
PENGUNAAN APLIKASI *ITEM AND TEST ANALYSIS* (ITEMAN) PADA SOAL *TRY OUT* UN IPA SMP TAHUN 2019**Nuril Huda¹, Tutik Sri Wahyuni²**

1 Jurusan Tadris Matematika, Fakultas Ilmu Tarbiyah dan Keguruan, UIN Maulana Malik Ibrahim, Jl. Gajayana No.50, Dinoyo, Kec. Lowokwaru, Kota Malang, Jawa Timur 65144, Indonesia

2 Jurusan Tadris Kimia, Fakultas Tarbiyah dan Ilmu Keguruan, IAIN Tulungagung, Jl. Mayor Sujadi No.46, Kudus, Plosokandang, Kec. Kedungwaru, Kabupaten Tulungagung, Jawa Timur 66221, Indonesia

Email : nurilhuda26@uin-malang.ac.id; tswahyuni@gmail.com

Abstract

The aims of this study are to (1) determine the use of ITEMAN application in analyzing try out of Junior High School Sciences National Exam; 2) find out the number of good questions based on classical test theory (CCT). This research is a descriptive study with a quantitative approach. The data source is from the Student Answer Sheet in Tulungagung Regency with a random sample of 1,777. Data collection techniques with documentation. The results of this study: 1) the use of ITEMAN application in analyzing the items obtained: a) the value of the reliability coefficient $\alpha = 0.84$ (very high), split-half (first-last) = 0.72 (high), and S-B (first-last) = 0.84 (very high) and measurement error of 2.41; b) difficulty level: 3 items (10.00%) are easy, 24 items (80.00%) are moderate, and 3 items (10.00%) are difficult; c) discrimination: 3 items (10.00%) are bad, 6 items (20.00%) are good enough, 5 items (15.67%) are moderate, and 16 items (53.33%) are good; and the effectiveness of the distractor: 27 items (90.00%) the distractor functionate, and 3 items (10.00%) the distractor does not function. 2) There are 19 items (63.33%) good questions based on CCT, they can be stored as question bank.

Keywords: ITEMAN; National Exam try out; Sciences; CCT.

Abstrak

Tujuan penelitian ini: 1) mengetahui penggunaan aplikasi Item and Test Analysis (ITEMAN) dalam menganalisis butir soal try out UN IPA SMP tahun 2019; 2) mengetahui jumlah soal yang baik berdasarkan teori tes klasik (TTK). Penelitian ini merupakan penelitian deskriptif dengan pendekatan kuantitatif. Sumber data dari Lembar Jawaban Siswa (LJS) di Kabupaten Tulungagung pada 26 Maret 2019 dengan populasi sebanyak 10.842 LJS dan sampel diambil secara random sebanyak 1.777 LJS. Teknik pengumpulan data dengan dokumentasi. Hasil penelitian ini: 1) penggunaan aplikasi ITEMAN dalam menganalisis butir soal diperoleh: a) nilai koefisien reliabilitas dengan formula ($\alpha = 0,84$ (sangat tinggi); split-half (first-last) = 0,72 (tinggi); dan S-B (first-last) = 0,84 (sangat tinggi) serta kesalahan pengukuran sebesar 2,41; b) tingkat kesukaran: 3 butir (10,00%) mudah, 24 butir (80,00%) sedang, dan 3 butir (10,00%) sukar; c) daya beda: 3 butir (10,00%) jelek, 6 butir (20,00%) belum memuaskan, 5 butir (15,67%) memuaskan, dan 16 butir (53,33%) sangat memuaskan; dan efektivitas distraktor: 27 butir (90,00%) distraktor berfungsi mengecoh peserta try out, dan 3 butir (10,00%) distraktor tidak berfungsi. 2) Ada 19 butir (63,33%) soal yang baik berdasarkan teori tes klasik pada soal try out UN IPA SMP tahun 2019, sehingga dapat disimpan sebagai bank soal.

