

DIAGNOSIS KESALAHAN SISWA BERBASIS PENSKORAN POLITOMUS MODEL *PARTIAL CREDIT* PADA MATEMATIKA

Awal Isgiyanto

Universitas Bengkulu

Jl. W.R. Supratman Kandang Limun Bengkulu 38371 A

awalunib@yahoo.co.id.

Abstrak

Penelitian ini bertujuan untuk menemukan informasi diagnostik dari kesalahan jawaban peserta pada Ujian Nasional (UN) Matematika. Informasi diagnostik yang ditemukan meliputi atribut yang mendasari butir soal, ketidaktuntasan atribut, dan jenis kesalahan yang dilakukan oleh peserta. Penelitian ini merupakan diagnosis *post-hoc*, yang digambarkan sebagai pendekatan *retrofitting*. Analisis butir soal dan respons butir pada UN mata pelajaran matematika untuk menemukan informasi diagnostik pada kategori isi, proses, dan keterampilan siswa SMP di Bantul Yogyakarta 2007/2008. Hasil penelitian (1) atribut yang mendasari butir soal pada matematika ada 47 atribut, meliputi 4 atribut isi, 36 atribut proses, dan 7 atribut keterampilan, (2) ketidaktuntasan atribut isi, proses, dan keterampilan yang tertinggi pada geometri dan pengukuran, (3) jenis kesalahan tertinggi pada bilangan, aljabar, dan geometri dan pengukuran adalah kesalahan konsep, dan jenis kesalahan tertinggi pada statistika dan peluang adalah kesalahan interpretasi bahasa, dan (4) penemuan informasi diagnostik data UN Matematika dapat dilakukan melalui mekanisme identifikasi atribut, pengembangan rubrik penskoran politomus, perhitungan ketidaktuntasan atribut, dan diagnosis kesalahan peserta tes.

Kata kunci: *diagnostik, matematika, partial credit model*

A DIAGNOSIS OF STUDENTS' MISTAKES BASED ON POLYTOMOUS SCORING PARTIAL CREDIT MODELS IN MATHEMATICS

Awal Isgiyanto

Bengkulu State University

Jl. W.R. Supratman Kandang Limun Bengkulu 38371 A

awalunib@yahoo.co.id.

Abstract

This study aims to reveal diagnostic information from the participants' incorrect answers in the National Examination (NE) of Mathematics. The diagnostic information includes the attributes underlying test items, the attribute exhaustiveness, and the types of mistakes that the participants made. This study was a post-hoc diagnostic study, described as the retrofitting approach. The analysis of test items and item responses in the NE of Mathematics aimed to reveal the diagnostic information in the Junior High School students' content, process, and skill categories. The results are that (1) there are 47 attributes underlying the mathematics test items, consisting of 4 content attributes, 36 process attributes, and 7 skill attributes, (2) the highest inexhaustiveness of the content, process, and skill attributes is in the topics of geometry and measurement, (3) most mistakes in the topics of numbers, algebra, geometry, and measurement are those of concepts, and most mistakes in statistics and probability are those of language interpretation, and (4) the findings of diagnostic information in the data on the NE of Mathematics can be revealed through the mechanism of identifying the attributes, developing a polytomous scoring rubric, finding out the attribute inexhaustiveness, and diagnosing mistakes the test participants.

Keywords: *diagnosis, mathematics, partial credit model*

Pendahuluan

Pengukuran merupakan prosedur penetapan skor siswa, yang dapat menunjukkan derajat yang dimiliki oleh siswa terhadap karakteristik yang sedang diukur. Penilaian merupakan proses menentukan hasil belajar berdasarkan suatu bukti hasil pengukuran (Mardapi, 2005: 75). Penilaian hasil belajar yang dilakukan oleh pemerintah diwujudkan dalam bentuk Ujian Nasional (UN).

Penelitian pendahuluan di Dinas Pendidikan, Pemuda, dan Olahraga Provinsi Daerah Istimewa Yogyakarta menunjukkan bahwa skor matematika peserta ujian (siswa) pada Ujian Nasional SMP Tahun Pelajaran 2007/2008 yang kurang dari 5 ada 34,33%. Siswa yang memperoleh skor matematika kurang dari 5 di Kota Yogyakarta ada 39,66%, Kulon Progo 37,31%, Bantul 34,88%, Sleman 24,75%, dan Gunung Kidul 38,61%. Hasil tersebut menunjukkan bahwa para siswa masih banyak melakukan kesalahan dalam menjawab butir soal Ujian Nasional Matematika.

