.5.1	Menentukan sendiri dasar untuk memisahkan makhluk hidup (hewan dan tumbuhan) atau bagian tubuhnya yang memiliki ciri yang berbeda dan menyatukannya berdasarkan ciri yang sama menurut hasil pengamatan. (*anchor*)	7. Perbedaan mencolok antara daun singkong dan daun jeruk adalah pada ..... a. pertulangannya* b. warnanya

Kode	Kelompok Indikator II-IV	Soal dengan dua pilihan jawaban
1.5.2	Menemukenali atau menyediakan unit-unit/alat pengukur yang benar untuk pengukuran-pengukuran pada umumnya	8. Penggaris berskala tidak dapat untuk alat pengukur .... a. tinggi badan b. berat badan*
1.5.3	Menemukan/ mencari sendiri penyebab ketidaktepatan dalam membaca skala pengukuran suhu tubuh makhluk hidup menggunakan termometer laboratorium atau termometer klinis.	9. Bila temanmu mengukur suhu tubuhnya menggunakan termometer dan menunjukkan angka 34 ^{o}C padahal temanmu tidak sakit maka kemungkinan penyebabnya adalah ... a. ruangannya luas b. alatnya rusak*
1.6.1	Mengestimasi/ memperkirakan sendiri secara kasar kesamaan ukuran dua tubuh atau bagian tubuh makhluk hidup.	10. Saat diberi tugas untuk mengamati itik manila di sekitar rumah, temanmu menemukan sejumlah itik manila betina terlihat sama dalam hal .... a. bentuk paruhnya* b. berat tubuhnya

Kode	Kelompok Indikator II-IV	Soal dengan dua pilihan jawaban
1.6.2	Menemukan sendiri cara menggunakan tubuh/organ tubuh makhluk hidup sebagai alat pengukur.	11. Apabila kita ingin mengukur panjang ubi ketela pohon yang terletak dilantai dengan teliti tetapi tidak tersedia pita ukur maka dapat menggunakan... a. depa tangan b. telapak kaki*
1.6.3	Memilih peralatan laboratorium yang sesuai dengan tugas yang dikerjakan untuk mengamati gejala kehidupan makhluk hidup.	12. Untuk melakukan pengamatan terhadap bentuk tubuh kutu kepala akan lebih mudah jika menggunakan ... a. kaca pembesar* b. cermin cembung
1.7.4	Memindahkan suatu material padat ataupun cairan dengan menggunakan tempat yang spesifik yang berhubungan dengan kegiatan yang berkaitan dengan makhluk hidup	13. Saat memindahkan air untuk mencairkan pupuk daun menggunakan sendok, air tumpah. Hal tersebut karena .... a. ukuran sendok takar kecil b. sendok takar dalam posisi miring*

Kode	Kelompok Indikator II-IV	Soal dengan dua pilihan jawaban
2.1.1	Mencari sendiri hal yang menyebabkan kesalahan dalam melakukan penyaringan untuk memperoeh ekstrak yang diperoleh dari makhluk hidup.	14. Saat kamu ingin memperoleh cairan buah mangga ternyata cairan yang diperoleh terlalu sedikit dibandingkan cairan mangga yang diperoleh temanmu pada waktu yang sama karena .... d. lubang-lubang pada alat penyaringanmu ada yang tersumbat* e. tangkai alat penyaringmu berbeda dengan milik temanmu
2.1.2	Menemukan sendiri hal yang menyebabkan kesalahan dalam menggunakan lensa pembesar/lup untuk mengamati gejala kehidupan makhluk hidup.	15. Kesalahan yang kamu lakukan ketika menggunakan kaca pembesar untuk mengamati bentuk telur kutu .... a. mata terlalu dekat atau terlalu jauh dari kaca pembesar * b. kaca pembesar yang kamu gunakan terlalu bersih

Kode	Kelompok Indikator II-IV	Soal dengan dua pilihan jawaban
2.1.3	Merumuskan sendiri kesimpulan berdasarkan data pengamatan terhadap makhluk hidup.	16. Tinggi 5 tanaman ketela pohon umur 1 bulan yang <u>tanpa pupuk</u> berturut-turut 20 cm, 21 cm, 22 cm, 20 cm, 19 cm sedangkan tinggi 5 tanaman ketela pohon yang dipupuk urea 10 gram tiap pohon berturut-turut 25 cm, 26 cm, 24 cm, 23 cm, 25 cm. Hal itu mendukung dugaan bahwa pemberian pupuk urea pada ketela pohon ..... c. kurang mempercepat pertumbuhan meningginya d. lebih mempercepat pertumbuhan meningginya*
2.2.1	Mencari sendiri fakta yang dapat mendukung dugaan/ prakiraan sementara (hipotesis) dalam suatu percobaan/ penyelidikan tentang gejala kehidupan makhluk hidup.	17. Apa yang kamu alami saat pagi hari dibandingkan pada saat siang hari? a. Tubuh mengeluarkan banyak keringat b. Tubuh mengeluarkan sedikit keringat *

