

IDENTIFIKASI KESULITAN PESERTA DIDIK DALAM BELAJAR MATEMATIKA DAN SAINS DI SEKOLAH DASAR

Heri Retnawati, Badrun Kartowagiran, Samsul Hadi, dan Kana Hidayati

Program Pascasarjana Universitas Negeri Yogyakarta

email: retnowati_heriuny@yahoo.co.id

Abstrak

Penelitian ini bertujuan untuk mengidentifikasi kesulitan peserta didik dalam pembelajaran matematika dan sains berdasarkan daya serap dan tingkat kesulitan butir tes. Penelitian ini merupakan penelitian deskriptif eksploratif. Data utama yang digunakan adalah respons peserta didik terhadap tes INAP 2007 untuk mata pelajaran matematika dan sains di kelas 5 SD. Tes ini merupakan dokumentasi Puspendik Balitbang Kementerian Pendidikan Nasional untuk wilayah Yogyakarta. Data dianalisis melalui pendekatan deskriptif kuantitatif dengan menggunakan daya serap (proporsi menjawab benar) pada pendekatan teori tes klasik dan tingkat kesulitan model Rasch pada pendekatan teori respons butir. Hasil penelitian menunjukkan bahwa (1) kesulitan peserta didik dalam pembelajaran matematika berdasarkan daya serap meliputi 88,57% dari keseluruhan materi yang seharusnya dikuasai peserta didik; (2) materi untuk mata pelajaran matematika yang dianggap sulit yang dianalisis melalui pendekatan teori respons butir adalah perbandingan dan skala; jarak, waktu, dan kecepatan; operasi hitung campuran; serta luas bangun datar; (3) kesulitan peserta didik dalam pembelajaran sains berdasarkan daya serap meliputi 80% dari keseluruhan materi yang seharusnya dikuasai peserta didik; (4) materi untuk mata pelajaran sains yang dianggap sulit yang diketahui melalui pendekatan teori respons butir adalah proses fotosintesis tumbuhan air, magnet, katrol, perubahan wujud, sistem aliran darah, adaptasi, lapisan bumi, dan tanda gunung berapi meletus.

Kata kunci: kesulitan peserta didik, daya serap, tingkat kesulitan butir tes

IDENTIFICATION OF STUDENT'S LEARNING DIFFICULTIES ON MATHEMATICS AND SCIENCE AT ELEMENTARY SCHOOL

Abstract

This study is aimed at identifying learners' difficulties in the mathematics and science classes based on absorption levels and test-item difficulties. Using the descriptive explorative research method, the main data were collected through students' responses to the 2007 INAP test of the mathematics and science subject for the fifth grade of the elementary school. The test was a document from the Research and Development Body of the Ministry of Education for Yogyakarta area. Data were analysed using proportion of correct answers to the test and Rasch model for item response theory. Research findings show the following results. (1) Based on the absorption level, students' difficulties in mathematics cover 88.57% of all the instructional materials that are expected to be mastered. (2) Based on the item response analyses, difficult materials are ratios and scales; distance, time, and speed; mix math operations; and flat shape width. (3) Students' difficulties in physics cover 80% of all the instructional materials that are expected to be mastered. (4) Difficult materials in physics are photosynthetic processes of water plants, magnets, crane, shape changes, blood circulation system, adaptation, earth layers, and signs of erupting volcanoes.

Keywords: learners' difficulty, absorption level, test-item difficulty level

PENDAHULUAN

Evaluasi ataupun survei pendidikan di Indonesia merupakan hal yang menarik untuk dicermati. Hasil survei internasional, misalnya *Trend International in Mathematics and Science Study* (TIMSS), menunjukkan bahwa kualitas pendidikan di Indonesia relatif rendah dibandingkan dengan negara lain. Pada TIMSS 1999, dari 38 negara yang diteliti prestasi peserta didik SLTP Indonesia menduduki ranking ke-34 yang berada di atas Cili, Maroko, Filipina, dan Afrika Selatan. Pada TIMSS 2003, prestasi peserta didik Indonesia berada pada ranking ke-35 dari 46 negara peserta yang melibatkan lebih dari 200.000 peserta didik. Demikian pula pada TIMSS 2007, hasilnya belum terlalu menggembirakan. Padahal TIMSS merupakan survei yang telah diakui oleh masyarakat dunia untuk mengetahui kompetensi matematika dan sains di antara para pelajar dari berbagai negara. Selain itu, TIMSS diakui representatif untuk mengetahui kompetensi peserta didik karena setiap negara peserta diwakili oleh ribuan pelajar, untuk mengukur kualitas pendidikan antar-negara. Pelajar-pelajar ini berasal dari berbagai sekolah dan berbagai provinsi yang ada di suatu negara di berbagai belahan dunia.

Data empiris tersebut menjelaskan bahwa kemampuan peserta didik Indonesia dalam bidang matematika dan sains masih sangat rendah. Sebagian besar memberikan penjelasan bahwa penyebab rendahnya kemampuan peserta didik ini disebabkan oleh kegagalan di dalam pembelajaran di sekolah. Di sisi lain, kegagalan ini sebagai gambaran umum akan rendahnya kualitas pendidikan di Indonesia.

Hasil survei TIMSS ini ditanggapi prihatin oleh berbagai pihak. Padahal secara teori, matematika dan sains merupakan pengetahuan dan ilmu yang sangat penting bagi manusia untuk menjalani hidup dan kehidupan. Namun, sudah menjadi rahasia umum bahwa keduanya dianggap sebagai

mata pelajaran yang dianggap sulit. Anggapan ini berbanding lurus dengan indikator rendahnya prestasi belajar peserta didik dalam matematika dan sains yang ditunjukkan oleh hasil survei internasional. Permasalahan ini akan menjadi lebih kompleks jika dikaitkan dengan variabel lain dalam pembelajaran matematika dan sains. Salah satu variabel di antaranya adalah sistem evaluasi yang mencakup karakteristik butir dan sistem administrasi tes.