Pada Peraturan Menteri Pendidikan Nasional Republik Indonesia Nomor 22 Tahun 2006 tentang Standar Isi untuk Satuan Pendidikan Dasar dan Menengah disebutkan bahwa ruang lingkup mata pelajaran Matematika pada satuan pendidikan SMP/MTs meliputi aspek-aspek (1) bilangan, (2) aljabar, (3) geometri dan pengukuran, dan (4) statistika dan peluang. Pada tingkat SMP, matematika berfungsi untuk mengembangkan kemampuan menghitung, mengukur, merumuskan dan menggunakan rumus matematika yang diperlukan dalam kehidupan sehari-hari melalui materi bilangan, aljabar, geometri dan pengukuran, dan statistika dan peluang.

Respons butir dikotomus mempunyai dua kategori skor jawaban, yaitu jawaban betul (skor 1) dan jawaban salah (skor 0) (Bond & Fox, 2007: 49; DeMars, 2010: 9). Model penskoran dikotomus tidak dapat digunakan untuk menemukan kesalahan yang dilakukan oleh siswa, karena semua *option* yang salah diberi skor 0. Kenyataannya, kesalahan yang dilakukan siswa dapat berbeda-beda. Untuk dapat menemukan kesalahan yang dilakukan siswa, diperlukan penskoran politomus.

Penskoran politomus adalah model respons butir yang mempunyai kemungkinan jawaban lebih dari dua kategori (Wells, Hambleton, & Urip

Purwono, 2008; 1). Model penskoran politomus meliputi *Graded Response Model (GRM)* dari Samejima, *Nominal Model (NM)* dari Bock, *Partial Credit Model (PCM)* dari Masters, *Generalized Partial Credit Model (GPCM)* dari Muraki, dan *Rating Scale Model (RSM)* dari Andrich (DeMars, 2010: 22; Thissen, Nelson, Rosa, et al., 2001: 143-149; Wells, Hambleton, & Purwono, 2008: 2-10; Hessen, 2009: 5-17). Penerapan penskoran politomus diharapkan dapat memberikan solusi keterbatasan penskoran dikotomus, terkait dengan diagnosis kesalahan siswa.

Kesulitan belajar didefinisikan sebagai tidak dimilikinya kompetensi tertentu oleh siswa terkait dengan pemahaman isi yang spesifik, proses kognitif, atau keterampilan dalam pemecahan masalah. Siswa yang mengalami kesulitan belajar dapat melakukan kesalahan dalam menjawab butir soal. Kesalahan yang dilakukan oleh siswa didefinisikan sebagai respons terhadap butir soal yang tidak sesuai dengan respons yang diharapkan (Gierl, 2007a). Diagnosis kesalahan yang dilakukan siswa didasarkan pada atribut. Atribut didefinisikan sebagai prosedur, proses, keterampilan, atau kompetensi yang harus dimiliki siswa untuk menyelesaikan butir soal (Gierl, 2007a; Gierl, 2007b; Gierl, Zheng, & Cui, 2008; Robert & Gierl, 2010). Atribut dalam kategori isi (*content*) adalah materi yang digunakan di dalam kerangka tes. Atribut kategori proses (*process*) merupakan keterampilan umum. Kategori proses ini merupakan keterampilan siswa yang diharapkan setelah mengikuti pelajaran dengan materi yang diberikan pada kategori isi. Atribut pada kategori keterampilan (*skill*) merupakan keterampilan yang spesifik.

Metode diagnostik digunakan untuk mengetahui kekuatan-kekuatan dan kelemahan-kelemahan (*strengths and weaknesses*) siswa (Samejima, 1995; Leighton, Gierl, & Hunka, 2004; Leighton & Gierl, 2007). Informasi diagnostik terkait dengan tidak dimilikinya atau tidak dikuasainya kompetensi tertentu yang diperlukan untuk menyelesaikan butir soal. Informasi tersebut berupa atribut yang mendasari butir soal, ketidaktuntasan atribut, dan kesalahan yang dilakukan oleh siswa. Demikian pentingnya rekomendasi kepada pemangku kepentingan untuk peningkatan kemampuan guru, pengembangan butir soal, dan untuk perbaikan proses pembelajaran matematika, maka perlu penelitian untuk mendiagnosis kesalahan siswa.

Berdasarkan uraian di atas, masalah penelitian yang akan diteliti cara mendiagnosis kesalahan siswa berbasis penskoran model *Partial Credit*. Tujuan penelitian ini untuk menemukan informasi diagnostik dari kesalahan jawaban siswa pada UN Matematika. Informasi diagnostik yang ingin ditemukan meliputi atribut yang mendasari butir soal, ketidaktuntasan atribut, dan kesalahan yang dilakukan oleh siswa.