Katakunci: ITEMAN; try out UN; IPA; TTK.

Dikirim: 27 September 2019; Diperbaiki: 25 April 2020; Dipublikasi: 30 Juni 2020.

PENDAHULUAN

Salah satu tujuan nasional negara Indonesia adalah mencerdaskan kehidupan bangsa. Tujuan ini dapat dicapai melalui pelaksanaan pendidikan. Standar Nasional Pendidikan (SNP) memuat tentang kriteria minimal sistem pendidikan yang berlaku di seluruh wilayah Indonesia. Hal ini dimuat dalam Peraturan Pemerintah No. 19 Tahun 2005. Selain itu, dalam SNP juga memuat delapan standar pendidikan, yaitu standar isi, standar proses, standar pendidik dan tenaga kependidikan, standar sarana dan prasarana pendidikan, standar pengelolaan, standar pembiayaan, standar kompetensi lulusan dan standar penilaian.

Standar kompetensi lulusan merupakan kualifikasi kemampuan lulusan yang mencakup pengetahuan, keterampilan dan sikap, sedangkan standar penilaian berkaitan dengan mekanisme, prosedur, dan penilaian hasil belajar peserta didik. Untuk mengetahui standar kompetensi lulusan, pemerintah menyelenggarakan Ujian Nasional. Hal ini sesuai dengan Peraturan Menteri Pendidikan dan Kebudayaan RI No 4 tahun 2018.

Ujian nasional yang diselenggarakan oleh pemerintah merupakan penilaian hasil belajar untuk mengukur kemampuan peserta didik selama belajar pada jenjang tertentu dalam rangka pengendalian mutu pendidikan nasional.

Salah satu hasil ujian nasional digunakan sebagai dasar untuk pertimbangan seleksi masuk jenjang pendidikan berikutnya. Pihak sekolah berusaha memberikan drill soal untuk fokus Ujian Nasional (UN) baik sebelum atau sesudah jam efektif di sekolah. Selain itu pihak sekolah juga memberikan latihan soal (*try out*) yang disusun sesuai dengan kisi-kisi UN yang ditetapkan oleh Badan Standar Nasional Pendidikan (BSNP). Upaya tersebut sejalan dengan teori belajar behavioristik. Pada teori behavioristik, pengembangan pembelajaran diarahkan pada penyelesaian tugas atau penguasaan pengetahuan secara sistematis (Budiyanto, 2016: 113). Peserta didik harus dibimbing, diberikan stimulus untuk melakukan apa yang mereka pelajari melalui drill soal. Dalam metode drill soal, peserta didik terlebih dahulu dibekali dengan pengetahuan secara teori, kemudian guru memberikan contoh soal dan latihan soal agar dikerjakan oleh peserta didik dengan bimbingan guru. Stimulus ini dilakukan secara berulang-ulang untuk menyempurnakan keterampilan dalam mengerjakan soal. Adanya stimulus tersebut akan saling berinteraksi sehingga mempengaruhi respon yang dihasilkan, yaitu diharapkan peserta didik mampu mengerjakan soal UN dengan baik.

Sejak tahun akademik 2015/2016, Pusat Penilaian Pendidikan (Puspendik) melibatkan guru-guru untuk menyusun kisi-kisi soal UN. Selanjutnya, kisi-kisi soal UN tersebut ditetapkan oleh BSNP menjadi rujukan dalam penyusunan soal UN (BSNP, 2017: 17). Tim pembuat soal terdiri dari dosen, ahli pengukuran, dan guru-guru di sekolah. Guru yang dapat bergabung sebagai tim penyusun soal UN harus mengikuti seleksi. Oleh karena itu, guru-guru di sekolah harus mengetahui kaidah dan tata cara penulisan soal yang baik dan benar sesuai dengan panduan penulisan soal 2017. Panduan tersebut dibuat oleh Pusat Penilaian Pendidikan (Puspendik) Balitbang Kemdikbud. Hal tersebut juga menjadi program kerja dari Musyawarah Guru Mata Pelajaran (MGMP) di setiap kabupaten agar dapat menyusun atau membuat soal latihan (*try out*) UN sesuai dengan kisi-kisi yang sudah ditetapkan.