Terdapat dua sistem utama evaluasi secara nasional yang ada di Indonesia, yakni UN (Ujian Nasional) dan INAP (*Indonesian National Assessment Programme*). INAP merupakan program pemerintah untuk mengetahui kualitas pendidikan di Indonesia. Data yang tersedia terkait dengan program ini meliputi data tentang guru, peserta didik, dan sekolah.

Hasil survei INAP pada data peserta didik diperoleh hasil tes (skor mentah) peserta didik untuk mengukur kemampuan matematika, bahasa Indonesia, sains, dan IPS di SD. Dengan menggunakan data ini, materi-materi yang dianggap sulit atau materi yang memiliki daya serap rendah dan tingkat kesulitan yang tinggi berdasarkan teori respons butir dapat diketahui. Berdasarkan hasil analisis ini kesulitan peserta didik dalam menguasai konsep matematika dan sains dapat diketahui.

Terkait dengan upaya peningkatan kualitas pendidikan di Indonesia, data-data INAP terkait dengan peserta didik dapat dimanfaatkan untuk mengidentifikasi kesulitan peserta didik dalam belajar matematika dan sains. Hasil identifikasi ini dapat dimanfaatkan sebagai saran dan masukan dalam pembelajaran yang dilaksanakan oleh guru di sekolah.

Berdasarkan fenomena di atas, diperlukan penelitian tentang identifikasi kesulitan peserta didik dalam belajar matematika dan sains, khususnya di jenjang SD (studi menggunakan respons peserta

didik terhadap tes inap 2007) dalam rangka memanfaatkan hasil asesmen nasional untuk perbaikan pembelajaran. Tujuannya adalah untuk mengidentifikasi kesulitan peserta didik dalam pembelajaran matematika dan sains berdasarkan daya serap tingkat kesulitan butir tes.

Secara teori, evaluasi dalam pendidikan dilaksanakan untuk memperoleh informasi tentang aspek yang berkaitan dengan pendidikan. Menurut Gronlund (1976), evaluasi dalam pendidikan memiliki tujuan untuk memberikan klarifikasi tentang sifat hasil pembelajaran yang telah dilaksanakan, informasi tentang ketercapaian tujuan jangka pendek yang telah dilaksanakan, masukan untuk kemajuan pembelajaran, informasi tentang kesulitan dalam pembelajaran, dan untuk memilih pengalaman pembelajaran di masa yang akan datang. Di samping itu, informasi evaluasi dapat digunakan untuk membantu memutuskan kesesuaian dan keberlangsungan dari tujuan pembelajaran, kegunaan materi pembelajaran, dan untuk mengetahui tingkat efisiensi dan keefektifan strategi pembelajaran (termasuk metode dan teknik belajar-mengajar) yang digunakan. Gronlund (1976) lebih lanjut menjelaskan bahwa evaluasi memiliki fungsi untuk membantu guru dalam penempatan peserta didik dalam kelompok-kelompok tertentu, perbaikan metode mengajar, mengetahui kesiapan peserta didik (sikap, mental, material), memberikan bimbingan dan seleksi dalam rangka menentukan jenis jurusan maupun kenaikan tingkat.

Dalam evaluasi pendidikan, diperlukan alat (instrumen). Salah satu alat yang digunakan untuk melakukan evaluasi adalah tes. Tes ini digunakan untuk mengetahui informasi tentang aspek psikologis tertentu. Tes merupakan suatu prosedur sistematis untuk mengamati dan menggambarkan satu atau lebih karakteristik seseorang dengan suatu skala numerik atau sistem kategorik. Berdasarkan hal ini, tes memberikan informasi

yang bersifat kualitatif dan kuantitatif.

Pelaksanaan evaluasi biasanya dilaporkan. Prosedur, tipe, atau teknik pelaporan hasil evaluasi dapat bervariasi, seperti dalam bentuk huruf, angka, lulus tidak lulus, dan sebagainya. Pada prinsipnya, laporan ini menggambarkan pencapaian/penguasaan peserta didik terhadap materi tertentu.

Laporan hasil evaluasi dapat disusun dan diinterpretasikan secara normatif dan secara kriteria. Evaluasi untuk kepentingan perbaikan hasil belajar mengajar atau penilaian formatif (dikenal pula sebagai *assesment for learning*), acuan yang sebaiknya digunakan adalah acuan kriteria. Hal ini disebabkan karena informasi yang diperoleh berupa penguasaan materi pelajaran yang telah dipelajari. Berdasarkan informasi ini, pendidik akan mengetahui teknik dan strategi mengajar agar materi pelajaran dapat diserap lebih baik.

Hasil evaluasi dapat pula dimanfaatkan untuk penentuan kelulusan. Pada keperluan ini, hasil ujian sebagai hasil pengukuran dipergunakan untuk menentukan seorang peserta didik lulus atau tidak lulus (penilaian sumatif).

Bagi peserta didik, pemanfaatan hasil evaluasi dipergunakan untuk mengetahui apakah dirinya sudah menguasai bahan yang disajikan guru, untuk mengetahui bagian mana yang belum dikuasai, sebagai penguatan bagi peserta didik yang sudah memperoleh skor tertinggi sehingga menjadi motivasi untuk belajar lebih giat, sebagai alat diagnosis bagi peserta didik yang bersangkutan, dengan mengetahui bagian mana yang sukar dikuasai peserta didik.