Metode Penelitian

Penelitian ini merupakan diagnosis *post-hoc* yang digambarkan sebagai pendekatan *retrofitting* (Gierl, 2007a). Pendekatan *retrofitting* dilakukan melalui analisis butir soal dan data respons butir pada UN Matematika. Subjek penelitian adalah peserta UN Matematika SMP Tahun Pelajaran 2007/2008 di wilayah Kabupaten Bantul, Daerah Istimewa Yogyakarta. Objek yang dipilih adalah 40 butir soal dan 1016 respons butir peserta UN Matematika. Sampel diambil secara *simple random sampling* diperoleh 12 sekolah atau 1016 peserta. Data *ex post facto* berupa respons butir peserta ujian dan butir soal pada UN Matematika SMP Tahun Pelajaran 2007/2008 dikumpulkan dengan teknik dokumentasi. Data diambil dari Dinas Pendidikan dan Olahraga Daerah Istimewa Yogyakarta sebagai basis analisis untuk menemukan informasi diagnostik.

Identifikasi atribut dilakukan melalui pendekatan *retrofitting* dengan cara menganalisis data berupa butir soal UN Matematika. Pendekatan *retrofitting* didasarkan pada atribut yang muncul di antara butir soal. Atribut yang mendasari butir soal sejalan dengan kompetensi dasar (KD). Atribut yang mendasari butir soal dikategorikan menjadi atribut kategori isi (I), atribut kategori proses (P), dan atribut kategori keterampilan (K). Hasil identifikasi atribut divalidasi oleh siswa, guru matematika SMP, dan pakar pendidikan matematika.

Data berupa butir soal UN Matematika dianalisis untuk mendapatkan Rubrik Penskoran. Rubrik Penskoran disusun melalui pendekatan *retrofitting* didasarkan pada atribut yang muncul di antara butir soal. Rubrik Penskoran divalidasi oleh siswa, guru matematika SMP, dan pakar pendidikan matematika. Validitas Rubrik Penskoran didasarkan pada hasil

validasi tersebut. Kesahihan skor yang diperoleh juga didasarkan pada validitas Rubrik Penskoran tersebut.

Suatu atribut dikatakan tidak tuntas, jika tidak dikuasai dan tidak diterapkan dengan tepat. Untuk menemukan ketidaktuntasan atribut didasarkan pada Rubrik Penskoran. Ketidaktuntasan atribut diperoleh pada data yang diskor politomus. Ketidaktuntasan atribut yang mendasari butir soal matematika diperoleh dengan menganalisis respons butir yang salah menjadi proporsi peserta yang melakukan kesalahan pada suatu atribut.

Data berupa butir soal dan respons butir UN Matematika dianalisis untuk menemukan jenis kesalahan yang dilakukan peserta. Analisis untuk menemukan kesalahan yang mengakibatkan peserta salah menjawab butir soal, diperoleh berdasarkan pada Rubrik Penskoran. Kesalahan yang telah ditemukan, untuk selanjutnya dilakukan klasifikasi menurut karakteristik jenis kesalahan.

Hasil Penelitian dan Pembahasan

1. Rubrik dan Model Penskoran

Untuk menjawab butir soal matematika diperlukan sejumlah tahapan penyelesaian. Pada setiap tahapan diperlukan penguasaan atribut yang mendasari butir soal yang bersangkutan. Penyusunan Rubrik Penskoran dilakukan melalui pendekatan *retrofitting*. Pendekatan *retrofitting* didasarkan pada atribut yang muncul di antara butir soal. Konstruksi suatu model kognitif setiap distraktor diidentifikasi untuk menemukan atribut yang mendasarinya. Identifikasi dilakukan melalui analisis tugas pada butir soal yang mewakili suatu *domain* tertentu.

Model penskoran politomus meliputi *Graded Response Model* (GRM) dari Samejima, *Nominal Model* (NM) dari Bock, *Partial Credit Model* (PCM) dari Masters, *Generalized Partial Credit Model* (GPCM) dari Muraki, dan *Rating Scale Model* (RSM) dari Andrich (DeMars, 2010: 22; Thissen, Nelson, Rosa, et al., 2001: 143-149; Wells, Hambleton, & Purwono, 2008: 2-10; Hessen, 2009: 5-17). Oleh karena, itu perlu inovasi model penskoran dengan mengembangkan model penskoran politomus pada butir soal UN Matematika.

Partial Credit Model (PCM) dikembangkan dari *Rasch Model (RM)* oleh Masters. *RM* untuk data skor dikotomus dikembangkan menjadi skor lebih dari dua kategori (politomus). *PCM* tidak mensyaratkan langkah penyelesaian butir tes harus berurutan dan tidak harus mempunyai kesulitan yang sama (De Ayala, 1993). *Threshold* pada penskoran *PCM* dari suatu kategori ke kategori berikutnya tidak selalu lebih besar. *PCM* juga cocok untuk butir yang diskor dalam kategori berjenjang, akan tetapi indeks kesukaran dalam setiap langkah tidak perlu terurut. Skor kategori yang lebih tinggi menunjukkan kemampuan yang lebih tinggi dari pada skor kategori yang lebih rendah. Respons butir pada penelitian ini dapat betul pada langkah tertentu, tetapi dapat salah pada langkah yang lain. Pada kasus respons seperti itu, disarankan untuk menggunakan penskoran *PCM*. Pada *RM* atau *PCM* diasumsikan bahwa parameter tingkat kesulitan butir merupakan satu-satunya karakteristik butir yang mempengaruhi kinerja peserta didik.

