

PENGEMBANGAN INSTRUMEN DIAGNOSTIK KOGNITIF PADA MATA PELAJARAN IPA DI SMP

¹⁾Yuli Prihatni, ²⁾Kumaidi, ³⁾Mundilarto

¹⁾Universitas Sarjanawiyata Tamansiswa, ²⁾Universitas Muhammadiyah Surakarta,

³⁾Universitas Negeri Yogyakarta

¹⁾yuliku7781@gmail.com, ²⁾kuma_426@yahoo.com, ³⁾mundilarto@uny.ac.id

Abstrak

Penelitian ini bertujuan untuk: (1) menghasilkan instrumen diagnostik kognitif pada mata pelajaran IPA materi kalor berdasarkan *learning continuum*, (2) menemukan karakteristik instrumen diagnostik kognitif pada mata pelajaran IPA materi kalor, berdasarkan *learning continuum*. Metode yang digunakan dalam penelitian ini adalah metode pengembangan tes diagnostik. Penelitian didahului dengan penyusunan *learning continuum*, kisi-kisi tes, hierarki materi prasyarat dan spesifikasi *item* pada materi IPA konsep kalor yang ditelaah melalui *Focus Group Discussion* (FGD). Uji coba dilaksanakan dengan subjek uji sebanyak 484 siswa kelas VII SMP di Kabupaten Sleman. Seleksi butir tes untuk mendapatkan fit tes menggunakan program *Quest* dan Program M Plus untuk uji kecocokan/kesesuaian model atribut dengan matriks Q. Hasil penelitian menunjukkan: (1) instrumen yang berhasil dikembangkan berbentuk tes pilihan ganda dengan alasan, berjumlah 28 butir soal dan menghasilkan tujuh Q matrik; (2) analisis dengan program *Quest* diperoleh butir fit dengan model dan program M Plus menemukan besarnya probabilitas pada setiap *latent class* pada tujuh Q matriks yang tersusun.

Kata kunci: *instrumen, diagnostik kognitif, IPA*

DEVELOPING THE INSTRUMENTS OF COGNITIVE DIAGNOSTIC FOR SCIENCE SUBJECT IN JUNIOR HIGH SCHOOL

¹⁾Yuli Prihatni, ²⁾Kumaidi, ³⁾Mundilarto

¹⁾Universitas Sarjanawiyata Tamansiswa, ²⁾Universitas Muhammadiyah Surakarta,

³⁾Universitas Negeri Yogyakarta

¹⁾yuliku7781@gmail.com, ²⁾kuma_426@yahoo.com, ³⁾mundilarto@uny.ac.id

Abstract

The study aimed at: (1) producing cognitive diagnostic instruments for science subjects of heat topic based on learning continuum (2) finding the characteristics of the instrument developed. The method used in this research is the development of diagnostic tests. The research was preceded by the preparation of learning continuum, test blueprint, hierarchy of prerequisite materials and test specification on materials of science concept which were examined through Focus Group Discussion (FGD). The test trial was implemented with 484 students of class VII SMP in Sleman as subjects. Selection of test items to get fit tests used Quest program and M Plus program to test the suitability/fitness model with the attribute matrix Q. The results show: (1) the instruments developed in the form of a multiple choice test with reason, totaling 28 items and seven Q matrix; (2) analysis using Quest program finds item fit with the model, and program M Plus find the probability in each latent class on seven Q matrix composed.

Keywords: *instruments, cognitive diagnostic, science*

Pendahuluan

Pelaksanaan pembelajaran Ilmu Pengetahuan Alam (IPA) di sekolah tidak lepas dari hakikat IPA itu sendiri. Fisika merupakan bagian dari IPA yang mempelajari gejala-gejala alam dan interaksinya. Pembelajaran IPA fisika adalah pembelajaran yang tidak mengabaikan hakikat IPA fisika sebagai sains (Taufik, Sukmadinata, Abdulkhak, & Tumbelaka 2010, p.33). Pelajaran fisika apabila melalui proses yang benar dapat melatih siswa mengembangkan daya pikir kritis dan sikap ilmiah. Sebagai ilmu pengetahuan yang mendasarkan pada pengamatan dan pengukuran, fisika dibawa kepada kecermatan, ketepatan, dan ketakterdugaan. Dalam mempelajari fisika, banyak siswa yang mengalami salah konsep (Suparno, 1998, p.96). Siswa cenderung hafal rumus fisika tanpa memahami konsepnya.

Rata-rata skor prestasi sains siswa Indonesia pada TIMSS 2007 adalah 433. Dengan skor tersebut siswa Indonesia menempati peringkat 35 dari 49 negara. Rata-rata skor siswa Indonesia pada TIMSS 2007 di bawah skor rata-rata yaitu 500, dan hanya mencapai *Low International Benchmark*. Capaian tersebut menunjukkan rata-rata siswa Indonesia hanya mampu mengenali sejumlah fakta dasar tetapi belum mampu mengkomunikasikan dan mengaitkan berbagai topik sains, apalagi menerapkan konsep-konsep yang kompleks dan abstrak (Efendi, 2010, p.385).

Kesulitan dan kelemahan siswa dalam bidang sains penting untuk dipecahkan. IPA merupakan ilmu pengetahuan alam yang berkaitan dengan cara mencari tahu tentang alam secara sistematis, sehingga IPA bukan hanya penguasaan kumpulan pengetahuan yang berupa fakta-fakta, konsep-konsep atau prinsip-prinsip saja tetapi juga merupakan suatu proses penemuan. Kemampuan guru dalam mendeteksi kesulitan dan kelemahan siswa merupakan tantangan dalam kegiatan pembelajaran IPA untuk mengantarkan siswa mencapai keberhasilan dalam belajar. Setiap siswa dalam kelas mempunyai karakteristik yang berbeda-beda. Perbedaan individual siswa dapat disebabkan pada cara ber-

pikirnya, oleh karena itu perbedaan individu perlu diperhatikan oleh guru dalam kegiatan pembelajaran. Pelaksanaan pembelajaran di kelas umumnya dilaksanakan dengan melihat siswa sebagai individu dengan kemampuan rata-rata dan kebiasaan yang kurang lebih sama. Kemampuan individual siswa kurang diperhatikan, sehingga siswa yang memiliki kelemahan akan terus tertinggal dalam pembelajaran. Proses belajar antara siswa yang memiliki kemampuan tinggi tentu tidak sama dengan siswa yang memiliki kemampuan sedang, bahkan dengan siswa yang kemampuannya kurang. Guru perlu memperhatikan individu siswa, kelemahan dan kesulitan siswa perlu diperhatikan. Pembelajaran IPA saat ini lebih menekankan pada runtutan materi pelajaran bukan pada proses berpikir dan psikologi kognitif siswa, sehingga dalam pembelajaran siswa banyak yang mengalami kesulitan dan salah konsepsi. Guru memegang peranan penting untuk mengatasi kesulitan siswa dan memperbaiki proses pembelajaran.

Kesulitan siswa dalam belajar erat kaitannya dengan perkembangan kognitif siswa. Menurut teori kognitif, belajar menunjukkan adanya jiwa yang sangat aktif, jiwa mengolah informasi yang diterima, tidak sekedar menyimpannya saja tanpa mengadakan transformasi (Gage & Berliner, 1984, p.267). Menurut teori ini anak memiliki sifat aktif, konstruktif, dan mampu merencanakan sesuatu. Anak mampu mencari, menemukan fakta, menganalisis, menafsirkan dan menarik kesimpulan. Berpikir pada umumnya diasumsikan sebagai proses kognitif, tindakan mental untuk memperoleh pengetahuan. Produk-produk berpikir seperti pikiran, pengetahuan, alasan serta proses yang lebih tinggi seperti penilaian dapat juga dihasilkan. Kaitan-kaitan kompleks dikembangkan melalui berpikir ketika digunakan sebagai bukti dari waktu ke waktu. Kaitan-kaitan ini dapat dihubungkan pada struktur yang terorganisasi dan diekspresikan oleh pemikir dalam beragam cara. Membangun model kognitif untuk mendesain ciri soal merupakan hal penting dalam pendekatan desain kognitif (Mislevy, 1994, pp.439-483).

Oleh karena itu, dalam penelitian ini penyusunan instrumen diagnostik kognitif memperhatikan proses berpikir siswa.