Kode	Kelompok Indikator II-IV	Soal dengan dua pilihan jawaban
2.2.2	Memperkirakan kemungkinan yang terjadi ketika suatu makhluk hidup saat sekarang sedang berada pada kondisi yang tidak mendukung berdasar fakta yang dihadapi	18. Jika tanaman hias yang kamu pelihara suatu saat terlambat kamu siram maka kemungkinannya akan menunjukkan gejala awal ... a. daun tanaman layu* b. daun tanaman gugur
2.3.1	Memperkirakan perubahan proses fisiologis yang terjadi di dalam makhluk hidup bila diberikan kondisi tertentu.	19. Jika kamu akan membandingkan keliling dua batang pohon tetapi kamu tidak memiliki pita ukur maka akan lebih tepat jika menggunakan …. a. kain yang diberi angka dengan jarak seperti pada pita ukur* b. kayu yang diberi angka dengan jarak seperti pada pita ukur

Kode	Kelompok Indikator II-IV	Soal dengan dua pilihan jawaban
2.3.1	Menentukan sendiri langkah kerja yang sesuai dalam suatu pengamatan, pengukuran, pemanasan, penghancuran dan penggunaan indikator kimia serta memilih sendiri peralatan yang sesuai dengan permasalahan yang dihadapi dalam kehidupan sehari - hari. (*anchor*)	20. Jika kamu akan memelihara anak ayam tetapi tidak memiliki kandang maka agar ayam tidak berkeliaran lebih baik kamu memasukkannya ke dalam …. a. bak plastik tertutup rapat b. kardus yang diberi lubang kecil*
1.1.1	Mencocokkan sendiri gambar suatu benda tak hidup dengan yang sesungguhnya atau sebaliknya untuk mengetahui keragaman penampakannya.	20. Selain warnanya, jika kamu melihat kursi dan fotonya maka yang sama persis adalah.... a. besarnya b. jumlah kakinya*
1.2.2	Mengidentifikasi/ mengenali sendiri keadaan yang berpotensi penuh resiko ketika melakukan pengamatan/ percobaan di sekolah yang berkaitan dengan benda tak hidup dengan yang sama dengan keadaan sehari-hari di rumah.	22. Jika <u>tidak</u> berhati-hati saat menyalakan korek api untuk menyalakan lilin, yang dapat terjadi adalah…. a. kebakaran* b. tangan tergores

Kode	Kelompok Indikator II-IV	Soal dengan dua pilihan jawaban
1.2.3	Memberi sendiri bagan/diagram suatu gejala benda tak hidup secara benar lengkap dengan labelnya.	23. Bagan urutan terjadinya peristiwa hujan yang benar adalah…. a. air menguap – membentuk awan – mengembun – terjadi hujan* b. air menguap – mengembun – membentuk awan – terjadi hujan
1.2.4	Membuat suatu tulisan yang berisi informasi tentang hasil pengamatan gejala benda tak hidup lengkap dengan judulnya.	24. Jika kamu diminta membuat suatu tulisan berupa hasil pengamatanmu tentang batu bata maka dapat kamu tulis antara lain yaitu .... a. umurnya b. warnanya*
1.5.1	Menyiapkan sendiri peralatan atau melakukan/ melanjutkan sendiri langkah yang harus ditempuh dalam melakukan pengamatan gejala benda tak hidup setelah guru memperagakan/ mendemosntrasikan.	25. Salah satu hal yang dapat menyebabkan gagalnya pengamatan terhadap bagian-bagian sepeda motor yang memerlukan contoh dari guru sebelum me-ngerjakannya yaitu pada saat diberi penjelasan kamu terus …. a. diam diri saja b. mengantuk*

Kode	Kelompok Indikator II-IV	Soal dengan dua pilihan jawaban
1.5.4	Menentukan sendiri dasar untuk memisahkan benda tak hidup atau bagiannya yang memiliki ciri berbeda berdasarkan data yang tersedia dalam buku/diberikan oleh guru.	26. Hasil pengamatan menunjukkan mobil memiliki empat roda, tiap roda juga memiliki ban yang di dalamnya diisi udara dengan cara memompanya, ada mesin untuk mengge-rakkannya sedasngkan sepeda memiliki dua roda, tiap roda memiliki ban yang di dalamnya diisi udara dengan cara memompanya, tanpa mesin untuk mengge-rakkannya. Berdasarkan informasi tersebut maka mobil dan motor dibedakan karena …. a. keberadaan bannya b. jumlah rodanya*