Untuk memperoleh *option* yang betul dapat diperoleh melalui penyelesaian yang bermacam-macam. Oleh karena itu, peserta memilih *option* yang betul diberi skor maksimum (skor 3). Pada tiga *option* yang salah dikategorikan menurut bobot kesalahannya menjadi kategori ringan (skor 2), kategori sedang (skor 1), dan kategori berat (skor 0).

Penskoran politomus *PCM* yang dikembangkan bertujuan untuk mengatasi kelemahan model penskoran dikotomus, yaitu tidak dapat digunakan untuk mengontrol distraktor dan untuk mengetahui kesalahan yang dilakukan oleh peserta, karena semua *option* yang salah diberi skor 0. Kenyataannya, kesalahan pada distraktor dapat berasal dari sumber kesalahan yang berbeda. Pengembangan model penskoran politomus ini dengan menyesuaikan *PCM* dari Masters dengan pertimbangan (1) model penskoran *PCM* dapat memberikan *credit* pada penerapan atribut dengan benar, (2) model penskoran *PCM* akan memberikan peluang meningkatkan akurasi pengukuran dibandingkan dengan model dikotomus, (3) antartahapan dalam menyelesaikan butir soal matematika tidak selalu mempunyai tingkat kesulitan yang sama, dan (4) kategori yang lebih tinggi tidak selalu mempunyai *threshold* yang

lebih tinggi, demikian juga kategori yang lebih rendah tidak selalu mempunyai *threshold* yang lebih rendah.

2. Atribut yang Mendasari Butir Soal Ujian Nasional Matematika

Atribut yang menyusun konstruk suatu model kognitif diidentifikasi melalui analisis tugas pada butir soal yang mewakili suatu *domain* tertentu. Untuk mengidentifikasi atribut yang mendasari butir soal UN Matematika digunakan pendekatan diagnosis *post-hoc*, yang digambarkan sebagai pendekatan *retrofitting* (Gierl, 2007b). Pendekatan ini dilakukan dengan cara menganalisis butir soal UN Matematika. Pendekatan *retrofitting* tersebut didasarkan pada atribut yang muncul di antara butir soal UN Matematika.

Sebaran hasil identifikasi atribut yang mendasari butir soal disajikan pada Tabel 1. Pada Tabel 1 tampak bahwa ada 47 atribut yang mendasari butir-butir soal UN Matematika, meliputi 4 atribut kategori isi, 36 atribut kategori proses, dan 7 atribut kategori keterampilan.

Tabel 1. Distribusi Atribut yang Mendasari Butir Soal

	Bilangan	Aljabar	Geometri	Statistika	Matematika
Atribut Isi	1	2	3	3	4
Atribut Proses	6	17	20	6	36
Atr. Keterampilan	0	4	4	1	7
Jumlah	7	23	27	10	47

Tiga puluh enam atribut kategori proses tersebut ada tiga atribut yang tidak termasuk KD. Atribut yang dimaksud adalah atribut (At.P.1) kompetensi melakukan proses pemikiran logis, atribut (At.P.2) kompetensi menentukan nilai bentuk aljabar, dan atribut (At.P.3) kompetensi melakukan konversi satuan panjang. Atribut (At.P.1) tersebut teridentifikasi pada pemecahan masalah butir soal 2, 3, 5, 6, 7, 8, 14, 15, 32, 33, 35, dan 39. Atribut (At.P.2) teridentifikasi pada butir soal 19. Atribut (At.P.3) teridentifikasi pada butir soal 32 dan 33.

Ketiga atribut tersebut tidak termasuk KD, akan tetapi merupakan kompetensi sangat penting yang diperlukan dalam penyelesaian butir soal UN Matematika Tahun Pelajaran 2007/2008. Untuk selanjutnya, ketiga atribut tersebut dapat dipertimbangkan dalam revisi KD mata pelajaran matematika SMP.