Bagi guru, pemanfaatan hasil evaluasi harian dapat digunakan untuk mengetahui sejauh mana peserta didik menguasai bahan pelajaran yang diajarkan guru, baik secara kelompok maupun individual; untuk mengetahui bagian mana saja dari materi pelajaran yang belum dikuasai peserta didik, terlebih bagian itu merupakan

prasyarat bagi bahan pelajaran selanjutnya, sehingga dapat melakukan upaya perbaikan; dapat memberikan gambaran baik peserta didik untuk memperkirakan pencapaian keberhasilan terhadap keseluruhan program yang akan dilaksanakannya.

Bagi pengelola pendidikan, mulai dari kepala sekolah, dinas, pengawas, kepala bidang, kepala kantor wilayah, direktur jenderal sampai menteri dapat mengambil manfaat dari hasil evaluasi. Setiap pengelola pendidikan dapat menemukan jawab dari pertanyaan/permasalahan mengenai apakah program pendidikan yang ditetapkan sudah tepat untuk suatu jenjang sekolah; apakah alat/sarana dan prasarana belajar sudah memadai untuk mencapai yang maksimal dari peserta didik; apakah metode penyajian yang disarankan dan petunjuk bagi guru sudah tepat; serta apakah kualitas pendidikan sudah tersebar merata.

Untuk dapat memanfaatkan hasil evaluasi, ada beberapa pendekatan teori yang dapat digunakan, yakni pendekatan teori tes klasik (*Classical Test Theory*) dan teori respons butir (*Item Response Theory*). Teori tes klasik atau disebut teori skor murni klasik (Allen & Yen, 1979) didasarkan pada suatu model aditif, yakni skor amatan merupakan penjumlahan dari skor sebenarnya dan skor kesalahan pengukuran. Ada beberapa parameter butir yang terkait dengan teori tes klasik, yakni proporsi menjawab benar, tingkat kesulitan, reliabilitas, daya pembeda, dan kesalahan pengukuran.

Tingkat kesukaran suatu butir soal yang disimbolkan dengan p_i merupakan salah satu parameter butir soal yang sangat berguna untuk menganalisis suatu tes. Dengan melihat parameter butir ini, akan diketahui seberapa baik kualitas butir soal. Jika p_i mendekati 0, soal tersebut terlalu sukar. Jika p_i mendekati 1, soal tersebut terlalu mudah sehingga perlu dibuang (butir tersebut tidak dapat membedakan kemampuan seorang peserta didik dengan peserta didik lainnya).

Allen dan Yen (1979) menyatakan bahwa secara umum indeks kesukaran suatu butir sebaiknya terletak pada interval 0,3 – 0,7. Pada interval ini, informasi tentang kemampuan peserta didik akan diperoleh secara maksimal. Dalam merancang indeks kesukaran suatu perangkat tes, perlu dipertimbangkan tujuan penyusunan perangkat tes tersebut. Untuk menentukan indeks kesukaran dari suatu butir pada perangkat tes pilihan ganda, digunakan persamaan sebagai berikut.

$$p_i = \frac{\Sigma B}{N} \dots\dots\dots (1)$$

dengan :

p_i = proporsi menjawab benar pada butir soal tertentu

ΣB = banyaknya peserta tes yang menjawab benar

N = jumlah peserta tes yang menjawab

Prosedur penyekoran di dalam evaluasi pendidikan dalam pendekatan teori tes klasik didasarkan atas jawaban yang benar. Jika peserta didik menjawab dengan benar tes pilihan ganda, diberi skor 1 dan jika menjawab dengan salah, diberi skor 0. Prosedur penyekoran semacam ini dinyatakan dengan skor total yang diperoleh peserta didik. Prosedur ini kurang memerhatikan interaksi antara setiap orang peserta didik dengan butir.

Pendekatan teori respons butir merupakan pendekatan alternatif yang dapat digunakan dalam menganalisis suatu tes. Terdapat dua prinsip yang digunakan pada pendekatan ini, yakni prinsip relativitas dan prinsip probabilitas. Pada prinsip relativitas, unit dasar dari pengukuran bukanlah peserta didik atau butir, tetapi lebih kepada kemampuan peserta didik relatif terhadap butir. Jika β_n merupakan indeks dari kemampuan peserta didik ke- n

pada *trait* yang diukur dan δ_i merupakan indeks dari tingkat kesulitan dari butir ke-*i* relatif yang terkait dengan kemampuan yang diukur, bukan β_n atau δ_i yang merupakan unit pengukuran, melainkan lebih didasarkan kepada perbedaan antara kemampuan dan dari peserta didik relatif terhadap tingkat kesulitan butir atau $(\beta_n - \delta_i)$ yang perlu dipertimbangkan. Sebagai alternatifnya perbandingan antara kemampuan terhadap tingkat kesulitan dapat digunakan. Jika kemampuan dari peserta didik melampaui tingkat kesulitan butir, respons peserta didik diharapkan benar. Jika kemampuan peserta didik kurang dari tingkat kesulitan butir, respons peserta didik diharapkan salah (Keeves dan Alagumalai, 1999).

Pada teori respons butir, prinsip probabilitas menjadi perhatian utama. Misalnya, kemampuan peserta didik ke-*n* dinyatakan dengan θ_n dan tingkat kesulitan dari butir dinyatakan dengan Δ_i . Sesuai dengan prinsip relativitas, jika $\theta_n > \Delta_i$ peserta didik diharapkan menjawab dengan benar, maka $\theta_n < \Delta_i$ peserta didik diharapkan menjawab salah.