Pengembangan instrumen diagnostik berkaitan dengan kegiatan penilaian. Pada Permendiknas No 20 tahun 2007 tentang standar penilaian dijelaskan bahwa penilaian adalah proses pengumpulan dan pengolahan informasi untuk menentukan pencapaian hasil belajar siswa. Penilaian tidak sekedar pengumpulan data siswa, tetapi juga pengolahannya untuk memperoleh gambaran proses dan hasil belajar siswa. Penilaian tidak sekedar memberi soal siswa kemudian selesai, tetapi guru harus menindaklanjutinya untuk kepentingan pembelajaran. Dunn, Morgan, Parry, & Reilly (2004, p.16) menyatakan bahwa tujuan dan peranan penilaian (*assessment*) dalam pembelajaran yaitu: (1) mendiagnosis kesulitan belajar siswa; (2) mengukur peningkatan dari waktu ke waktu; (3) menentukan penguasaan siswa terhadap pengetahuan dan keterampilan tertentu; (4) menentukan ranking siswa dari keseluruhan siswa dalam kelas; (5) mengevaluasi metode pembelajaran; (6) mengevaluasi efektivitas program pembelajaran.

Penilaian merupakan bagian terpenting dari proses pembelajaran. Tujuan penilaian memberikan masukan yang berupa informasi secara komprehensif tentang hasil belajar siswa, baik dilihat ketika saat kegiatan pembelajaran berlangsung maupun dilihat dari hasil akhirnya, dengan menggunakan berbagai cara penilaian sesuai dengan kompetensi yang diharapkan dapat dicapai siswa. Untuk menilai sejauhmana siswa telah menguasai beragam kompetensi, diperlukan berbagai jenis penilaian. Dengan demikian, penilaian merupakan rangkaian kegiatan untuk memperoleh data, menganalisis, dan menafsirkan data tentang proses dan hasil belajar siswa yang dilakukan secara sistematis dan berkesinambungan, sehingga menjadi informasi yang bermakna dalam pengambilan keputusan

Penilaian di kelas dilakukan oleh guru bertujuan untuk mengetahui kemajuan dan hasil belajar siswa, membandingkan kemampuan siswa, mendiagnosa kesulitan belajar,

memberikan umpan balik/perbaiki proses belajar mengajar, dan penentuan kenaikan kelas. Melalui penilaian dapat diperoleh informasi yang akurat tentang penyelenggaraan pembelajaran dan keberhasilan belajar siswa, guru, serta proses pembelajaran itu sendiri. Berdasarkan informasi itu, dapat dibuat keputusan tentang pembelajaran, kesulitan siswa dan upaya bimbingan yang diperlukan serta keberadaan kurikulum itu sendiri. Penilaian memiliki tujuan yang sangat penting dalam pembelajaran, diantaranya untuk *grading*, seleksi, mengetahui tingkat penguasaan kompetensi, bimbingan, diagnosis, dan prediksi.

Seorang guru harus memberikan peluang bagi siswa dan guru untuk mendapatkan dan menggunakan informasi tentang kemajuan menuju tujuan pembelajaran. Banyak kegiatan atau aktivitas yang terjadi di dalam kelas dilakukan guru dan siswa yang dapat digambarkan sebagai penilaian. Artinya, tugas dan pertanyaan siswa dapat digunakan untuk menunjukkan pengetahuan mereka, pemahaman dan keterampilan yang dimiliki siswa. Apa yang dilakukan siswa kemudian diamati dan diinterpretasikan dan dapat digunakan untuk memperbaiki proses pembelajaran. Proses penilaian ini merupakan bagian penting dari praktek di kelas yang dilaksanakan sehari-hari dan melibatkan guru dan siswa dalam refleksi, dialog dan pengambilan keputusan (Leung & Mohan 2004, pp.335-359).

Menurut Nitko (1996, pp.284-285) pada praktiknya penilaian diagnostik di kelas memiliki dua kawasan yaitu: (1) untuk mengidentifikasi target pembelajaran yang belum dikuasai siswa; dan (2) untuk menemukan penyebab-penyebab atau alasan-alasan yang membuat siswa belum dapat menguasai target-target pembelajaran. Penilaian yang dilakukan dalam penelitian ini menggunakan instrumen tes.

Penilaian dalam penelitian ini dilakukan dengan menggunakan instrumen yang berupa tes diagnostik. Dalam penelitian ini penilaian dikhususkan sebagai alat diagnosis kognitif siswa pada mata pelajaran IPA SMP, penilaian digunakan untuk menunjukkan

letak kelemahan dan kesulitan belajar yang dialami siswa dan kemungkinan prestasi yang bisa dikembangkan. Hal ini akan membantu guru untuk memberikan pembelajaran yang sesuai dengan proses berpikir siswa.

Mengamati kesalahpahaman dalam pembelajaran memainkan peran penting dalam pengembangan kemampuan mental anak-anak. Oleh karena itu, meningkatkan kemampuan belajar anak merupakan aspek penting dari mendidik anak (Afaq & Al-Mashari, 2010, p.76). Dengan demikian, perlu dirancang tes diagnostik yang dapat mengidentifikasi kelemahan dan kekuatan siswa dalam belajar. Tes diagnostik menurut Ebel (1979, p.375) adalah rancangan untuk mengetahui kekurangan-kekurangan yang khusus atau kegagalan dalam belajar dalam beberapa subjek atau pelajaran seperti membaca dan aritmatika. Pendapat senada dinyatakan oleh Gronlund yang menyatakan bahwa tes diagnostik adalah tes yang dirancang untuk mengetahui sebab kegagalan siswa dalam mengajar. Tes diagnostik memiliki dua fungsi utama, yaitu, mengidentifikasi masalah atau kesulitan yang dialami siswa, dan merencanakan tindak lanjut berupa upaya-upaya pemecahan sesuai masalah atau kesulitan yang telah teridentifikasi.

Weeden, Winter & Brounfoot (2002, p.20) menyatakan bahwa tes diagnostik adalah tes yang digunakan untuk mengidentifikasi masalah-masalah spesifik yang dialami siswa. Tes diagnostik dapat menunjukkan dan memetakan kelemahan-kelemahan siswa dan ketidakajegan pengetahuan. Tes diagnostik seharusnya mampu mengungkap mengapa siswa memberikan respon seperti yang mereka lakukan (Gorin, 2007, p.174). Sesungguhnya tes diagnostik juga dapat membantu penilaian guru dalam memperdalam pemahaman konseptual siswa mereka dari topik tertentu (Russell, 2009, p.414). Dengan tes diagnostik, kelemahan siswa dapat dideteksi dan guru diharapkan dapat memberi bantuan yang tepat sesuai dengan apa yang dibutuhkan oleh siswa bila siswa mengalami kesulitan.

Kesulitan siswa dalam pembelajaran penting untuk dipecahkan. Dengan demikian, belajar dapat menghasilkan perubahan-perubahan dalam memperoleh kemajuan pe-ngetahuan pemahaman, ketrampilan dan nilai. Belajar menurut psikologi kognitif berarti belajar dengan memperhatikan tahapan-tahap proses berpikir yang mengacu pada psikologi kognitif. Anderson & Krathwohl (2001, p.39) menetapkan empat jenis pengetahuan yakni pengetahuan faktual, konseptual, prosedural dan metakognitif. Pengetahuan faktual meliputi elemen dasar yang harus diketahui siswa untuk mempelajari disiplin ilmu atau untuk menyelesaikan masalah-masalah dalam disiplin ilmu tersebut. Pengetahuan konseptual meliputi hubungan-hubungan antar elemen dalam sebuah struktur besar yang memungkinkan elemen-elemennya berfungsi secara bersama-sama. Pengetahuan prosedural meliputi bagaimana melakukan sesuatu, mempraktikkan metode-metode penelitian dan kriteria-kriteria untuk menggunakan ketrampilan, algoritma, teknik dan metode. Pengetahuan metakognitif meliputi pengetahuan tentang kognisi secara umum dan kesadaran dan pengetahuan tentang kognisi diri sendiri.