3. Ketidaktuntasan Atribut

Suatu atribut dikatakan tidak tuntas jika tidak dikuasai dan tidak diterapkan dengan tepat. Ketidaktuntasan atribut yang mendasari butir soal diperoleh dengan menganalisis respons butir yang salah menjadi proporsi peserta yang melakukan kesalahan pada suatu atribut.

a. Ketidaktuntasan Atribut Isi

Ketidaktuntasan atribut isi pada butir soal matematika mempunyai rerata 25,24. Ketidaktuntasan atribut isi berturut-turut mulai yang tertinggi adalah pada submateri (1) geometri dan pengukuran, (2) aljabar, (3) statistika dan peluang, dan (4) bilangan. Ketidaktuntasan atribut isi berturut-turut mulai yang tertinggi adalah (1) atribut (At.I.3), yaitu konsep dasar geometri dan pengukuran, (2) atribut (At.I.2), yaitu konsep dasar aljabar, (3) atribut (At.I.4), yaitu konsep dasar statistika dan peluang, dan (4) atribut (At.I.1), yaitu konsep dasar geometri dan pengukuran.

b. Ketidaktuntasan Atribut Proses

Ketidaktuntasan atribut proses yang tertinggi terjadi pada atribut (At.P.28), yaitu kompetensi menghitung luas permukaan dan volume kubus, balok, prisma dan limas, dan terendah pada atribut (At.P.8), yaitu kompetensi menyelesaikan persamaan linear satu variabel. Ketidaktuntasan atribut proses tertinggi terjadi pada atribut (At.P.27), yaitu kompetensi mengidentifikasi sifat-sifat kubus, balok, prisma dan limas serta bagian-bagiannya, dan terendah terjadi pada atribut (At.P.8), yaitu kompetensi menyelesaikan persamaan linear satu variabel. Ketidaktuntasan atribut kategori proses yang tertinggi berturut-turut adalah atribut (At.P.27), (At.P.16), (At.P.28), dan (At.P.15) merupakan atribut pada submateri geometri. Hasil penelitian mendukung hasil penelitian yang dilakukan

Tatsuoka, Corter, & Tatsuoka (2004) bahwa para siswa kurang menguasai dalam geometri. Oleh karena itu, pembelajaran pada geometri perlu mendapat perhatian yang lebih serius dibandingkan dengan submateri yang lain.

c. Ketidaktuntasan Atribut Keterampilan

Ketidaktuntasan atribut keterampilan pada matematika mempunyai rerata 23,34. Ketidaktuntasan atribut keterampilan berturut-turut mulai yang tertinggi adalah pada submateri (1) geometri dan pengukuran, (2) aljabar, dan (3) statistika dan peluang. Ketidaktuntasan atribut kategori keterampilan pada butir soal matematika yang tertinggi terjadi pada atribut (At.K.5), yaitu keterampilan menggunakan gambar bangun geometri untuk menentukan panjang sisi-sisi segitiga siku-siku; diikuti atribut (At.K.1), yaitu keterampilan menggunakan perbandingan untuk pemecahan masalah; dan atribut (At.K.3), yaitu keterampilan membuat atau menggunakan sketsa grafik fungsi aljabar sederhana pada sistem koordinat Cartesius.

4. Jenis Kesalahan Siswa

Siswa yang mengalami kesulitan belajar dapat melakukan kesalahan dalam menjawab butir soal. Kesalahan peserta dalam menjawab butir soal didefinisikan sebagai respons terhadap butir soal yang tidak sesuai dengan respons yang diharapkan (Gierl, 2007a). Diagnosis kesalahan yang dilakukan siswa merupakan aktivitas untuk menemukan respons peserta yang tidak sesuai dengan pola respons yang diharapkan atau pola respons yang ideal, berdasarkan gejala yang berupa respons terhadap butir soal matematika.

Penemuan kesalahan didasarkan pada Rubrik Penskoran. Identifikasi kesalahan peserta difokuskan pada atribut-atribut yang tidak dikuasai dan tidak diterapkan dengan tepat. Kesalahan siswa dapat diketahui dari tahapan penyelesaian pada setiap distraktor butir soal UN. Hasil analisis yang dilakukan siswa diklasifikasikan menjadi 4 jenis: (1) kesalahan konsep, (2) kesalahan interpretasi bahasa, (3) kesalahan prosedur, dan (4) kesalahan berhitung. Penemuan jenis kesalahan disajikan pada uraian berikut.

a. Kesalahan Konsep

Kesalahan konsep didefinisikan sebagai penerapan aksioma, definisi, teorema, atau sifat-sifat yang tidak tepat. Kesalahan konsep yang dilakukan peserta pada butir soal 13 sebagai berikut.

“Himpunan penyelesaian dari $-5 - 7x \leq 7 - x$ untuk x bilangan bulat adalah.