Dalam teori respons butir, selain asumsi-asumsi yang telah diuraikan sebelumnya, hal penting yang perlu diperhatikan adalah pemilihan model yang tepat. Pemilihan model yang tepat akan mengungkap keadaan yang sesungguhnya dari data tes sebagai hasil pengukuran. Ada tiga model hubungan antara kemampuan dengan parameter butir, yakni model 1 parameter (model Rasch), model 2 parameter, dan model 3 parameter. Model Rasch dituliskan sebagai berikut:

$$P_i(\theta) = \frac{e^{(\theta-b_i)}}{1 + e^{(\theta-b_i)}}, \text{ dengan } i : 1, 2, 3, \dots, n \dots (2)$$

Keterangan:

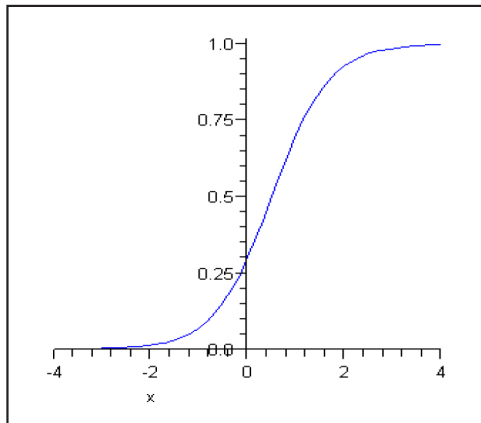
$P_i(\theta)$: probabilitas peserta tes yang memiliki kemampuan θ dipilih secara acak dapat menjawab butir *i* dengan benar

- θ : tingkat kemampuan subyek (sebagai variabel bebas)
- b_i : indeks kesukaran butir ke-*i*
- e : bilangan natural yang nilainya mendekati 2,718
- n : banyaknya butir dalam tes

Parameter b_i merupakan suatu titik pada skala kemampuan agar peluang menjawab benar sebesar 50%. Misalnya, suatu butir tes mempunyai parameter $b_i = 0,3$. Artinya, diperlukan kemampuan minimal 0,3 pada skala untuk dapat menjawab benar dengan peluang 50%. Semakin besar nilai parameter b_i , semakin besar kemampuan yang diperlukan untuk menjawab benar dengan peluang 50%. Dengan kata lain, semakin besar nilai parameter b_i , semakin sulit butir soal tersebut.

Hubungan peluang menjawab benar $P_i(\theta)$ dengan tingkat kemampuan peserta (θ) dapat digambarkan sebagai kurva karakteristik butir (*item characteristic curve* atau *ICC*). Gambar 1 berikut merupakan ilustrasi kurva karakteristik butir untuk model Rasch (1 parameter, 1P), dengan butir 1 ($b=-0,5$), butir 2 ($b=0$) dan butir 3 ($b=0,5$).

Pada model logistik dua parameter, probabilitas peserta tes untuk dapat menjawab benar suatu butir soal ditentukan oleh dua karakteristik butir, yakni indeks kesukaran butir (b_i) dan indeks daya beda butir (a_i). Parameter a_i merupakan indeks daya pembeda yang dimiliki butir ke-*i*. Pada kurva karakteristik, a_i proporsional terhadap koefisien arah garis singgung (*slope*) pada titik $\theta = b$. Butir soal yang memiliki daya pembeda yang besar memunyai kurva yang sangat menanjak, sedangkan butir soal yang memunyai daya pembeda kecil memunyai kurva yang sangat landai. Secara teoretis, nilai a_i ini terletak antara $-\infty$ dan $+\infty$. Pada butir yang baik nilai ini memunyai hubungan positif dengan performen pada butir dengan kemampuan yang diukur, dan a_i terletak antara 0 dan 2 (Hambleton & Swaminathan, 1985).



Gambar 1. Kurva Karakteristik Butir untuk Model 1P, dengan $b=0,5$

Menurut Hambleton, Swaminathan, & Rogers (1991) dan Hullin, *et all* (1983), secara matematis model logistik dua parameter dapat dituliskan sebagai berikut.

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

dengan $i : 1,2,3, \dots, n$ (3)

Keterangan :

- θ : tingkat kemampuan peserta tes
- $P_i(\theta)$: probabilitas peserta tes yang memiliki kemampuan θ dapat menjawab butir i dengan benar
- a_i : indeks daya pembeda
- b_i : indeks kesukaran butir ke- i
- e : bilangan natural yang nilainya mendekati 2,718
- n : banyaknya butir dalam tes
- D : faktor penskalaan yang harganya 1,7

Selain model respons butir dikotomi, ada model lain yang dapat digunakan untuk menyekor respons peserta terhadap suatu butir tes, yakni model politomi. Model-model politomi pada teori respons butir antara lain *nominal respons model* (NRM), *rating scale model* (RSM), *partial credit model* (PCM), *graded respons model* (GRM) dan *generalized*

partial credit model (GPCM) (Van der Linden & Hambleton, 1997).

Model respons butir politomous dapat dikategorikan menjadi model respons butir nominal dan ordinal, tergantung pada asumsi karakteristik tentang data. Model respons butir nominal dapat diterapkan pada butir yang mempunyai alternatif jawaban yang tidak terurut (*ordered*) dan adanya berbagai tingkat kemampuan yang diukur. Pada model respons ordinal terjadi pada butir yang dapat diskor ke dalam banyaknya kategori tertentu yang tersusun dalam jawaban. Skala Likert diskor berdasarkan pedoman penyekoran kategori respons terurut, yang merupakan penyekoran ordinal. Butir-butir tes matematika dapat diskor menggunakan sistem parsial kredit. Langkah-langkah menuju jawaban benar dihargai sebagai penyekoran ordinal. Model penyekoran yang paling sering dipakai ahli salah satunya adalah *Graded Respons Model* (GRM).