Pengembangan tes diagnostik harus memenuhi dua asumsi yaitu (1) tes mampu menganalisa keterampilan atau pengetahuan untuk dimasukkan ke dalam komponen sub-keterampilan atau sub-pengetahuan, (2) butir tes yang dikembangkan mampu mengukur sub-ketrampilan atau subpengetahuan tersebut (Mehrens & Lehmann, 1984, p.410). Dengan tes diagnostik, kemampuan siswa dapat dilihat pada level/tahapan pengetahuannya. Tahapan pengetahuan siswa dapat diketahui apabila *learning continuum* sudah dirumuskan. *Learning continuum* yaitu urutan logis yang menunjukkan keterkaitan materi secara vertikal dalam pelajaran. Pada kompetensi-kompetensi yang akan diukur diidentifikasi sejumlah kompetensi yang harus dikuasainya terlebih dahulu, kompetensi yang harus dikuasai adalah kemampuan pra-syarat.

Penelitian tentang tes diagnostik di Indonesia sudah mulai berkembang, Kusaeri (2012) mampu mengembangkan instrumen

tes diagnostik dengan model DINA menggunakan studi simulasi, membangun *syntak* dengan menggunakan CDM. Proses membangkitkan data dan analisis dalam penelitian ini menggunakan paket *software* Mplus dari Muthen & Muthen. Hasil studi simulasi menunjukkan bentuk matriks ideal. Suatu matriks dikatakan ideal bila matriks tersebut menghasilkan struktur *latent class* yang mak-simum. Penelitian Duskri, Kumaidi & Suryanto (2014, p.44) mengembangkan tes diagnostik kesulitan belajar matematika di SD. Penelitian ini dapat memunculkan informasi hasil tes secara klasikal dan individual, grafik ketuntasan belajar, profil individual, analisis salah konsepsi dan saran remedial. Namun, keterbatasan penelitian ini adalah *software* komputer yang dirancang untuk melakukan analisis hasil tes diagnostik kesulitan belajar masih berbasis *Microsoft excel* 2007 dan dapat dikembangkan dengan program lain yang lebih interaktif. Penelitian Isgiyanto (2011) menggunakan atribut dalam penyusunan instrumen, namun atribut yang tersusun tidak diidentifikasi *latent class* nya.

Tes diagnostik yang dikembangkan dalam penelitian ini adalah tes diagnostik kognitif pada mata pelajaran IPA SMP materi kalor. Pengembangan tes berdasar pada *Learning Continuum*, dan menggunakan atribut dalam penyusunan matriks Q dan diidentifikasi *latent class* nya. Kegunaan *Learning Continuum* menurut *Northwest Evaluation Association* (2003, p.4) adalah sebagai panduan untuk mengetahui perbedaan tingkat prestasi siswa. Pendidik harus menyadari bahwa perbedaan prestasi dan tingkat pencapaian siswa saat ini harus menjadi pertimbangan dominan ketika pengelompokan dalam memberikan instruksi dan memilih bahan/materi. *Learning Continuum* akan membantu guru mengetahui kelemahan siswa yang disebabkan karena siswa belum menguasai materi prasyarat dan akan mengembangkan potensi siswa apabila siswa telah menguasai kompetensi yang ditetapkan. Atribut didefinisikan sebagai suatu prosedur proses atau kemampuan siswa untuk menyelesaikan suatu *item* (Gierl, 2007,

p.327). Atribut adalah karakteristik laten dari peserta tes/responden (Rupp, Templin & Henson, 2010, p.315).

Tes diagnostik memiliki karakteristik yang berbeda dengan butir soal tes yang lain. Pada tes diagnostik jawaban atau respons yang diberikan oleh siswa harus memberikan informasi yang cukup untuk menduga masalah atau kesulitan yang dialaminya (memiliki fungsi diagnosis). Tes diagnostik dirancang dengan memperhatikan kognitif siswa. Tes diagnosis kognitif yang dikembangkan mengacu pada proses kognitif Bloom yang dikembangkan Anderson & Krathwohl (2001, p.39) dengan menggunakan tahapan vertikal dan tahapan horizontal. Tahapan vertikal yaitu dengan memperhatikan dimensi pengetahuan yang meliputi pengetahuan faktual, pengetahuan konseptual dan pengetahuan prosedural. Pengetahuan metakognitif tidak dikembangkan dalam penelitian ini dikarenakan merupakan pengetahuan diri. Dalam memecahkan masalah sains beberapa pengetahuan dapat berjalan bersama-sama pada proses kognitif yang berbeda. Proses kognitif merupakan tahapan secara horizontal, dalam tahapan ini terdiri dari: mengingat, memahami, mengaplikasikan, dan menganalisis.

Tujuan penelitian ini adalah (1) menghasilkan tes diagnostik kognitif pada mata pelajaran IPA materi kalor berdasarkan *learning continuum*; (2) menemukan karakteristik tes diagnostik kognitif pada mata pelajaran IPA materi kalor, berdasarkan *learning continuum*.

Metode Penelitian

Penelitian ini menggunakan metode penelitian pengembangan tes diagnostik yang bertujuan untuk penilaian kognitif menurut Nichols (1994, p.587) yaitu: (1) mengkaji konstruksi teori yang substansif sebagai hal yang mendasar dalam pengembangan tes berdasar penelitian atau *review* penelitian; (2) melakukan seleksi desain pengukuran untuk mengkonstruksi butir yang dapat direspon oleh peserta tes berdasarkan pengetahuan, ketrampilan yang spesifik atau karakteristik lain sesuai teori; (3) melakukan administrasi

tes yang meliputi beberapa aspek diantaranya menyusun bentuk tes, teknologi yang digunakan untuk membuat tes, analisis situasi lingkungan pada waktu pengetesan, membuat kisi-kisi tes, menulis tes dan sebagainya; (4) membuat skoring hasil tes yaitu penentuan nilai tes yang telah dilakukan melalui proses analisis butir soal; (5) melakukan revisi yaitu proses penyesuaian dan perbaikan sesuai dengan tujuan tes diagnostik dan standar yang diharapkan.

Tes diagnostik kognitif dikembangkan sesuai dengan fokus penelitian melalui tiga tahapan yaitu (1) tahap perencanaan tes, (2) tahap uji coba, dan (3) tahapan pengembangan *prototype software computer based testing* agar penelitian ini dapat langsung diaplikasikan di sekolah. Penelitian pengembangan secara umum, prosesnya mencakup: membangun kerangka konseptual, diikuti dengan merancang sistem, dan berakhir dengan membangun prototipe untuk pengujian dan evaluasi (Ellis & Levy, 2010). Tahap perencanaan dimulai dengan studi literatur (kaji pustaka dan hasil penelitian terdahulu) dan studi lapangan. Studi literatur dilakukan untuk memahami lebih jauh tentang tes diagnostik kognitif. Diperlukan teori-teori baru yang menjelaskan penyebab yang mendasari masalah (Ellis & Levy, 2010, p.108). Pada studi literatur ini peneliti mengumpulkan temuan riset dan informasi yang mendukung terhadap produk yang direncanakan. Penyusunan instrumen tes diagnostik kognitif pada mata pelajaran IPA Fisika dimulai dengan kegiatan *workshop* yang diikuti oleh 6 orang guru yang tergabung dalam MGMP IPA kabupaten Sleman, pengawas dan pakar pengukuran. Pada tahap ini peneliti dan peserta *workshop* merumuskan *learning continuum* penguasaan konsep materi IPA SMP. *Learning continuum* yang dibuat diharapkan dapat menggambarkan perkembangan dan peningkatan kemampuan siswa untuk menguasai konsep IPA Fisika mulai dari konsep dasar menuju konsep lanjut. *Learning continuum* dapat digunakan untuk mengidentifikasi kompetensi dasar dan merumuskan indikator materi IPA SMP yaitu materi Kalor.

Tes yang dibuat dalam penelitian ini adalah tes diagnostik kognitif yang bertujuan untuk mengetahui kelemahan-kelemahan siswa dalam mata pelajaran IPA fisika materi kalor. Hasil tes diagnostik kognitif dapat digunakan sebagai dasar untuk memberikan tindak lanjut berupa perlakuan yang tepat dan sesuai dengan kelemahan yang dimiliki siswa. Agar tujuan tes dapat tercapai maka pada tahapan ini juga disusun atribut tes diagnostik kognitif pada materi kalor. Atribut adalah proses kognitif atau ketrampilan dasar yang dibutuhkan untuk memecahkan tes. Atribut dapat dilihat sebagai sumber dari kompleksitas kognitif pada hasil tes. Penyusunan atribut penting dilakukan sebagai landasan untuk menggambarkan kemampuan kognitif peserta tes. Atribut pada instrumen tes diagnostik kognitif ini dibuat berdasarkan pada *learning continuum* dan hierarki materi prasyarat. Pemilihan atribut dilakukan secara hati-hati karena berimplikasi pada proses mengkonstruksi Q matriks.