- A. $\{-1, 0, 1, 2, 3, \dots\}$ C. $\{\dots, -6, -5, -4, -3, -2\}$
B. $\{-2, -1, 0, 1, 2, \dots\}$ D. $\{\dots, -7, -6, -5, -4, -3\}$ “

Kesalahan konsep yang dilakukan peserta pada butir soal 13 adalah peserta tidak memahami (1) notasi pertidaksamaan, (2) sifat-sifat pertidaksamaan, dan (3) menyatakan himpunan dapat mengakibatkan peserta melakukan kesalahan dalam menjawab butir soal. Definisi, aksioma, teorema, atau sifat-sifat merupakan satu kesatuan yang membentuk suatu konsep matematika. Aksioma dan definisi disusun melalui model matematika, untuk selanjutnya melalui proses berpikir deduktif didapat suatu teorema. Indikator kesalahan konsep adalah peserta tidak dapat menerapkan teorema, rumus, konsep, atau sifat-sifat dengan tepat.

b. Kesalahan Interpretasi Bahasa

Kesalahan interpretasi bahasa didefinisikan sebagai ketidaktepatan peserta dalam memberi arti atau memberi makna suatu kalimat atau ungkapan matematika. Jenis kesalahan interpretasi bahasa yang dilakukan peserta pada butir soal 15 berikut.

“Petugas lalu lintas melakukan pemeriksaan terhadap pengendara kendaraan bermotor. Hasilnya 25 orang memiliki SIM A, 30 orang memiliki SIM C, 17 orang memiliki SIM A dan SIM C, sedangkan 12 orang tidak memiliki SIM A maupun SIM C. Banyaknya pengendara bermotor yang diperiksa adalah ...

- A. 50 orang B. 60 orang C. 72 orang D. 84 orang”

Kesalahan dalam menginterpretasikan bahasa yang dilakukan peserta pada butir soal 15 adalah peserta melakukan kesalahan dalam

memaknai banyaknya pengendara bermotor yang diperiksa adalah $25+30+17-12$, dalam memahami atau memaknai banyaknya pengendara bermotor yang diperiksa adalah $25+30+17$, dalam memahami atau memaknai banyaknya pengendara bermotor yang diperiksa adalah $25+30+17+12$. Kesalahan tersebut menunjukkan bahwa peserta tidak tuntas pada atribut kompetensi melakukan proses pemikiran logis, konsep dasar aljabar: himpunan, penggunaan himpunan, kompetensi menggunakan konsep himpunan dalam pemecahan masalah, dan keterampilan menyajikan himpunan dengan diagram Venn. Peserta melakukan kesalahan pada salah satu atau seluruh tentang (1) keanggotaan suatu himpunan, (2) irisan dua himpunan, (3) gabungan dua himpunan, (4) komplemen suatu himpunan, (5) selisih dua himpunan, (6) bilangan cardinal, atau (7) diagram Venn dapat mengakibatkan peserta melakukan kesalahan dalam menjawab butir soal. Indikator kesalahan interpretasi bahasa adalah (1) peserta tidak dapat memaknai bahasa sehari-hari ke dalam bahasa matematika dengan benar, dan (2) peserta tidak dapat memaknai bahasa matematika ke dalam bahasa sehari-hari dengan benar.

c. Kesalahan Prosedur

Kesalahan prosedur didefinisikan sebagai penerapan langkah-langkah atau tahapan tertentu dalam suatu pemecahan masalah yang tidak tepat. Kesalahan prosedur yang dilakukan peserta pada butir soal 6 sebagai berikut.

“Suatu hari Tono memperkirakan persediaan makanan untuk 60 ekor ayam akan habis dalam 12 hari. Bila hari itu ia membeli lagi 20 ekor ayam, maka persediaan makanan tersebut akan habis dalam waktu ...

- A. 4 hari B. 9 hari C. 16 hari D. 36 hari”

Pemecahan masalah butir soal 6 dilakukan melalui prosedur dengan menentukan banyak hari persediaan makanan ayam tersebut akan habis sama dengan banyak makanan yang tersedia dibagi dengan

banyaknya ayam. Selanjutnya, untuk menentukan banyaknya hari persediaan makanan ayam tersebut akan habis sama dengan banyaknya makanan yang tersedia dibagi dengan banyaknya ayam. Indikator kesalahan prosedur adalah (1) peserta tidak dapat menerapkan algoritma dengan tepat, (2) peserta tidak dapat melakukan manipulasi algoritma dengan betul, dan (3) peserta tidak dapat menggunakan penalaran dengan benar.

d. Kesalahan Berhitung

Kesalahan berhitung didefinisikan sebagai penerapan pengerjaan yang tidak sesuai dengan kaidah-kaidah perhitungan, atau penjabaran dalam penyelesaian masalah. Kesalahan berhitung yang dilakukan peserta pada butir soal 12 sebagai berikut.

“Hasil dari $\frac{1}{x} - x$ adalah ...

A. $\frac{1-x}{x}$ B. $\frac{x-1}{x}$ C. $\frac{x^2-1}{x}$ D. $\frac{1-x^2}{x}$ ”

Kesalahan berhitung yang dilakukan peserta pada butir soal 12 adalah peserta tidak memahami: (1) untuk menyamakan penyebut kedua pecahan, maka dipilih KPK sebagai penyebutnya, KPK penyebut x dan 1 adalah x , dan (2) pengurangan pada pecahan dilakukan dengan menyamakan penyebutnya, kemudian mengurangi pembilangnya, dapat mengakibatkan peserta melakukan kesalahan dalam menjawab butir soal. Indikator kesalahan berhitung adalah (1) peserta tidak dapat melakukan komputasi dengan betul, (2) peserta tidak dapat menerapkan operasi hitung atau operasi bentuk aljabar dengan tepat, dan (3) peserta tidak dapat melakukan perhitungan dengan cermat.