Ilmu Pengetahuan Alam (IPA) merupakan satu dari empat mata pelajaran yang diujikan pada UN di jenjang Sekolah Menengah Pertama (SMP)/Madrasah Tsanawiyah (MTs). Matapelajaran IPA di jenjang SMP/MTs mencakup materi fisika, biologi, dan kimia. Menyongsong pelaksanaan UN jenjang SMP/MTs tahun 2019, MGMP IPA di bawah koordinasi Dikpora Bidang Pendidikan Dasar Kabupaten Tulungagung menyusun soal *try out* UN. Proses penyusunan soal *try out* UN mata pelajaran IPA oleh MGMP disesuaikan dengan kisi-kisi yang ditetapkan BSNP dan memenuhi validitas isi. Meskipun hanya sebagai latihan, soal-soal *try out* harus dibuat sebaik mungkin, karena baik buruknya soal akan mempengaruhi perolehan hasil belajar siswa. Hasil *try out* diharapkan dapat memberikan gambaran tentang kesiapan siswa menghadapi UN. Oleh karena itu, soal yang akan diujikan sebaiknya dianalisis terlebih dahulu melalui tahap telaah kualitatif dan empiris (kuantitatif). Huda (2015) mengemukakan bahwa analisis butir soal secara kualitatif dilakukan menggunakan pendekatan Teori Tes Klasik (TTK) dan teori tes modern. Oleh karena soal *try out* UN IPA Kabupaten Tulungagung tersebut belum diketahui karakteristik butir soal berdasarkan Teori Tes Klasik (TTK), maka perlu dilakukan pengujian secara empiris terhadap soal-soal tersebut.

Pengujian secara empiris dilakukan untuk mengetahui analisis kuantitatif butir soal. Salah satu program yang digunakan dalam analisis kuantitatif butir soal adalah ITEMAN. *Item and Test Analysis* (ITEMAN) merupakan aplikasi *Windows* yang dirancang untuk memberikan laporan rinci butir tes dan analisis pengujiannya menggunakan teori tes klasik (Sutrisno, dkk; 2018: 5). Teori tes klasik (TTK) merupakan teori yang menggunakan model matematika yang sangat sederhana untuk menunjukkan hubungan antara skor amatan, skor sebenarnya, dan skor kesalahan, dengan Rumus $X = T + E$. Model ini diikuti dengan sejumlah asumsi untuk menyederhanakan formula dalam mengestimasi indeks kehandalan dan kesahihan suatu instrumen (Mardapi, 2012: 52). Selanjutnya, Mardapi juga menjelaskan asumsi-asumsi dari teori tes klasik yang mudah dipahami yaitu 1) instrumen hanya mengukur satu dimensi. Instrumen atau tes yang dirancang menguji IPA, maka dimensi yang diukur juga hanya IPA; 2) tidak ada hubungan antara skor murni dan skor kesalahan; 3) Tidak ada hubungan antara kesalahan pada pengukuran pertama dengan kesalahan pada pengukuran kedua; 4) Tidak ada hubungan antara skor murni pada pengukuran pertama dengan kesalahan pada pengukuran kedua; 5) Tidak ada hubungan antara skor murni pada pengukuran kedua dengan kesalahan pada pengukuran pertama; dan 6) Rata kesalahan pengukuran pada populasi adalah 0.