Respons peserta terhadap butir j dengan model GRM dikategorikan menjadi $m+1$ skor kategori terurut, $k=0,1,2,\dots,m$ dengan m merupakan banyaknya langkah dalam menyelesaikan dengan benar butir j , dan indeks kesukaran dalam setiap langkah juga terurut. Hubungan parameter butir dan kemampuan peserta dalam GRM untuk kasus homogen (a_j sama dalam setiap langkah) dinyatakan oleh Muraki & Bock (1997) sebagai berikut.

$$P^*(\theta) = P^*_{jk}(\theta) - P^*_{jk+1}(\theta) \dots \dots \dots (4)$$

$$P^*_{jk}(\theta) = \frac{\exp[Da_j(\theta - b_{jk})]}{1 + \exp[Da_j(\theta - b_{jk})]} \dots \dots \dots (5)$$

- Dengan $P^*_{j0}(\theta) = 1$ dan $P^*_{jm+1}(\theta) = 0$
- a_j : indeks daya beda butir j
- θ : kemampuan peserta,
- b_{jk} : indeks kesukaran kategori k butir j
- $P^*_{jk}(\theta)$: probabilitas peserta berkemampuan θ yang memperoleh skor kategori k pada butir j

$P_{jk}^*(\theta)$: probabilitas peserta berkemampuan θ yang memperoleh skor kategori k atau lebih pada butir j
 D : faktor skala

METODE

Penelitian ini merupakan penelitian eksploratif dengan menggunakan pendekatan kuantitatif. Penelitian ini menggunakan dokumentasi INAP 2007, yang berupa butir-butir soal, perangkat tes, dan karakteristik butir INAP 2007, khususnya butir tes matematika dan sains. Pada penelitian ini data yang digunakan adalah data SD kelas 6 di Daerah Istimewa Yogyakarta. Penelitian ini dilaksanakan selama 10 bulan, yakni sejak bulan Februari 2010 sampai dengan November 2010. Analisis data dilakukan secara kuantitatif, dimulai dengan menganalisis keseluruhan butir untuk mengetahui daya serapnya (berdasarkan teori tes klasik) dan tingkat kesulitannya (berdasarkan teori respons butir). Pada data dikotomi estimasi parameter butir dilakukan dengan pendekatan teori respons butir dikotomi. Pada data politomi estimasi parameter butir dilakukan dengan pendekatan teori respons butir politomi. Estimasi parameter butir dan parameter kemampuan dilakukan dengan bantuan *software* Parscale dari SSI (Muraki & Bock, 1997).

HASIL PENELITIAN DAN PEMBAHASAN

Identifikasi kesulitan yang dialami peserta didik didasarkan pada dua pendekatan, yakni daya serap peserta didik dan tingkat kesulitan butir dengan model Rasch. Daya serap peserta didik diketahui dengan proporsi menjawab benar. Proporsi peserta didik yang menjawab benar merupakan tingkat kesulitan butir dengan pendekatan teori tes klasik. Adapun tingkat kesulitan butir dengan model Rasch merupakan teori respons butir dengan pendekatan teori respons butir. Untuk mata pelajaran matematika, masing-masing diuraikan secara terpisah.

Hasil analisis untuk proporsi menjawab benar pada tes matematika disajikan pada Tabel 1. Berdasarkan tabel tersebut, diperoleh gambaran bahwa dari 35 butir soal yang terdiri atas 20 butir pilihan ganda, 10 butir isian singkat, dan 5 butir essay (*constructed response*), hanya sebanyak empat butir soal yang proporsi menjawab benarnya lebih dari 0,65. Dengan kata lain, sebanyak 34 butir soal proporsi menjawab benarnya kurang dari 65%.

Mengingat kurikulum pada saat ini yang dipakai adalah kurikulum berbasis kompetensi, yang menjadi titik tekan evaluasi pun adalah kompetensi peserta didik dalam segi keutantasan belajar. Peserta didik dikatakan mengalami kesulitan jika proporsi menjawab benar kurang dari 65% atau sebesar 0,65. Pada keadaan ini, peserta didik belum dapat dikatakan belajar tuntas untuk mata pelajaran matematika, khususnya materi kelas 5.

Dengan melihat jumlah butir soal yang dikuasai peserta didik sebanyak empat butir (11,43%) dari 35 butir, artinya sebesar 88,57% peserta didik masih mendapatkan kesulitan dalam mengerjakan soal. Kesulitan ini bervariasi dan terjadi pada hampir seluruh materi pada standar kompetensi di kelas 5. Materi yang dianggap sulit meliputi materi bilangan dan operasinya, perbandingan, bangun datar, serta dimensi tiga. Sementara itu, materi yang tergolong ke dalam kategori mudah adalah materi simetri.

Tingkat kesulitan terdapat pula pada bentuk butir soalnya. Dengan mencermati Tabel 1, 10 butir yang memiliki tingkat kesulitan paling tinggi terdapat pada bentuk soal *Constructed Response (CR)* (4 butir soal, 40%), butir isian singkat (5 butir soal, 50%), dan butir pilihan ganda (1 butir soal, 10%). Butir-butir lain yang memiliki tingkat kesulitan didominasi oleh butir isian singkat dan pilihan ganda. Sementara itu, butir-butir yang memiliki tingkat kemudahan adalah butir-butir pilihan ganda.