Penemuan jenis kesalahan yang dilakukan peserta pada UN Matematika adalah:

- a. kesalahan konsep yang ditemukan adalah (1) siswa tidak dapat menerapkan teorema, rumus, konsep, atau sifat-sifat dengan tepat,

- (2) siswa tidak dapat menuliskan rumus atau teorema dengan benar, dan (3) siswa tidak dapat merumuskan suatu konsep atau model matematika dengan betul,
- b. kesalahan interpretasi bahasa yang ditemukan adalah (1) siswa tidak dapat memaknai bahasa sehari-hari ke dalam bahasa matematika dengan benar, (2) siswa tidak dapat menginterpretasikan notasi, simbol, grafik, tabel atau gambar ke dalam bahasa matematika dengan benar, dan (3) siswa tidak dapat memaknai bahasa matematika ke dalam bahasa sehari-hari dengan benar,
 - c. kesalahan prosedur yang ditemukan adalah (1) siswa tidak dapat menerapkan algoritma dengan tepat, (2) siswa tidak dapat melakukan manipulasi algoritma dengan betul, dan (3) siswa tidak dapat menggunakan penalaran dengan benar,
 - d. kesalahan berhitung yang ditemukan adalah (1) siswa tidak dapat melakukan komputasi dengan betul, (2) siswa tidak dapat menerapkan operasi hitung atau operasi bentuk aljabar dengan tepat, dan (3) siswa tidak dapat melakukan perhitungan dengan cermat.

Penemuan jenis dan frekuensi kesalahan untuk setiap submateri bilangan, aljabar, geometri dan pengukuran, dan statistika dan peluang disajikan pada Tabel 2.

Tabel 2. Frekuensi Kesalahan Konsep, Kesalahan Interpretasi Bahasa, Kesalahan Prosedur, dan Kesalahan Berhitung

Submateri	Kesalahan			
	Konsep	Interpretasi Bahasa	Prosedur	Berhitung
Bilangan	1561	815	569	1171
Aljabar	4538	2764	1249	3071
Geometri	5741	3497	2218	1794
Statistika	829	1073	392	335
Matematika	12669	8149	4428	6371

Pada Tabel 2 tampak bahwa frekuensi kesalahan tertinggi yang terjadi pada bilangan, aljabar, geometri dan pengukuran adalah jenis kesalahan konsep. Frekuensi kesalahan tertinggi yang terjadi pada statistika dan peluang adalah jenis kesalahan interpretasi bahasa. Jenis kesalahan tertinggi yang terjadi pada matematika adalah jenis kesalahan konsep.

Simpulan

Simpulan berdasarkan hasil analisis dan pembahasan sebagai berikut.

1. Atribut yang mendasari butir soal pada matematika ada 47 atribut, meliputi 4 atribut isi, 36 atribut proses, dan 7 atribut keterampilan.
2. Ketidaktuntasan atribut isi, proses, dan keterampilan yang tertinggi pada geometri dan pengukuran.
3. Kesalahan menjawab butir soal diklasifikasikan menjadi jenis kesalahan konsep, kesalahan interpretasi bahasa, kesalahan prosedur, dan kesalahan berhitung. Jenis kesalahan tertinggi pada bilangan, aljabar, geometri dan pengukuran adalah jenis kesalahan konsep, dan kesalahan tertinggi pada statistika dan peluang adalah jenis kesalahan interpretasi bahasa.
4. Penemuan informasi diagnostik data UN Matematika dapat dilakukan melalui mekanisme identifikasi atribut yang mendasari butir soal, pengembangan rubrik penskoran politomus, perhitungan ketidaktuntasan atribut, dan diagnosis kesalahan yang dilakukan siswa.

Saran dan Rekomendasi

Berdasarkan simpulan disarankan dan direkomendasikan sebagai berikut.

1. Kepada pendidik atau pengguna informasi diagnostik
Informasi tentang atribut yang mendasari butir soal, ketidaktuntasan atribut, dan kesalahan yang dilakukan siswa disarankan untuk dimanfaatkan sebagai umpan balik di dalam perbaikan dan peningkatan kualitas proses pembelajaran.

2. Kepada pemangku kepentingan

Penemuan mekanisme untuk mendayagunakan butir soal dan respons butir pada bentuk soal objektif pilihan ganda menjadi sumber informasi edukatif. Oleh karena, itu kepada pemangku kepentingan direkomendasikan sebagai berikut.

a. Untuk peningkatan kompetensi guru.