Asumsi-asumsi teori tes klasik sebagaimana disebutkan di atas, memungkinkan untuk dikembangkan dalam rangka pengembangan berbagai formula yang berguna dalam melakukan pengukuran psikologis. Reliabilitas, kesalahan pengukuran, tingkat kesukaran, daya beda dan efektivitas distraktor adalah formula penting dalam pemilihan butir soal. Reliabilitas merupakan koefisien yang menunjukkan tingkat keajegan atau konsistensi hasil pengukuran suatu tes (Azwar, 2013:37). Konsisten hasil pengukuran tersebut mengacu pada

penggunaan alat ukur yang sama untuk orang yang berbeda atau pada waktu yang berbeda tetapi kondisi yang sama. Kesalahan pengukuran (*standard error of measurement/ SEM*) merupakan estimasi tentang bagaimana seorang peneliti evaluasi memperkirakan kesalahan dari tes yang telah dibuat (Sukardi, 2010:50). Kesalahan pengukuran pada umumnya dapat juga menunjukkan tingkat reliabilitas tes. Jika nilai kesalahan pengukuran suatu tes yang telah dibuat kecil, berarti reliabilitas tes tersebut tinggi. Sebaliknya, jika nilai kesalahan pengukuran besar, berarti bahwa tes yang telah dibuat mempunyai reliabilitas rendah. SEM menyajikan ukuran variabilitas dalam skor sebagai dasar dari adanya kesalahan pengukuran.

Tingkat kesukaran ditentukan sebagai persentase atau proporsi dari pengambil tes yang menjawab benar (Reynolds, Livingston, & Wilson, 2010:148). Tingkat kesukaran dianalisis dengan tujuan untuk melihat apakah soal yang dibuat sudah sesuai dengan kemampuan peserta didik yang mengikuti tes. Proporsi peserta didik menjawab benar sedikit (kecil) maka dapat disimpulkan bahwa tingkat kesukaran itu tinggi, sebaliknya apabila proporsi menjawab benar banyak (besar) maka dapat disimpulkan bahwa tingkat kesukaran tersebut rendah.

Daya beda merupakan kemampuan item dalam membedakan antara siswa yang memiliki kemampuan tinggi dengan siswa yang memiliki kemampuan rendah (Azwar, 2013:137). Daya beda butir soal yang baik, ketika tes atau butir soal mengukur hal yang sama bahwa setiap peserta tes yang mampu dapat menjawab soal dengan benar dan peserta tes yang tidak mampu akan menjawab salah. Analisis distraktor mengizinkan penyusun tes untuk memeriksa grup yang berkemampuan tinggi dan grup berkemampuan rendah memilih masing-masing pada pilihan ganda yang telah disediakan (Reynolds, Livingston, & Willson, 2010:157). Pada soal pilihan ganda alternatif jawaban yang tidak benar bisa dijadikan pengecoh (distraktor) bagi siswa yang sebenarnya tidak mengetahui alternatif jawaban yang benar. Distraktor akan memberikan informasi kepada guru tentang siswa yang benar-benar menguasai materi pelajaran dengan siswa yang kurang menguasai materi pelajaran.

Program ITEMAN yang digunakan dalam menganalisis butir soal ini dapat digunakan untuk: (1) menganalisis data file (format ASCII) jawaban butir soal yang dihasilkan melalui *manual entry* data atau dari mesin *scanner*; (2) menskor dan menganalisis data soal pilihan ganda dan skala Likert untuk 30.000 siswa dan 250 butir soal; (3) menganalisis sebuah tes yang terdiri dari 10 skala (subtes) dan memberikan informasi tentang kuantitatif setiap butir (daya pembeda, tingkat kesukaran, efektivitas distraktor), reliabilitas (KR-20/Alpha, *split-half*, dan *Spearman-Brown (S-B)*), *standar error of measurement*, *mean*, *variance*, *standar deviasi*, *skew*, *kurtosis* untuk jumlah skor pada jawaban benar, skor minimum dan maksimum, skor median, dan frekuensi distribusi skor.