Bentuk tes yang digunakan dalam penelitian ini adalah bentuk tes objektif dengan beralasan. Dengan dilengkapi alasan, akan diketahui cara berfikir siswa dan langkah-langkah siswa dalam menyelesaikan soal. Siswa yang benar-benar menguasai akan dapat memberikan alasan jawabannya. Kelebihan tes obyektif adalah sebagai berikut: (a) jumlah materi yang dapat ditanyakan relatif lebih luas dibandingkan dengan materi yang dapat dicakup soal bentuk lainnya. Jumlah soal yang ditanyakan umumnya relatif banyak; (b) dapat mengukur berbagai jenjang; (c) penskorannya mudah, cepat, objektif, dan dapat mencakup ruang lingkup bahan dan materi yang luas dalam satu tes untuk suatu kelas atau jenjang; (d) sangat tepat digunakan apabila pesertanya banyak; (e) reliabilitas soal pilihan ganda relatif lebih tinggi dibandingkan dengan soal uraian. Namun tes obyektif juga mempunyai kelemahan salah satunya yaitu hanya dapat mengukur sampai dengan tahap analisis. Oleh karena itu dalam penelitian ini proses kognitif yang diukur sampai pada tahap analisis.

Uji coba penelitian pengembangan tes diagnostik kognitif pada mata pelajaran IPA

materi kalor dilaksanakan pada 484 siswa di Tujuh SMP yang mempunyai karakteristik berbeda di Kabupaten Sleman yaitu SMPN 1 Sleman, SMPN 1 Kalasan, SMPN 1 Turi, SMPN 4 Tempel, SMPN 4 Gamping, SMPN 2 Cangkringan dan SMP Muhammadiyah Pakem. Sampel uji coba harus mempunyai karakteristik yang kurang lebih sama dengan karakteristik peserta tes yang sesungguhnya. Ujicoba ini di gunakan sebagai sarana untuk memperoleh data empirik tentang tingkat kebaikan soal yang telah disusun. Melalui ujicoba data diperoleh tentang: realibilitas, validitas, tingkat kesukuran, pola jawaban, efektivitas pengecoh, daya beda dan lain-lain (Mardapi, 2008, p.95).

Analisis data pengembangan instrumen tes, pada tahap pertama menggunakan validasi isi produk oleh para ahli (*experts*) melalui proses FGD untuk mengetahui sejauh mana naskah tes disagnostik kognitif mata pelajaran IPA SMP mencerminkan tujuan pengembangan. Analisis data secara kuantitatif dengan bantuan program *Quest*, CDM dan menggunakan Program *M Plus*. Program *Quest* digunakan untuk mendapatkan parameter butir soal yaitu nilai tingkat kesulitan butir untuk melengkapi informasi diagnostik kognitif pada instrumen yang telah disusun yaitu dengan mengkaitkan tahap *learning continuum* dan dimensi kognitif. Program *Quest* dapat membandingkan *ability* dan *difficulty* yang tidak dapat diketahui apabila dianalisis menggunakan analisis secara klasik. Dengan demikian, hasil analisis program *Quest* dapat membantu memberikan informasi diagnostik tentang kelemahan dan kesulitan siswa. Program *Quest* juga dapat digunakan untuk menentukan *fit item* dengan model. Dengan demikian program *Quest* dapat membuat *item* yang terstandar. Batas suatu *item* dinyatakan *fit* dengan model jika memiliki infit MNSQ diantara 0,77 sampai dengan 1,30 (Adam & Khoo, 1996, p.30). Persyaratan Fit statistik pada program *Quest*, yakni jika nilai rata-rata infit MNSQ mendekati 1,0 dengan simpangan baku 0,00.

Program CDM dalam penelitian ini digunakan untuk membangkitkan *syntak* M Plus. Program M Plus digunakan untuk me-

nemukan informasi yang lebih banyak tentang *latent class*, dengan cara menyimulasikan ukuran matriks Q dan banyaknya peserta tes. Penelitian Simulasi menggunakan studi simulasi *Monte Carlo (M Plus)*. Pengklasifikasian profil atribut dilakukan untuk melihat kecenderungan pola jawaban peserta tes diagnostik kognitif sehingga dapat diketahui atribut-atribut yang telah dikuasai oleh kebanyakan peserta tes dan atribut-atribut yang belum dikuasai oleh peserta tes. Hal ini disampaikan oleh Rupp, Templin & Henson (2010, p.232) bahwa mengestimasi parameter peserta tes sama halnya dengan mengklasifikasi profil atribut peserta tes.

Hasil Penelitian dan Pembahasan

Model Instrumen diagnostik kognitif yang dikembangkan berupa (1) *Learning Continuum* pada materi kalor mata pelajaran IPA SMP yang menjadi landasan dalam penyusunan kisi-kisi tes, (2) hierarki materi prasyarat mata pelajaran IPA SMP pada materi kalor, (3) *Item Specification* yang dikembangkan berisi tentang uraian yang menunjukkan karakteristik diagnostik kognitif yang harus dimiliki setiap butir tes, (4) atribut soal dan Q matrik, (5) perangkat tes diagnostik kognitif Mata Pelajaran IPA SMP yang terdiri atas 28 soal pilihan ganda dengan alasan dan (6) *software* yang dilengkapi panduan Tes Diagnostik Kognitif pada Mata Pelajaran IPA SMP khususnya pada materi kalor. *Software* komputer yang dikembangkan merupakan program aplikasi berbasis *website* yang menggunakan pemograman HTML dipadu dengan PHP dan menggunakan *database* jenis *MySQL*.

Secara keseluruhan instrumen hasil validasi para pakar mempunyai nilai baik dan sangat baik. Data yang diperoleh dari hasil FGD kemudian disusun dalam Q matriks dan atribut tes diagnostik kognitif mata Pelajaran IPA SMP. Atribut bersifat hierarkis, hal ini menggambarkan kemampuan prasyarat secara langsung antar atribut (Gierl, 2007, p.328). Pada penelitian ini atribut yang disusun berjumlah 17 dan disusun dalam 7 Q matrik.

Data yang sudah terkumpul di validasi secara empirik menggunakan *Rasch model* melalui program *Quest*. Seleksi butir dengan program *Quest* digunakan untuk mendapatkan nilai tingkat kesulitan butir dan fit tes. Analisis dengan *Rasch Model* digunakan untuk mengukur satu konstruks teori pada satu waktu pada hierarki skala logit. Batas suatu *item* dinyatakan fit dengan model jika memiliki infit MNSQ diantara 0,77 sampai dengan 1,30 (Adam & Khoo, 1996, p.30). Persyaratan Fit statistik pada program *Quest*, yakni jika nilai rata-rata infit MNSQ mendekati 1,0 dengan simpangan baku 0,00. Instrumen tes diagnostik kognitif pada mata pelajaran IPA memiliki jumlah butir soal yang dianalisis sebanyak 28 butir.

Pada penelitian ini 27 *item* berada di dalam garis vertikal atau memiliki infit MNSQ diantara 0,77 sampai dengan 1,30. Butir nomor 7 berada pada infit MNSQ 1,31. Namun karena butir soal nomor 7 tidak berada jauh dari garis vertikal, dan mempunyai selisih infit MNSQ 0,01 maka pada penelitian ini butir 7 tetap digunakan dengan direvisi. Instrumen penelitian ini telah memenuhi persyaratan Fit statistik pada program *Quest*, yakni dengan nilai rata-rata infit MNSQ 1,00 dengan simpangan baku 0,11. Nilai reliabilitas sampel pada penelitian ini sebesar 0,99 Menurut Wright & Masters (1982, p.106) Nilai reliabilitas berdasarkan estimasi *item* disebut reliabilitas sampel. Dengan demikian, semakin tinggi nilainya, semakin banyak *item* yang fit dengan model. Sedangkan reliabilitas tes dalam penelitian ini diperoleh sebesar 0,72 yaitu dengan melihat nilai reliabilitas berdasarkan estimasi *case* atau testi. Nilai reliabilitas berdasarkan estimasi *case/teste* disebut reliabilitas tes (Wright & Masters, 1982). Reliabilitas tes sebesar 0,72 memberikan informasi bahwa pengukuran menggunakan instrumen tes diagnostik kognitif pada pelajaran IPA memberikan hasil yang konsisten. Gronlund (1982) menyatakan besarnya indeks reliabilitas suatu alat ukur khususnya pada pengukuran pendidikan seperti tes buatan guru sekitar 0,6-0,85. Semakin tinggi nilai reliabilitas tes, semakin banyak sampel untuk uji coba yang memberikan informasi

yang diharapkan, begitu juga sebaliknya. Semakin rendah nilai reliabilitas tes maka semakin sedikit sampel untuk ujicoba yang memberikan informasi yang diharapkan.