Kode	Kelompok Indikator II-IV	Soal dengan dua pilihan jawaban
1.5.8	Menemukan sendiri penyebab ketidaktepatan dalam melakukan pengukuran benda tak hidup menggunakan alat yang berskala naik turun.	27. Perhatikan gambar timbangan di bawah ini! Jika timbangan pada gambar saat digunakan untuk menimbang sekantung gula pasir berada dalam keadaan terlalu miring maka yang terjadi adalah.... a. hasil timbangannya akan salah* b. hasil timbangannya akan tetap benar
1.6.2	Menemukan sendiri cara menggunakan tubuh/organ tubuh makhluk hidup sebagai alat pengukur.	28. Jika kita tidak memiliki penggaris berskala maka untuk mengukur panjang meja kita dapat menggunakan .... a. jengkal tangan* b. jari tangan

Kode	Kelompok Indikator II-IV	Soal dengan dua pilihan jawaban
1.7.2	Memilih sendiri cara kerja untuk mengurangi kesalahan dalam melakukan percobaan yang berhubungan dengan benda tak hidup.	29. Agar hasil percobaan dapat dipercaya maka untuk mengetahui akibat perbedaan tingkat pemanasan terhadap pemuaian panjang besi maka …. a. disediakan potongan-potongan besi yang panjangnya sama* b. digunakan potongan besi yang besar dan kecil
1.7.3	Mencari sendiri hal yang menyebabkan kesalahan dalam membuat campuran atau larutan bahan yang digunakan dalam percobaan yang berhubungan dengan benda tak hidup.	30. Jika kamu akan membuat minuman teh manis maka <u>tidak boleh</u> …. a. terlalu lama kamu aduk-aduk b. terlalu sedikit menambahkan gula*
1.7.4	Menemukan sendiri langkah-langkah kerja yang aman dalam menggunakan peralatan laboratorium yang mudah terbakar.	31. Jika kita harus praktik berkelompok menggunakan alkohol sebaiknya …. a. menjauhkannya dari api seperti nyala lilin atau korek* b. jangan terlalu dekat dengan kelompok lain

Kode	Kelompok Indikator II-IV	Soal dengan dua pilihan jawaban
1.7.5	Merumuskan sendiri kesimpulan yang sesuai dengan data hasil pengamatan/ penyelidikan yang ada dan informasi yang ada dalam buku pelajaran yang berkaitan dengan benda tak hidup.	32. Hasil pengamatan menunjukkan dibandingkan batu sungai, batu bata lebih lunak, warnanya juga lebih cerah, dan bangunnya teratur yaitu berbentuk balok. Hasil dari membaca buku ternyata batu bata dibuat untuk membangun bangunan tembok rumah. Di kawasan tertentu yang tidak tersedia batu bata juga digunakan batu kali untuk membuat tembok rumah. Kesimpulan yang dapat ditarik yaitu batu bata digunakan sebagai batu buatan .... a. pengganti batu sungai* b. tandingan batu sungai
2.1.1	Merumuskan sendiri dugaan/ prakiraan sementara dalam suatu percobaan berdasarkan hasil percobaan/ penyelidikan tentang gejala benda tak hidup.	33. Jika kamu menjemur baju basah di bawah sinar matahari maka pakaian tersebut akan …. a. tetap basah b. menjadi kering*

Kode	Kelompok Indikator II-IV	Soal dengan dua pilihan jawaban
2.1.2	Menemukan sendiri cara mengantisipasi risiko serta mengambil tindakan pencegahan dan cara kerja yang sesuai dalam melakukan percobaan/ penyelidikan tentang gejala pada benda tak hidup.	34. Agar tangan tidak terkena minyak oli ketika kamu praktik IPA maka yang terbaik kamu menge-nakan sarung tangan yang terbuat dari …. a. kain b. karet*
2.2.4	Memilih sendiri hal-hal yang dapat mempengaruhi hasil suatu percobaan/ penyelidikan, menentukan sendiri data yang diperoleh sesuai hasil percobaan/ penyelidikan dan memilih sendiri cara menyajikan data hasil percobaan tentang gejala benda tak hidup.	35. Saya akan melakukan percobaan untuk mengetahui akibat lama penjemuran terhadap tingkat kekeringan kain maka ada banyak hal yang harus saya siapkan, diantaranya yaitu akan saya pilih …. a. lokasi penjemuran yang berbeda-beda, kemudian setelah data saya peroleh akan saya sajikan dalam bentuk tabel b. satu lokasi penjemuran, kemudian setelah data saya peroleh akan saya sajikan dalam bentuk diagram*

Prof. Bambang Subali, M.S.
Prof. Kumaidi, M.A., Ph.D.
Dr. Nonoh Siti Aminah, M.Pd.

Pengembangan Model Asesmen Konfirmatori Penguasaan Metode Ilmiah Pada Mata Pelajaran IPA di Sekolah Dasar

ISBN 602498019-1
9786024980191

UNY PRESS
Jl.H.Affandi (Jl.Gejayan), Gg. Alamanda,
Kompleks FT-UNY, Kampus Karangmalang, Yogyakarta,
Kode Pos: 55281,Telp.(0274)589346,
unypress.yogyakarta@gmail.com