Tujuan penelitian ini: 1) mengetahui penggunaan aplikasi ITEMAN dalam menganalisis butir soal *try out* UN IPA SMP Kabupaten Tulungagung tahun 2019 berdasarkan teori tes klasik (reliabilitas, kesalahan pengukuran, tingkat kesukaran, daya beda, dan efektivitas distraktor); 2) mengetahui jumlah soal yang baik berdasarkan teori tes klasik pada soal *try out* UN IPA SMP tahun 2019. Temuan penelitian ini diharapkan dapat memberikan manfaat khususnya kepada tim MGMP IPA Kabupaten Tulungagung yang menyusun soal *try out* UN tersebut.

METODE PENELITIAN

Penelitian ini merupakan penelitian deskriptif. Penelitian ini menggambarkan data kuantitatif yang diperoleh menyangkut keadaan subjek atau respons dari suatu populasi, sehingga pendekatan yang digunakan adalah pendekatan kuantitatif. Sumber data penelitian ini berasal dari lembar jawaban peserta *try out* UN SMP tahun 2019 yang dilaksanakan serentak di Kabupaten Tulungagung pada 26 Maret 2019. Jumlah soal *try out* UN IPA SMP tahun 2019 adalah 30 soal, terdiri 12 soal fisika (No. 91 sampai 102); 14 soal biologi (No. 103 sampai 115); dan 4 soal kimia (No. 116 sampai 120). Sumber data tersebut oleh peneliti dijadikan sebagai populasi penelitian sebanyak 10.842 lembar jawab siswa. Jumlah sampel diambil secara *random* sebanyak 1.777 lembar jawab siswa. Jumlah sampel yang diambil 16% dari populasi, karena populasi lebih dari 100 subjek dapat diambil 10-15% dari populasi (Arikunto: 2010, 112). Teknik pengumpulan data yang digunakan adalah teknik dokumentasi.

Aplikasi ITEMAN versi 4.3 digunakan untuk menganalisis butir soal berdasarkan teori tes klasik (TTK). Parameter analisis butir soal secara empiris meliputi tingkat kesukaran, daya beda, efektivitas distraktor, reliabilitas, dan kesalahan pengukuran. Butir soal yang baik dapat dipakai kemudian hari/ disimpan sebagai bank soal.

Analisis reliabilitas suatu tes ada beberapa cara atau formula, yaitu formula belah dua, Kuder-Richardson, dan Cronbach *alpha*. Hasil perhitungan dari formula tersebut disebut dengan koefisien reliabilitas. Dari *output* ITEMAN versi 4.3 ada 3 formula yaitu *Alpha* (KR-20), *split-half*, dan *Spearman-Brown (S-B)* dengan kriteria yang digunakan pada Tabel 1.

Tabel 1. Kriteria Koefisien Reliabilitas

Koefisien Reliabilitas	Makna
$r < 0,20$	Sangat Rendah
$0,20 < r \leq 0,40$	Rendah
$0,40 < r \leq 0,60$	Sedang
$0,60 < r \leq 0,80$	Tinggi
$0,80 < r \leq 1,00$	Sangat Tinggi

Sumber: (Setyawarno, 2017:15)

Analisis tingkat kesukaran butir soal dengan menggunakan program komputer ITEMAN versi 4.3 dapat dilihat pada kolom *P* dengan kriteria yang digunakan pada Tabel 2.

Tabel 2. Kriteria Tingkat Kesukaran

Kriteria <i>P</i>	Keterangan
$P > 0,70$	Mudah
$0,30 \leq P \leq 0,70$	Sedang
$P < 0,30$	Sukar

Sumber: (Surapranata, 2004:21)

Analisis daya beda butir soal dengan menggunakan program komputer ITEMAN versi 4.3 dapat dilihat pada kolom *Rpbis* dengan kriteria yang digunakan pada Tabel 3.