Uji kecocokan fit kesesuaian model atribut dengan Q matriks dengan menggunakan program M Plus. Templin (2009) mengelompokan uji fit model dalam tiga kelompok yaitu; Fit mutlak, fit relatif dan fit *item*. Uji Fit perlu dilakukan sebelum menginterpretasikan dan melakukan desiminasi hasil (Rupp, Templin & Henson, 2010, p.265). Uji fit model atribut dengan Q matriks pada penelitian ini menggunakan nilai *entropy*. Nilai *entropy* 1 berarti seluruh peserta tes diklasifikasikan secara tepat (*Good fit*) ke dalam kelasnya. Nilai *entropy* 0 berarti semua peserta tes memiliki peluang sama diklasifikasikan ke dalam semua kelas (*poor fit*). Idealnya dalam mendesain Q matriks harus berkorespondensi dengan struktur *latent class* (de la Torre & Karelitz, 2009, p.233). Nilai *entropy* pada Q matriks yang telah tersusun dalam penelitian ini terendah adalah pada Q matriks 7 sebesar 0,472 hal ini berarti sebesar 47,2,6 % peserta tes diklasifikasikan secara tepat (*good fit*) ke dalam kelasnya. Nilai *entropy* terbesar terdapat pada Q matriks 6 yaitu sebesar 0,812 berarti sebanyak 81,2% peserta tes diklasifikasikan secara tepat (*good fit*) ke dalam kelasnya.

$$Q6 = \begin{vmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{vmatrix} \quad Q7 = \begin{vmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{vmatrix}$$

Untuk menentukan Fit relatif menggunakan *information criteria*. Informasi ini digunakan untuk membandingkan dua Q matriks yang memiliki atribut yang berbeda. Dengan melihat nilai *Akaike Information Criteria* (AIC), *Bayesian Information Criteria* (BIC) dan *sample size adjusted BIC*. Nilai paling rendah dari *information criteria* menunjukkan paling baik memberikan struktur *latent class*. Menurut Nylund, Asparouhov, & Muthen (2007) penggunaan nilai BIC lebih tepat jika dibandingkan dengan *AIC*, hal ini dikarenakan hasil *AIC* seringkali lebih tinggi diban-

dingkan dengan nilai sebenarnya dan BIC lebih konsisten.

Pada Tabel 1 didapat bahwa Q matriks 6 mempunyai nilai *AIC*, *BIC* dan *Sample-Size Adjusted BIC* lebih rendah jika dibandingkan dengan Q matriks yang lain. Artinya ukuran matriks 3x3 lebih ideal dibandingkan matriks lain karena mampu memberikan informasi struktur *laten class* secara maksimum. Penelitian yang dilakukan oleh Kusaeri (2012, p.127) menyebutkan bahwa nilai paling rendah pada ketiga *information criteria*, menunjukkan bahwa Q matriks itu paling ideal karena mampu memberikan informasi struktur *latent class* secara maksimum. Pada penelitian ini Q matriks 6 dan Q matriks 7 adalah matriks dengan ukuran 3x3. Namun mempunyai susunan yang berbeda. Ukuran matriks 3x3 bukan merupakan satu-satunya yang menentukan baik tidaknya sebuah model namun juga dipengaruhi oleh pola dari matriks tersebut. Seperti yang disampaikan De la Torre, & Karelitz (2009, p.233) bahwa dengan memperhatikan aspek kesesuaian, maka dalam mendesain Q matrik harus berkorespondensi dengan struktur *latent class*.

Tabel 1. *Information Criteria* pada Q matrik

| Matrik | <i>Information Criteria</i> | | |
|--------|-----------------------------|-----------------------|---------------------------------|
| | <i>Akaike (AIC)</i> | <i>Bayesian (BIC)</i> | <i>Sample-Size Adjusted BIC</i> |
| Q1 | 2387,752 | 2450,484 | 2402,875 |
| Q2 | 2415,299 | 2478,031 | 2430,422 |
| Q3 | 2910,025 | 3014,577 | 2935,229 |
| Q4 | 1414,467 | 1477,199 | 1429,590 |
| Q5 | 2710,149 | 2814,701 | 2735,353 |
| Q6 | 1356,752 | 1411,119 | 1369,858 |
| Q7 | 1951,507 | 2005,874 | 1964,613 |

Pada penelitian ini Q1 terdiri atas 3 atribut dan 4 *item*, Q2 terdiri atas 3 atribut dan 4 *item*, Q3 terdiri atas 4 atribut dan 5 *item*, Q4 terdiri atas 3 atribut dan 4 *item*, Q5 terdiri atas 4 atribut dan 5 *item*, Q6 terdiri atas 3 atribut dan 3 *item* dan pada Q7 terdiri atas 3 atribut dan 3 *item*. De La Torre & Karelitz (2009) menyarankan agar banyak-

nya *item* di dalam Q matriks lebih dari banyaknya atribut. Rupp & Templin (2008, p.83) mengungkapkan bahwa setiap penambahan sejumlah atribut pada Q matriks seharusnya panjang tes juga ditambah agar dapat memberikan informasi yang reliabel tentang semua variabel laten.

Idealnya Q matriks dijadikan sebagai dasar dalam menyusun kisi-kisi selama proses penyusunan *item*. Seperti penelitian yang telah dilakukan Kusaeri (2012, p. 150) bahwa *item* tes yang tidak dibangun berdasarkan Q matrik, kurang ada kesesuaian antara Q matriks dengan struktur *latent class* sehingga akan mempengaruhi rendahnya nilai *entropy*. Pada penelitian ini Q matriks dikonstruksi berdasarkan hierarki materi prasyarat dan *learning continuum* sehingga terdapat Q matriks dengan *entropy* rendah.

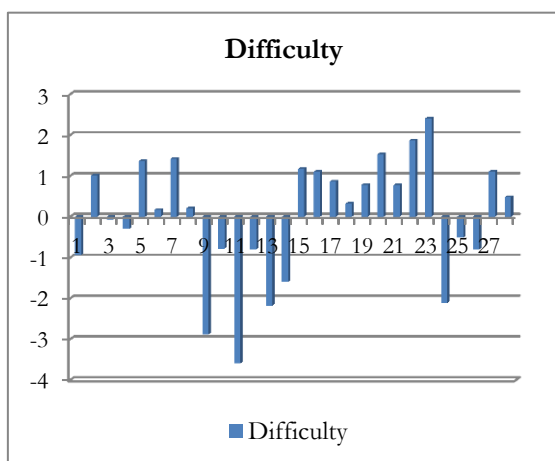
Untuk mengetahui Fit *item* dengan menggunakan program M Plus yaitu didasarkan pada dua informasi, yakni *univariate information* dan *bivariate information*. Sebuah *item* dikatakan fit apabila informasi yang diperoleh dari *univariate information*, data hasil observasi dengan data yang diprediksi model bernilai sama. Sementara *bivariate information* menghasilkan statistik fit dari pasangan *item*. Berdasarkan hasil analisis M Plus diketahui bahwa semua informasi yang diperoleh dari *univariate information*, data hasil observasi dengan data yang diprediksi model bernilai sama.

Hasil tes diagnostik kognitif yang telah diujicobakan di tujuh SMP di Kabupaten Sleman, diperoleh informasi bahwa pada soal jenis pengetahuan faktual memiliki rata-rata >70% pada seluruh sekolah yang menjadi subjek penelitian. Pada sekolah berkategori baik rata-rata siswa yang dapat mengerjakan soal jenis pengetahuan faktual >90% namun pada pengetahuan koseptual <65% dan pada pengetahuan prosedural <50%. Sekolah dengan kategori sedang, rata-rata siswa yang dapat mengerjakan soal pada jenis pengetahuan faktual >80%, namun pada pengetahuan koseptual <40% dan pada pengetahuan prosedural <30%. Pada sekolah dengan kategori kurang, rata-rata siswa yang dapat mengerjakan soal pada

jenis pengetahuan faktual sebesar >75% namun pada pengetahuan koseptual <40% dan pada pengetahuan prosedural <30%.