Model Rasch digunakan untuk mengestimasi tingkat kesulitan butir soal. Hasil estimasi tingkat kesulitan disajikan pada Tabel 2. Butir soal dikatakan mudah jika tingkat kesulitan berada pada skala kurang

dari -2,00. Butir yang tingkat kesulitannya lebih besar dari 2,00 merupakan butir soal dengan tingkat kesulitan tinggi, sedangkan yang tidak termasuk ke dalam kedua-duanya merupakan butir dengan tingkat kesulitan

Tabel 1. Proporsi Menjawab Benar Mapel Matematika Berdasarkan Teori Tes Klasik

No. Butir	Daya Serap	Materi Terkait Kompetensi yang Diukur
CR1	0,0215311	Waktu
Isi5	0,043062201	Operasi bilangan pada pecahan
Isi9	0,052631579	Luas layang-layang
CR3	0,053827751	KPK
CR4	0,061004785	Operasi campuran
CR2	0,08492823	Waktu
PG13	0,098086124	Perubahan suhu
Isi7	0,126794258	Skala
Isi8	0,131578947	Perbandingan
Isi6	0,148325359	Perbandingan
PG20	0,155502392	Jarak
Isi10	0,157894737	Luas bangun datar
CR5	0,159090909	FPB
Isi3	0,188995215	Operasi campuran
PG7	0,232057416	Luas layang-layang
PG17	0,236842105	Kerangka balok
PG4	0,277511962	Operasi perkalian
PG1	0,279904306	Pengurangan
PG12	0,289473684	Luas bangun datar
PG16	0,315789474	Panjang rusuk kubus
PG18	0,330143541	Perkalian pada pecahan
PG5	0,339712919	Kuadrat penjumlahan
PG9	0,351674641	Waktu
Isi2	0,385167464	Volume kubus
PG14	0,392344498	Jaring-jaring kubus
Isi4	0,423444976	FPB
PG19	0,459330144	Sifat bangun datar
PG10	0,468899522	Besar sudut yang dibentuk oleh jarum jam
PG8	0,473684211	Perkalian, pembagian, dan persen
PG6	0,476076555	Akar dari suatu bilangan
Isi1	0,488038278	Sumbu simetri bangun persegi
PG11	0,655502392	Bangun datar
PG2	0,708133971	KPK
PG15	0,796650718	Simetri lipat suatu bangun
PG3	0,801435407	Penjumlahan dan pembagian

sedang. Butir yang termasuk pada kategori sulit yakni butir yang mengukur kompetensi peserta didik untuk penguasaan materi perbandingan, jarak, waktu, serta kecepatan, bilangan dan operasinya.

Seperti halnya pada teori tes klasik, tingkat kesulitan butir yang tinggi terjadi pada butir soal yang berjenis pilihan ganda dan isian singkat. Materi yang dirasakan sulit oleh peserta didik yakni perbandingan dan skala,

Tabel 2. Hasil Estimasi Tingkat Kesulitan Butir Tes Matematika dengan Teori Respons Butir

No. Butir	Materi Terkait Kompetensi yang Diukur	Tingkat Kesulitan
PG1	Pengurangan	1,373
PG2	KPK	-1,311
PG3	Penjumlahan dan pembagian	-2,021
PG4	Operasi perkalian	1,397
PG5	Kuadrat penjumlahan	0,987
PG6	Akar dari suatu bilangan	0,129
PG7	Luas layang-layang	1,762
PG8	Perkalian, pembagian, dan persen	0,153
PG9	Waktu	0,893
PG10	Besar sudut yang dibentuk oleh jarum jam	0,182
PG11	Bangun datar	-0,963
PG12	Luas bangun datar	1,307
PG13	Perubahan suhu	3,206
PG14	Jaring-jaring kubus	0,619
PG15	Simetri lipat suatu bangun	-1,972
PG16	Panjang rusuk kubus	1,121
PG17	Kerangka balok	1,718
PG18	Perkalian pada pecahan	1,028
PG19	Sifat bangun datar	0,248
PG20	Jarak	2,447
Isi1	Sumbu simetri bangun persegi	0,067
Isi2	Volume kubus	0,672
Isi3	Operasi campuran	2,129
Isi4	FPB	0,447
Isi5	Operasi bilangan pada pecahan	4,240
Isi6	Perbandingan	2,529
Isi7	Skala	2,775
Isi8	Perbandingan	2,749
Isi9	Luas layang-layang	4,033
Isi10	Luas bangun datar	2,397
CR1	Waktu	3,302
CR2	Waktu	2,168
CR3	KPK	2,607
CR4	Operasi campuran	2,559
CR5	FPB	1,661

Tabel 3. Proporsi Menjawab Benar untuk Sains

No. Butir	Daya Serap	Materi Terkait Kompetensi yang Diukur
CR5	0,104492	Tanda-tanda gunung merapi meletus
CR1	0,126953	Proses terjadinya daun
Isi4	0,152344	Alat perlindungan diri pada tanaman
Isi1	0,164063	Sistem aliran darah
PG22	0,169922	Proses pengembunan
Isi7	0,195313	Lapisan utama bumi
PG17	0,203125	Katrol untuk mengangkat beban
Isi6	0,214844	Proses perubahan wujud zat
CR4	0,245117	Tata surya
CR3	0,257813	Proses perubahan wujud zat
PG13	0,283203	Cara pembuatan magnet secara induksi
PG21	0,285156	Proses penguapan
PG5	0,287109	Proses terjadinya tumbuhan air
CR2	0,287109	Cara melindungi diri tanaman kaktus
Isi10	0,404297	Sifat kaca
PG8	0,40625	Cara hewan melindungi dirinya
Isi9	0,421875	Penyebab terjadinya gempa bumi
PG23	0,425781	Magnet
Isi5	0,449219	Sifat plastik
PG10	0,451172	Perubahan kimia
PG4	0,460938	Peredaran darah manusia
PG6	0,462891	Proses fotosintesis
PG11	0,472656	Reaksi kimia antara air dan udara
PG18	0,482422	Batuan
PG20	0,488281	Pelapukan batuan
PG1	0,498047	Sistem kekebalan tubuh manusia
PG2	0,535156	Sistem pencernaan
Isi2	0,556641	Proses fotosintesis
PG25	0,587891	Pencegahan banjir
PG24	0,601563	Pelestarian penyusut
PG15	0,615234	Keseimbangan pada benda
PG3	0,636719	Dampak kekurangan mineral
PG19	0,664063	Pelapukan batuan
PG14	0,671875	Cara pembuatan magnet dengan cara aliran listrik
PG16	0,679688	Keseimbangan pada benda
PG7	0,703125	Penyesuaian makhluk hidup terhadap lingkungannya
Isi3	0,708984	Fungsi kaki hewan elang
Isi8	0,755859	Akibat penebangan hutan
PG12	0,806641	Magnet
PG9	0,951172	Cara hewan melindungi dirinya

jarak, waktu, dan kecepatan, operasi hitung campuran, dan luas bangun datar.