Tabel 3. Kriteria Daya Beda

Kriteria	Keterangan
0,40 – 1,00	Baik (tanpa revisi)
0,30 – 0,39	Sedang (tidak perlu direvisi)
0,20 – 0,29	Cukup Baik (perlu direvisi)
Negatif – 0,19	Buruk

Sumber: (Azwar, 2013:140)

Informasi efektivitas distraktor juga dapat diperoleh pada program komputer komputer ITEMAN versi 4.3 yaitu pada kolom *Prop*. Distraktor dikatakan berfungsi apabila nilai dari *Prop* pada setiap pilihan ganda memiliki nilai yang lebih besar dari nilai 0,05 (Mardapi, 2008:143). Nilai proporsi dari opsi setiap butir soal yang memiliki nilai kurang dari nilai *Prop*, maka distraktor tersebut perlu direvisi. Analisis butir soal berdasarkan teori tes klasik (TTK) menggunakan kriteria: a) item yang baik, apabila mempunyai tingkat kesukaran $0,30 \leq P \leq 0,70$ (sedang), daya beda minimal 0,30 (sedang); dan distraktor untuk tiap respons minimal 5%; dan b) item tidak baik, apabila salah satu parameter tersebut tidak terpenuhi.

HASIL PENELITIAN DAN PEMBAHASAN

Penggunaan aplikasi ITEMAN versi 4.3 berdasarkan teori tes klasik (TTK) untuk analisis butir soal *try out* UN IPA SMP tahun 2019. Analisis tersebut menghasilkan informasi tentang butir soal yang baik dapat dipakai kemudian hari atau disimpan sebagai bank soal berdasarkan parameter butir soal, yaitu reliabilitas dan kesalahan pengukuran, tingkat kesukaran, daya beda, serta efektivitas distraktor.

Hasil *output* dari aplikasi ITEMAN versi 4.3 untuk analisis reliabilitas direkap pada Tabel 4.

Tabel 4. Rekap Nilai Koefisien Reliabilitas

Formula	Nilai Koefisien Reliabilitas	Keterangan
<i>Alpha/ KR-20</i>	0,84	Sangat tinggi
<i>Split-Half (First-Last)</i>	0,72	Tinggi
<i>Split-Half (Odd-Even)</i>	0,75	Tinggi
<i>S-B (First-Last)</i>	0,84	Sangat tinggi
<i>S-B (Odd-Even)</i>	0,85	Sangat tinggi

\Berdasarkan Tabel 4, nilai koefisien reliabilitas soal *try out* UN IPA SMP tahun 2019 termasuk sangat tinggi dan kesalahan pengukurannya sebesar 2,41 artinya adanya indikasi yang mempengaruhi skor sebenarnya dengan skor amatan, indeks reliabilitas dapat digunakan dengan mengestimasi seberapa besar skor berfluktuasi sebagai kesalahan pengukuran.

Hasil *output* dari aplikasi ITEMAN versi 4.3 untuk analisis tingkat kesukaran butir soal dapat dilihat pada kolom *P* yang direkap pada Tabel 5.

Tabel 5. Rekap Tingkat Kesukaran Butir Soal

Kriteria	Jumlah	Persentase	Materi
Mudah	3	10,00	3 soal biologi
Sedang	24	80,00	11 soal fisika, 10 soal biologi, 3 soal kimia
Sukar	3	10,00	1 soal fisika, 1 soal biologi, 1 soal kimia
Total	30	100,00	

Berdasarkan Tabel 5, bahwa 80% butir soal *try out* UN IPA SMP tahun 2019 memiliki tingkat kesukaran sedang. Soal dengan tingkat kesukaran mudah yaitu soal No. 103, 107, dan 114; soal dengan tingkat kesukaran sedang yaitu No. 91, 92, 93, 94, 95, 96, 97, 98, 99, 101, 102, 104, 105, 106, 108, 109, 111, 112, 113, 115, 116, 117, 118, dan 120; soal dengan tingkat kesukaran sukar yaitu No.100, 110, dan 119.

Hasil *output* dari aplikasi ITEMAN versi 4.3 untuk analisis daya beda butir soal dapat dilihat pada kolom *Rpbis* yang direkap pada Tabel 6.