Pada jenis pengetahuan faktual diperoleh informasi diagnostik yaitu adanya siswa yang masih belum menguasai kemampuan prasyarat kalor (suhu dan zat), dan tidak dapat membedakan antara menyublim dan mengkristal. Pada jenis pengetahuan konseptual diperoleh informasi diagnostik bahwa masih banyak siswa yang belum dapat menjelaskan konsep perpindahan kalor, menganalisis perubahan suhu dan wujud zat menggunakan gambar, dan belum dapat menganalisis penerapan manfaat perpindahan kalor. Pada jenis dimensi prosedural diperoleh informasi diagnostik bahwa masih banyak siswa yang belum dapat menentukan jumlah kalor pada zat yang diketahui kalor jenisnya, menentukan kapasitas kalor pada zat yang diketahui massa dan kalor jenisnya, mengukur kenaikan suhu pada zat yang dipanaskan dengan kalor tertentu, menganalisis proses perubahan wujud zat menggunakan grafik dan menganalisis hukum kekekalan energi pada persamaan kalor (Azas Black).

Tingkat kesukaran butir soal dapat menambah informasi diagnostik. Dari 28 soal pilihan ganda dengan alasan yang dikerjakan oleh siswa diperoleh tingkat kesukaran butir yang disajikan pada Gambar 1.



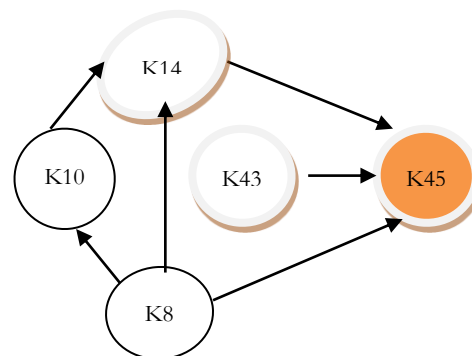
Gambar 1. *Item Difficulty* Tes Diagnostik

Pada Gambar 1 dapat dilihat bahwa soal yang paling sulit adalah *item* soal nomor 23 dan *item* soal paling mudah adalah *item*

soal nomor 11. Apabila dianalisis menggunakan Q matrik, maka *item* soal nomor 23 berada pada Q matriks 5. Atribut yang menyusun Q matriks 5 adalah atribut A3, A8 dan A12. Atribut A3 yaitu satuan dan besaran-besaran kalor, Atribut A8 yaitu tentang kalor laten dan A12 dan atribut A12 yaitu tentang Azas Black. Matriks Q 5 berukuran 5x4. Besarnya kelas laten pada matriks Q5 adalah $2^4 = 16$.

$$Q_5 = \begin{array}{c|cccc} & A_3 & A_5 & A_8 & A_{12} \\ \hline I_{17} & 1 & 0 & 0 & 0 \\ I_{20} & 1 & 1 & 1 & 0 \\ I_{21} & 1 & 1 & 1 & 0 \\ I_{22} & 1 & 0 & 0 & 1 \\ I_{23} & 1 & 0 & 1 & 1 \end{array}$$

Untuk dapat menjawab soal nomor 23 maka siswa harus menguasai materi prasyarat seperti yang ditampilkan pada Gambar 2.



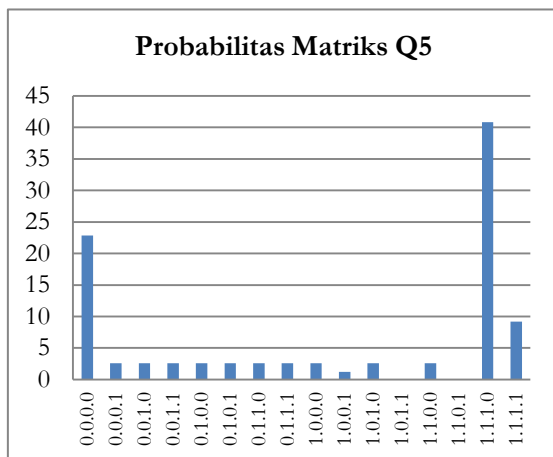
Gambar 2. Gambar 2. Hierarki Materi prasyarat *item* soal nomor 23

Keterangan:

- K8 : Menjelaskan proses perubahan suhu pada zat
- K10: Menjelaskan konsep perpindahan kalor
- K14: Menentukan konversi satuan pada besaran kalor
- K43: Menjelaskan hukum kekekalan energi pada persamaan kalor (Azas Black)
- K45: Menentukan suhu campuran pada zat yang berbeda.

Kesulitan yang dialami siswa terutama dalam menentukan konversi dan menghitung besaran-besaran kalor dan kalor la-

ten. Selain itu, siswa juga belum paham terhadap konsep Azas Black. Hasil diagnostik ini digunakan oleh guru untuk menentukan tindakan sesuai dengan yang dibutuhkan siswa. *Item* pada matriks Q5 adalah *item* dengan jenis pengetahuan prosedural yang mempunyai tingkat kesukaran tinggi. Untuk dapat menyelesaikan soal dalam bentuk prosedural, siswa harus memahami bagaimana menyelesaikan soal dengan langkah-langkah dan prosedur yang tepat, mengkonversi satuan, menggunakan rumus sesuai dengan konsep yang benar. Oleh karena itu, diperlukan pengetahuan prasyarat untuk dapat menyelesaikan soal ini, siswa tidak cukup hanya menghafal rumusnya, namun harus paham terhadap prosedur dan teknik penyelesaiannya.



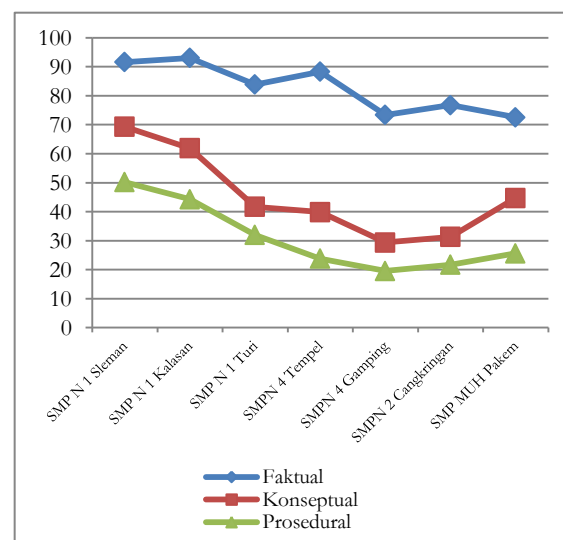
Gambar 3. Probabilitas Q Matriks 5

Guru harus membiasakan dan melatih siswa mengerjakan soal-soal dengan jenis pengetahuan prosedural. Untuk mendapatkan pengetahuan prosedural yang terkoneksi secara baik dengan pengetahuan konseptualnya tentu saja kepada siswa perlu ditanamkan konsep-konsep secara baik, dikaitkannya konsep tersebut dengan konsep yang lain, dan dilatihkan bagaimana merepresentasikan konsep dengan simbol, serta dilatihkan bagaimana menggunakan aturan atau prosedur untuk menyelesaikan permasalahan fisika.