Hasil analisis untuk proporsi menjawab benar pada tes sains disajikan pada Tabel 3.

Berdasarkan tabel tersebut, dapat diperoleh gambaran bahwa dari 40 butir soal dengan rincian 25 butir pilihan ganda, 10 butir isian singkat, dan 5 butir essay (*constructed*

Tabel 4. Tingkat Kesulitan Butir Tes Sains Berdasarkan Teori Respons Butir

No. Butir	Materi Terkait Kompetensi yang Diukur	Tingkat Kesulitan
PG1	Sistem Kekebalan tubuh	0,439
PG2	Sistem pencernaan	-0,226
PG3	Dampak kekurangan mineral	-1,788
PG4	Peredaran darah manusia	0,860
PG5	Proses terjadinya fotosintesi tumbuhan air	3,150
PG6	Proses fotosintesis	0,752
PG7	Penyesuaian makhluk hidup terhadap lingkungannya	-2,581
PG8	Cara hewan melindungi dirinya	1,222
PG9	Cara hewan melindungi dirinya	0,000
PG10	Perubahan kimia	0,952
PG11	Reaksi kimia antara air dan udara	0,502
PG12	Magnet	-3,918
PG13	Cara pembuatan magnet secara induksi	2,984
PG14	Cara pembuatan magnet dengan cara aliran listrik	-2,069
PG15	Keseimbangan pada benda	-1,413
PG16	Keseimbangan pada benda	-2,349
PG17	Katrol untuk mengangkat beban	4,358
PG18	Batuan	1,113
PG19	Pelapukan batuan	-2,160
PG20	Pelapukan batuan	0,313
PG21	Proses penguapan	2,741
PG22	Proses pengembunan	3,329
PG23	Magnet	0,279
PG24	Pelestarian penyu	-1,095
PG25	Pencegahan banjir	-1,190
Isi1	Sistem aliran darah	2,768
Isi2	Proses fotosintesis	-3,320
Isi3	Fungsi kaki hewan elang	-3,554
Isi4	Alat perlindungan diri pada tanaman	3,084
Isi5	Sifat plastik	-0,482
Isi6	Proses perubahan wujud zat	3,331
Isi7	Lapisan utama bumi	4,148
Isi8	Akibat penebangan hutan	-4,126
Isi9	Penyebab terjadinya gempa bumi	0,069
Isi10	Sifat kaca	0,265
CR1	Proses terjadinya daun	6,178
CR2	Cara melindungi diri tanaman kaktus	5,638
CR3	Proses perubahan wujud zat	1,848
CR4	Tata surya	1,756
CR5	Tanda-tanda gunung berapi meletus	10,639

response), sebanyak 8 butir soal yang proporsi menjawab benar lebih dari 0,65. Dengan kata lain, pada sebanyak 32 butir soal, proporsi menjawab benar kurang dari 65%.

Dengan mencermati lebih lanjut atas banyaknya butir yang dikuasai peserta didik hanya sebanyak 8 dari 40 butir atau sebanyak 20%, dapat diartikan bahwa sebanyak 80% peserta didik merasa kesulitan dalam mengerjakan soal-soal sains. Tingkat kesulitan ini bervariasi dan terjadi pada hampir seluruh materi pada standar kompetensi di kelas 5. Sementara itu, untuk materi magnet dan adaptasi terhadap lingkungan tergolong ke dalam materi yang berkategori yang mudah.

Tingkat kesulitan juga terdapat pada bentuk butir soalnya. Dengan mencermati kembali Tabel 3, 10 butir yang sulit terjadi karena dipengaruhi pula bentuk soalnya. Seluruh butir *Constructed response* (5 butir soal) merupakan materi sulit. Sebanyak 8 butir isian singkat (dari 10 butir soal), meskipun sebanyak 19 butir pilihan ganda (dari 25 butir soal) merupakan butir yang sulit.

Dengan menggunakan model Rasch, tingkat kesulitan butir soal dapat diestimasi. Hasil estimasi tingkat kesulitan disajikan pada Tabel 4. Tingkat kesulitan butir bervariasi, meskipun mengukur kompetensi dasar yang sama.

Seperti halnya pada teori tes klasik, tingkat kesulitan butir yang tinggi terjadi pada butir soal yang berjenis pilihan ganda dan isian singkat. Materi yang mudah adalah materi tentang menyesuaikan makhluk hidup terhadap lingkungan, magnet, aliran listrik, keseimbangan, pelapukan batuan, proses fotosintesis, adaptasi, dan akibat penebangan hutan. Materi yang sulit adalah proses fotosintesis tumbuhan air, magnet, katrol, perubahan wujud, sistem aliran darah, adaptasi, lapisan bumi, dan tanda gunung berapi meletus.