Tabel 6. Rekap Daya Beda Butir Soal

Kriteria	Jumlah	Persentase	Materi
Baik (tanpa revisi)	16	53,33	9 soal fisika, 6 soal biologi, 1 soal kimia
Sedang (tidak perlu direvisi)	5	15,67	1 soal fisika, 3 soal biologi, 1 soal kimia
Cukup Baik (perlu direvisi)	6	20,00	2 soal fisika, 3 soal biologi, 1 soal kimia
Buruk	3	10,00	2 soal biologi, 1 soal kimia
Total	30	100,00	

Berdasarkan Tabel 6, bahwa 70,00% butir soal *try out* UN IPA SMP tahun 2019 mampu membedakan siswa yang memiliki kemampuan tinggi dengan rendah. Soal dengan daya beda baik yaitu No. 92, 93, 94, 96, 97, 98, 99, 101, 102, 104, 105, 106, 108, 111, 113, dan 117; soal dengan daya beda sedang yaitu No. 91, 103, 109, dan 112; soal dengan daya beda cukup baikn yaitu No. 95, 100, 114, 115, 116, dan 118; soal dengan daya beda buruk yaitu No. 107, 110, dan 119.

Hasil *output* dari aplikasi ITEMAN versi 4.3 untuk analisis efektivitas distraktor dapat dilihat pada kolom *Prop* pada rekapnya Tabel 7.

Tabel 7. Rekap Efektivitas Distraktor

Keputusan	Jumlah	Persentase	Materi
Efektif	27	90,00	12 soal fisika, 10 soal biologi, 3 soal kimia
Tidak Efektif	3	10,00	2 soal biologi, 1 soal kimia
Total	30	100	

Berdasarkan Tabel 7. bahwa 90,00% distraktor yang dibuat tim MGMP IPA Kabupaten Tulungagung pada soal *try out* UN IPA SMP tahun 2019 efektif mengecoh para peserta *try out*. Soal yang memiliki distraktor tidak efektif yaitu No. 104, 105, dan 118. Selain ketiga soal tersebut, distraktor opsi jawaban pilihan ganda efektif mengecoh peserta didik.

Butir soal yang baik harus memenuhi parameter tingkat kesukaran (sedang), daya beda (sedang), dan distraktor berfungsi sesuai rekap Tabel 8.

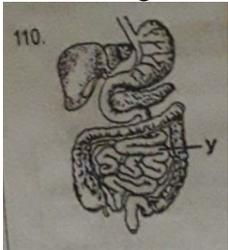
Tabel 8. Rekap Butir Soal *Try Out* UN IPA SMP Tahun 2019

Kesimpulan	Jumlah	Persentase	Materi
baik	19	63,33	10 soal fisika, 7 soal biologi, 2 soal kimia
tidak baik	11	36,67	2 soal fisika, 7 soal biologi, 2 soal kimia
Total	35	100	

Tabel 8 menunjukkan bahwa 19 butir (63,33%) soal *try out* UN IPA SMP tahun 2019 memiliki karakteristik baik, sehingga dapat dipakai kemudian hari/ disimpan sebagai bank soal. Soal yang memiliki karakteristik baik yaitu No. 91, 92, 93, 94, 96, 97, 98, 99, 101, 102, 105, 106, 108, 109, 111, 112, 113, 117, dan 120. Soal memiliki karakteristik tidak baik yaitu No. 95, 100, 103, 104, 107, 110, 114, 115, 116, 118, dan 119.

Analisis butir soal *try out* UN IPA SMP tahun 2019 yaitu No. 110 termasuk soal yang tidak baik karena tingkat kesukaran (sukar) dan daya beda (jelek) pada sub mapel biologi sesuai dengan Gambar 1.

Perhatikan gambar organ pencernaan berikut.



110.