Soal yang mempunyai *item difficulty* terendah adalah *item* soal nomor 11. Soal ini merupakan soal ingatan dan jenis soal faktual. Soal faktual cenderung mudah karena

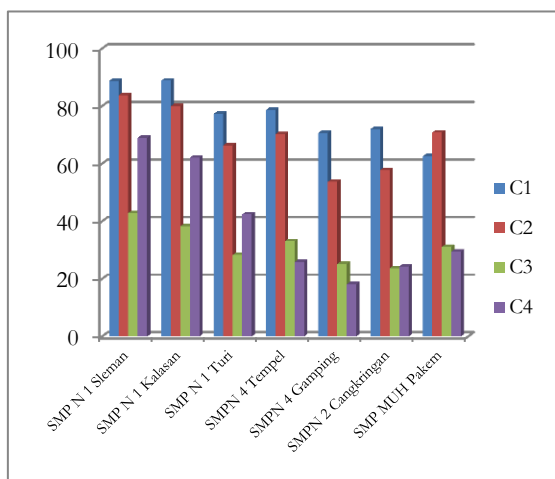
dalam kehidupan sehari-hari siswa sering menjumpainya bahkan mempraktikannya. Pada soal nomor 11 sebesar 97,7% siswa dapat menyelesaikan soal dan dapat menyampaikan alasannya dengan benar. Pada pengetahuan faktual cenderung lebih dikuasai siswa dari pada pengetahuan jenis konseptual dan prosedural. Pengetahuan faktual berkaitan dengan pengalaman yang dialami siswa dalam kehidupan sehari-hari. Apa yang dialami oleh siswa akan lebih mudah tersimpan dalam otak siswa. Oleh karena itu, dalam pembelajaran IPA, guru seharusnya senantiasa mengkaitkan materi pelajaran dengan peristiwa sehari-hari.

Pada penelitian ini soal dengan jenis pengetahuan faktual lebih banyak dikuasai oleh siswa dibandingkan dengan jenis pengetahuan prosedural. Siswa lebih menguasai soal yang berbentuk ingatan dan hafalan tanpa memahami konsepnya. Siswa cenderung hafal dengan rumus-rumus fisika dan tahapan-tahapan penyelesaiannya tanpa memahami konsep dan prosedur yang benar. Dengan demikian, apabila diberikan soal yang berbeda, siswa akan mengalami kesulitan. Hal ini sejalan dengan penelitian Oksuz (2007, p.13) menyatakan bahwa siswa sulit menghubungkan pengetahuan konseptual yang dimiliki ke dalam pengetahuan prosedural sehingga siswa cenderung menghafalkan tahap pengerjaan soal, tidak berdasarkan pemahaman terhadap konsep terkait.



Gambar 4. Hasil Tes Diagnostik Dilihat dari Jenis Pengetahuan

Hasil tes diagnostik kognitif yang telah diujicobakan di beberapa sekolah menunjukkan bahwa siswa pada setiap sekolah masih mengalami kesulitan pada tiap dimensi kognitif. Pada pelajaran IPA Fisika siswa disekolah cenderung hafal rumus namun tidak paham konsepnya. Sehingga pada soal-soal dengan jenis prosedural yang disusun dengan tipe yang berbeda maka siswa akan mengalami kesulitan. Dengan diketemukannya kesulitan siswa pada setiap sekolah berkaitan dengan pengetahuan konseptual dan prosedural, maka guru menjadi salah satu faktor yang mempunyai peranan, artinya kesalahan dan kesulitan yang dialami siswa kemungkinan dipengaruhi oleh cara guru dalam mengajar dan cara guru dalam memberikan evaluasi.



Gambar 5. Grafik Hasil Tes Diagnostik Berdasar Proses Pengetahuan

Grafik menunjukkan bahwa pada tiap sekolah masih mengalami kesulitan pada aspek aplikasi dan analisis. Hal ini seperti penelitian yang dilakukan Fauzan (2010, p.227) bahwa hampir semua siswa mengalami kesulitan dalam mengerjakan butir-butir soal yang termasuk analisis. Pada mata pelajaran IPA Fisika seharusnya soal aplikasi dan analisis sudah dibiasakan diberikan oleh guru sehingga siswa tidak mengalami kesulitan, namun pada penelitian ini aspek analisis dan aplikasi siswa masih sangat lemah. Dengan demikian, guru pada setiap sekolah perlu memperkaya dan melatih siswa dengan menggunakan soal-soal yang mampu

menjadikan siswa berpikir kritis dengan memperhatikan *learning continuum* agar dapat meningkatkan kualitas pembelajaran.

Pengembangan tes diagnostik kognitif akan lebih bermakna apabila dapat digunakan oleh guru. Pada penelitian ini guru dapat memanfaatkan hasil tes diagnostik kognitif sebagai dasar untuk memberikan tindak lanjut berupa perlakuan yang tepat dan sesuai dengan kelemahan yang dimiliki siswa melalui hasil pengembangan *software* komputer yang dilengkapi dengan *password*. Hasil tes diagnostik kognitif yang telah dikembangkan dengan *software* komputer terdiri atas dua informasi yaitu (1) banyaknya *item* yang dijawab dengan benar dan *item* yang dapat dijawab siswa berdasarkan jenis pengetahuannya, (2) informasi tentang materi yang sudah dikuasai dan belum dikuasai siswa berdasarkan *learning continuum* dan materi prasyarat yang belum diketahui siswa.

Penelitian ini dapat menjawab pertanyaan mengapa peserta tes tidak menguasai suatu materi sekaligus dapat menginterpretasikan hasil tes diagnostik kognitif pada mata pelajaran IPA SMP sehingga mampu mendiagnosis kesulitan belajar siswa. Kesulitan siswa dalam pembelajaran IPA Fisika materi kalor akan dapat dipecahkan, karena melalui *latent class* diketahui materi apa saja yang tidak dikuasai siswa. Guru dapat memberikan penanganan terutama letak kelemahan dan kesulitan belajar yang dialami siswa dan kemungkinan prestasi yang bisa dikembangkan. Hal ini sesuai dengan yang disampaikan Maddalena (2008) bahwa pengajar tidak hanya mampu mengidentifikasi dimana dan bagaimana siswa sering salah paham terhadap konsep yang sangat rumit, tetapi juga bahwa mereka dapat mengidentifikasi cara terbaik untuk memperbaikinya.

Tes diagnostik kognitif yang dikembangkan dalam penelitian ini dipadu *software* komputer yang dapat membantu guru untuk memberikan pembelajaran yang sesuai dengan proses berpikir siswa sehingga kualitas pembelajaran IPA meningkat. Penelitian ini sejalan dengan hasil penelitian Afaq & Al-Mashari (2010, p.85) yang telah mengembangkan alat diagnostik bertujuan untuk

membantu siswa dengan mendiagnosis kesulitan belajar. Alat yang dikembangkan memiliki kemampuan menguji kemampuan pelajar untuk menjawab pertanyaan dengan benar dengan mengevaluasi respon otomatis dan kemudian menghasilkan daftar kelemahan atau kesalahan yang ditemukan. Umpan balik yang diberikan dapat digunakan oleh instruktur/guru untuk meningkatkan keterampilan pelajar dengan berfokus pada kelemahan yang dilaporkan, dan menghasilkan tes diagnostik yang terstandar. Tes diagnostik kognitif yang tersusun dilengkapi dengan *software* komputer, sehingga tes diagnostik ini dapat langsung digunakan oleh guru untuk menghasilkan informasi yang dapat mendiagnosis kelemahan dan kekuatan siswa sehingga guru dapat memberikan tindak lanjut dan penanganan yang sesuai agar dapat meningkatkan kualitas pembelajaran.

Simpulan dan Saran

Tahapan pengembangan instrumen diagnostik kognitif pada mata pelajaran IPA SMP materi kalor dimulai dengan perumusan *Learning Continuum*. Tahapan berikutnya adalah penyusunan Kisi-kisi tes, Hierarki Materi Prasyarat, Spesifikasi Tes, Soal dan Q matriks yang diturunkan dari Atribut. Karakteristik instrumen diagnostik kognitif yang tersusun telah memenuhi validitas isi melalui *expert judgement* dengan FGD.

Hasil pengembangan pada revisi akhir terdiri atas 28 soal yang berbentuk pilihan ganda dengan alasan. Instrumen diagnostik kognitif pada mata pelajaran IPA materi kalor telah memenuhi persyaratan yaitu *fit* dengan model, dengan menggunakan program *Quest*. Instrumen penelitian ini memiliki *mean infit MNSQ* 1,01 dengan simpangan baku 0,09. Reliabilitas tes dengan rumus *alpha Cronbach* menunjukkan tes memiliki reliabilitas tinggi yaitu 0,72. Uji fit kesesuaian model atribut dengan Q matriks menggunakan program M Plus berhasil menyusun 7 buah Q matriks dengan *Latent class* pada setiap matriks adalah 2^n dengan n adalah jumlah atribut. Penelitian ini berhasil menemukan besarnya propabilitas pada se-

tiap *latent class* pada tujuh Q matriks yang tersusun.