Peserta pada tes INAP adalah peserta didik kelas 6 SD. Materi yang digunakan sebagai bahan tes adalah bahan pelajaran

yang asumsinya telah dipelajari peserta didik di kelas 4 dan kelas 5. Kisi-kisi tes disusun berdasarkan standar isi meliputi standar kompetensi dan kompetensi dasar yang ditetapkan pemerintah. Butir-butir tes ini telah divalidasi oleh Pusat Penilaian Pendidikan Balitbang Diknas tahun 2007.

Mencermati tingkat kesulitan butir tes, baik ditinjau dari proporsi menjawab benar berdasarkan teori tes klasik maupun berdasarkan teori respons butir untuk mata pelajaran matematika dan sains, dapat dikatakan hasil belajar peserta didik selama ini belum memuaskan. Hasil belajar yang belum memuaskan ini disebabkan adanya indikator sebagian besar materi yang diteskan belum dikuasai secara tuntas oleh peserta didik, yang mengindikasikan belum berhasilnya pembelajaran yang dilakukan oleh guru selama di sekolah.

Belum berhasilnya pembelajaran yang dilakukan oleh guru dapat disebabkan oleh berbagai hal. Penyebab tersebut di antaranya selama ini pembelajaran ditekankan pada pencapaian kompetensi dengan pembelajaran *student center*, belum tentu dilakukan oleh guru. Terlebih lagi keterlibatan peserta didik untuk menemukan konsep matematika dan sains, dan penekanan pembelajaran sebagai proses dalam pendidikan. Misalnya, masih kurangnya pelaksanaan praktikum sains di sekolah dasar. Dengan belum dilaksanakannya kedua hal tersebut, kebermaknaan konsep matematika dan sains dalam diri peserta didik menjadi kurang. Sebagai akibatnya, penguasaan konsep matematika dan sains belum seperti yang diharapkan.

Bentuk soal yang digunakan pada TIMSS juga berbeda dengan tes biasa. Biasanya peserta didik lebih sering dihadapkan pada tes pilihan ganda dan jarang menggunakan soal isian singkat dan uraian. Hal ini juga menyebabkan peserta didik terbiasa dengan jawaban yang disediakan dan menjadi kurang terbiasa dengan mengonstruksi jawaban sendiri melalui penalaran. Hal inilah

yang menyebabkan peserta didik merasa lebih mudah dihadapkan butir soal yang berbentuk pilihan ganda dibandingkan dengan bentuk soal isian singkat maupun *constructed response*.

SIMPULAN

Berdasarkan hasil penelitian dapat disimpulkan sebagai berikut. *Pertama*, kesulitan peserta didik dalam pembelajaran matematika berdasarkan daya serap meliputi 88,57% dari keseluruhan materi yang seharusnya dikuasai peserta didik. *Kedua*, materi yang dirasakan sulit untuk mata pelajaran matematika oleh peserta didik yang diketahui dengan pendekatan teori respons butir yakni perbandingan dan skala, jarak, waktu, dan kecepatan, operasi hitung campuran, dan luas bangun datar. *Ketiga*, kesulitan peserta didik dalam pembelajaran sains berdasarkan daya serap meliputi 80% dari keseluruhan materi yang seharusnya dikuasai peserta didik. *Keempat*, materi yang sulit untuk pelajaran sains yang diketahui dengan pendekatan teori respons butir yaitu proses fotosintesis tumbuhan air, magnet, katrol, perubahan wujud, sistem aliran darah, adaptasi, lapisan bumi, tanda gunung berapi meletus.

Terkait dengan hasil penelitian dapat disarankan hal-hal berikut. *Pertama*, terkait dengan sebagian besar peserta didik mengalami kesulitan dalam pembelajaran matematika dan sains, terlebih lagi peserta didik kelas 5 sekolah dasar, maka pendidik perlu merefleksikan. *Kedua*, kembali pembelajaran yang telah dilaksanakan dan kemudian merencanakan perbaikan. *Ketiga*, dalam pembelajaran, perlu digunakan pembelajaran yang menekankan pendekatan poses dan berpusat pada peserta didik, sehingga berbagai kompetensi matematika dan sains dapat dikuasai. *Keempat*, pembelajaran perlu dilakukan dengan Perbaikan pembelajaran dapat pula menggunakan hasil-hasil evaluasi yang telah dilaksanakan, baik oleh pemerintah maupun

sekolah (*assessment for learning*). *Kelima*, dalam pelaksanaan pendidikan, khususnya dalam evaluasi, perlu digunakan macam-macam jenis tes, tidak hanya pilihan ganda saja, misalnya menggunakan uraian, sehingga peserta didik dapat melatih kemampuan penalarannya untuk mengkonstruksi sendiri jawabannya.

DAFTAR PUSTAKA

- Allen, M. J. & Yen, W. M. 1979. *Introduction to Measurement Theory*. Monterey, CA: Brooks/Cole Publishing Company.
- Gronlund, N.E. 1976. *Measurement and Evaluation in Teaching*. New York: Macmillan Publishing Co.
- Hambleton, R.K. & Swaminathan, H. 1985. *Item Response Theory*. Boston, MA: Kluwer Inc.
- Keeves, J.P. dan Alagumalai, S. 1999. "New Approaches to Measurement". Dalam Masters, G.N. dan Keeves, J.P. (Eds). *Advances in Measurement in Educational Research and Assessment*. Amsterdam: Pergamon.
- Muraki, E., & Bock, R.D. 1997. *Parscale 3: IRT Based Test Scoring and Item Analysis for Graded Items and Rating Scales*. Chicago: Scientific Software Inc.
- Van der Linden, W.J. dan Hambleton, R.K. 1997. "Item Response Theory: Brief History, Common Models and Extensions". Dalam Van der Linden, W.J. dan Hambleton, R.K. (Eds). *Handbook of Item Response Theory*. New York: Springer.