Agar menyelenggarakan pelatihan bagi guru dengan materi yang difokuskan pada diagnostik kesalahan jawaban siswa pada tes objektif pilihan ganda pada matematika.

b. Untuk pengembangan butir soal

Agar mengambil kebijakan (1) pengembangan butir soal objektif pilihan ganda dikonstruksi dalam format tes objektif pilihan ganda yang disiapkan untuk diskor politomus, (2) penyusunan butir soal didasarkan pada atribut yang mendasarinya, yang dikembangkan dari kompetensi dasar (KD) agar penggunaan atribut bervariasi dan duplikasi penggunaan atribut dapat diminimalkan, dan (3) pemilihan distraktor didasarkan pada atribut yang mendasari butir soal untuk mengurangi bias dan untuk mendiagnosis ketidaktuntasan atribut, jenis kesalahan yang dilakukan oleh siswa.

3. Kepada para peneliti

Para peneliti disarankan untuk melakukan penelitian yang diarahkan pada penskoran model *Partial Credit*

Daftar Pustaka

- Bond, T. G., & Fox, C. M. (2007). *Applying the rasch model: fundamental measurement in the human sciences*. (2nd Ed.). Mahwah: Lawrence Erlbaum Associates, Publishers.
- De Ayala, R. J. (1993). Methods, plainly speaking: An introduction to polytomous item response theory models. *Measurement and Evaluation in Counseling and Development*, 25, 172-189

- DeMars, C. (2010). *Item response theory*. New York: Oxford University Press, Inc.
- Depdiknas. (2006). *Peraturan Menteri Pendidikan Nasional Nomor 22, Tahun 2006, tentang Standar Isi untuk Satuan Pendidikan Dasar dan Menengah*
- Gierl, M. J. (2007a). Making diagnostic inferences about cognitive attributes using the rule-space model and attribute hierarchy method. *Journal of Educational Measurement*, 44(4), 325-340.
- _____. (6 February 2007b). *Using the attribute hierarchy method to make diagnostic inferences about examinees' cognitive skills*. Makalah disajikan dalam Association of Test Publishers Annual Meeting Palm Springs, CA.
- Gierl, M. J., Yinggan Zheng, & Ying Cui. (2008), Using the attribute hierarchy method to identify and interpret cognitive skills that produce group differences. *Journal of Educational Measurement*, 45 (1), 65-89.
- Hessen, D. J. (Juni, 2009). *Applying IRT using R*. Makalah disajikan dalam Pelatihan Aplikasi Program R dalam Bidang Psikometri dan Penilaian Pendidikan, di PPs Universitas Negeri Yogyakarta.
- Leighton, J. P., & Gierl, M. J. (2007). Defining and evaluating models of cognition used in educational measurement to make inferences about examinees' thinking processes. *Educational Measurement: Issues and Practices*, 26,3-16.
- Leighton, J. P., Gierl, M. J., & Hunka, S. M. (2004). The attribute hierarchy method for cognitive assessment: A variation on tatsuoaka's rule-space approach. *Journal of Educational Measurement*, 41 (3), 205-237.
- Mardapi, Djemari. (2005). Pengembangan sistem penilaian berbasis kompetensi. Dalam HEPI (Eds) *Rekayasa Sistem Penilaian Dalam*

- Rangka Meningkatkan Kualitas Pendidikan* (pp.71-85). Yogyakarta: HEPI
- Roberts, M. R., & Gierl, M. J. (2010). Developing score reports for cognitive diagnostic assessment. *Educational Measurement: Issues and Practice*. 29 (3), 25-38.
- Samejima, F. (1995). A Cognitive method using latent trait models: competency space approach and its relationship with DiBello and Stout's unified cognitive-psychometric diagnosis model. Dalam P.D. Nichols, S.F. Chipman & R.L. Breman. *Cognitively Diagnostic Assessment* (pp. 391-410). New Jersey: Lawrence Erlbaum Associates Publishers.
- Tatsuoka, K., Corter, J.E., & Tatsuoka, C. (2004). Patterns of diagnosed mathematical content and process skills in TIMSS-R a cross a sample of 20 countries. *American Educational Research Journal*, 41(4), 901-926.
- Thissen, D., Nelson, L., Rosa, K., et al. (2001). Item response theory for items scored in more than two categories. Dalam D. Thissen & H. Wainer. *Test Scoring* (pp. 141-184). New Jersey: Lawrence Erlbaum Associates Publishers.
- Wells, C. S., Hambleton, R.K., & Urip Purwono. (18-24 Juni 2008). *Polytomous response IRT models and applications*. Makalah disajikan dalam Pelatihan Asesmen Pendidikan dan Psikologi (Psikometri), di PPs Universitas Negeri Yogyakarta.