Saran yang diberikan penelitian ini untuk Guru IPA SMP yaitu agar dapat menerapkan produk Instrumen Tes Diagnostik Kognitif pada materi Kalor, sehingga guru dapat mendiagnosis kesulitan belajar siswa dan dapat memberikan bantuan sesuai dengan yang dibutuhkan siswa. Tindak lanjut yang dilakukan guru dalam pembelajaran hendaknya lebih menanamkan konsep fisika dan prosedur yang tepat dan memperhatikan materi prasyarat yang harus dikuasai siswa. Dengan demikian, kualitas pembelajaran dapat meningkat. Bagi peneliti yang tertarik untuk mengembangkan model tes diagnostik kognitif lebih lanjut, perlu ketelitian dan kecermatan dalam menyusun atribut yang harus selaras dengan *learning continuum* dan hierarki materi prasyarat.

Pemerintah khususnya Badan Standar Nasional Pendidikan (BSNP), diharapkan memperhatikan hierarki materi prasyarat dan *learning continuum* dalam menyusun standar kompetensi dan kompetensi dasar agar sesuai dengan struktur berpikir siswa. Dengan demikian, anak dapat belajar sesuai dengan tahap berpikirnya. Instrumen tes diagnostik kognitif akan lebih mudah dimanfaatkan jika dikembangkan dengan model aplikasi yang lengkap dan spesifik. Oleh karena itu, bagi yang berminat mengembangkan dapat melakukan penelitian lanjutan secara lebih mendalam dan komprehensif baik dalam bidang kajian IPA ataupun kajian pada bidang yang lain dapat mengembangkannya dengan model yang lebih lengkap dan spesifik.

Daftar Pustaka

- Adam, R.J. & Khoo, S.T. (1996). *Quest. The Interactive Test Analysis System*. Version 2.1. Melbourne: ACER.
- Afaq, A & Al-Mashari, A. (2010). On the development of a computer based diagnostic assessment tool to help in teaching and learning process. *International Journal of Education and Development using Information and Communi-*

- cation Technology(IJEDICT). Vol 6. Issue 1.76-87.*
- Anderson, L. W., & Krathwohl, D. R. (Eds). (2001). A taxonomy for learning, teaching and assessing: *A revision of Bloom's Taxonomy of educational objectives*: Complete edition, New York: Longman.
- De La Torre, J. & Karelitz, T.M (2009). Impact of Diagnosticity on the Adequacy of Models for Cognitive Diagnosis under a Linear Attribute Structure: A Simulation Study. *Journal of Educational Measurement Winter*. 46 (4):450-469.
- Dunn, L., Morgan.C., Parry, S. R., & Reilly, M.O. (2004). *The Student assessment Handbook*. New York. Taylor & Francise. Library
- Duskri, M., Kumaidi, K., & Suryanto, S. (2014). Pengembangan Tes Diagnostik Kesulitan Belajar Mtematika di SD. *Jurnal Penelitian Dan Evaluasi Pendidikan*, 18(1), 44-56. Retrieved from <http://journal.uny.ac.id/index.php/jp ep/article/view/2123/1768>.
- Ebel, Robert, L. (1979). *Essentials of Educational Measurement*. New Jersey: Prentice. Hall.Inc.
- Efendi, R. (2010). Kemampuan Fisika Siswa Indonesia Dalam TIMSS (Trend of International on Mathematic and Science Study). *Prosiding Seminar Nasional Fisika 2010*. ISBN: 978-979-98010-6-7.
- Ellis, T.J & Levy, Y. (2010). *A Guide for Novice Researchers: Design and Development Research Methods*. Proceedings of Informing Science & IT Education Conference (InSITE), 108-118.
- Fauzan. (2010). Pengembangan tes diagnostik kesulitan belajar matematika di SMA. *Disertasi doktor*, tidak diterbitkan, Universitas Negeri Yogyakarta, Yogyakarta.
- Gage, N.L & Berliner, D.C. (1991). *Educational Psychology*. 5th ed. Boston: Houghton Mifflin Company
- Gierl, M.J. (2007). Making Diagnostic Inferences About Cognitive Attributes Using the Rule Space Model and Attribute Hierarchy Method. *Journal of Educational Measurement*, 44(4), 325-340.
- Gorin, J.S. (2007). Test construction and diagnostic testing. Dalam Leighton, J.P. & Gierl, M.J.(eds). *Cognitive diagnostic assessment for education theory and applications*.(p.174). Cambridge. Cambridge University Press.
- Gronlund, Norman, E. (1982). *Measurement and Evaluation in Teaching*. New York: Macmilan Publishing.co.
- Isgiyanto, A. (2011). Diagnosis Kesalahan Siswa Berbasis Penskoran Politomus Model Partial Credit Pada Matematika. *Jurnal Penelitian dan Evaluasi Pendidikan*, 15 (2), 308-325.
- Kusaeri. (2012). Pengembangan tes diagnostik dengan menggunakan model DINA unuk mendapatkan informasi salah konsepsi dalam aljabar. *Disertasi doktor*, tidak diterbitkan, Universitas Negeri Yogyakarta, Yogyakarta.
- Leung, C., & Mohan, B. (2004). Teacher formative assessment and talk in classroom contexts: *assessment as discourse and assessment of discourse*. <http://itj.sagepub.com>.
- Maddalena, T. (2008). Issues of power and equity in two models of self-assessment. *Teaching in Higher Education*. 13 (1) 81-92.
- Mardapi, D. (2008). *Teknik penyusunan instrumen tes dan non tes*. Yogyakarta: Mitra Cendikia Offset.
- Mehrens, W.A. & Lehman, I.J. (1984). *Measurement and evaluation in education and Psychology*. New York. Holt, Rinehart and Winson, Inc.

- Mislevy, R.J. (1994). *Evidence and Inference in educational assessment*. *Psychometrica*, 59, 439-483.
- Nichols, P.D. (1994). A framework for developing cognitively diagnostic assessment. *Review of Educational*. 64 (4), 575-603.
- Nitko, J.A. (1996). *Of Educational Assessment of Student*. New Jersey: Prentice Hall. Inc.
- NWEA. (2003). *Idaho State Aligned Learning Continuum Release 1.0* diambil dari <https://id.scribd.com/doc/244834242/ID-State-Learning-Continuum>
- Nylund, K.L., Asparouhov, T. & Muthen, B.O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling A Monte Carlo simulation study. *Structural Equation Modeling*, 14, 535-569.
- Oksuz, C. (2007). Childrens understanding of equality and the equal symbol. Adnan Menderes. Diambil dari cumalioksuz@adu.edu.tr
- Suparno, P. (1998). Miskonsepsi (Konsep alternatif) Siswa SMU dalam Bidang Fisika. Dalam Sumaji. *Pendidikan sains yang Humanis*. 95-111. Kanisius: Yogyakarta.
- Permendiknas Nomor 20 Tahun 2007 tentang Standar penilaian.
- Rupp, A. A., & Templin, J. L. (2008). Unique characteristics of diagnostic classification models: A comprehensive review of the current state-of-the-art. *Measurement: Interdisciplinary Research and Perspectives*, 6(4), 219-262.
- Rupp, A.A., Templin, J. & Henson, R.A. (2010). *Diagnostic measurement: Theory, methods and applications*. New York: The Guilford Press.
- Russell, M. & O'dwyer. L.M. (2009). Diagnosing students' misconceptions in algebra: results from an experimental pilot. *Behavior Research Methods*. 41(2), 414-424.
- Taufik, M., Sukmadinata, Abdulkhak, I., Tumbelaka, B. Y. (2010). Desain model pembelajaran untuk meningkatkan kemampuan pemecahan masalah dalam pembelajaran IPA (fisika) sekolah menengah pertama di kota bandung. *Berkala Fisika, Vol 13, No 2 (2010): Edisi Kbusus, E31-E44*
- Templin, J. (2009). Estimation of diagnostic models with Mplus. Slide presentasi disampaikan dalam *National Council on Measurement in education, Diagnostic Modeling*, tanggal 12 Februari 2014. <http://jtemplin.coe.uga.edu/teaching>.
- Weeden, P. Winter, J. & Broadfoot, P. (2002). *Assessment what's in for school?*. Routledge falmer: London
- Wright, B.D & Masters, GN. (1982). *Rating Scale analysis*. Chicago: Mesa Press.