Pada bagian (Y) terjadi proses pencernaan

- Amilum menjadi maltosa oleh enzim tripsin
- Amilum menjadi maltosa oleh enzim amilase
- Maltosa menjadi glukosa oleh enzim amilase
- Maltosa menjadi glukosa oleh enzim tripsin

Gambar 1. Contoh Butir Soal yang Tidak Baik No. 110

Berdasarkan Gambar 1, peserta *try out* banyak memilih opsi B, sedangkan kunci jawabannya C karena di dalam usus, maltase yang termasuk jenis enzim amilase menghidrolisis maltosa menjadi glukosa, yang nantinya akan diserap ke dalam sirkulasi darah melalui dinding usus halus. Hal ini menunjukkan bahwa sebagian besar peserta didik masih belum memahami konsep macam-macam enzim dalam pencernaan yang terdapat dalam mulut, pankreas dan usus halus. Bahkan mahasiswa tahun pertama jurusan pendidikan biologi di Turki juga belum memahami dengan baik tentang pencernaan kimiawi yang melibatkan enzim, seperti enzim ptialin yang terdapat di dalam mulut, serta enzim-enzim yang dihasilkan di dalam usus halus dan pankreas (Oztas, 2015: 121). Namun, sebagian besar mahasiswa tersebut memahami pencernaan mekanik yang terjadi ketika makanan dikunyah di dalam mulut dan gerakan peristaltik.

Analisis butir soal *try out* UN IPA SMP tahun 2019 yaitu No. 119 termasuk soal yang tidak baik karena tingkat kesukaran (sukar) dan daya beda (jelek) pada mapel kimia sesuai dengan Gambar 2.

119. Berikut ini yang merupakan campuran adalah

- santan, udara, dan bensin
- udara, bensin, dan susu
- bensin, susu, dan santan
- udara, santan, dan susu

Gambar 2. Contoh Butir Soal yang Tidak Baik No. 119

Berdasarkan Gambar 2, peserta *try out* banyak memilih opsi C, sedangkan kunci jawabannya D karena bensin termasuk dalam senyawa hidrokarbon. Hal ini menunjukkan bahwa sebagian besar peserta didik masih belum memahami perbedaan klasifikasi zat yang terdiri dari unsur, senyawa dan campuran. Temuan ini sesuai dengan hasil penelitian Wahyuni (2010: 52) yang menyimpulkan bahwa pemahaman siswa kelas IX SMP pada klasifikasi materi yang terdiri dari unsur, senyawa dan campuran diperoleh rata-rata sebesar 14,9% dan termasuk dalam kriteria sangat rendah. Lebih lanjut ia menambahkan bahwa kesulitan tersebut terutama terdapat pada aspek klasifikasi materi pada tingkat mikroskopik. Penelitian lain juga menemukan bahwa 52% dari siswa kelas IX SMP di Turki tidak memahami perbedaan antara unsur, senyawa dan campuran dalam klasifikasi materi sehingga mereka juga tidak dapat memahami metode pemisahannya, apakah secara fisika atau kimia (Tuysuz, 2009: 629).

Peserta didik menganggap bahwa udara bukan merupakan campuran, sedangkan konsep yang benar adalah udara terdiri dari beberapa gas, seperti gas oksigen (O_2), karbondioksida (CO_2), karbonmonoksida (CO), uap air (H_2O), dan sebagainya. Wahyuni (2010; 54) juga menemukan bahwa sebagian siswa kelas IX SMP

menganggap gas bukan materi karena gas tidak berwarna dan tidak terlihat (*gas is not matter because they are invisible*).

KESIMPULAN DAN SARAN

A. Kesimpulan

Berdasarkan hasil penelitian dan pembahasan pada butir soal *try out* UN IPA SMP tahun 2019 dapat disimpulkan sebagai berikut:

1. Penggunaan aplikasi ITEMAN dalam menganalisis butir soal diperoleh: a) nilai koefisien reliabilitas dengan formula ($\alpha = 0,84$ (sangat tinggi); *split-half (first-last)* = 0,72 (tinggi); dan S-B (*first-last*) = 0,84 (sangat tinggi)) serta kesalahan pengukuran sebesar 2,41; b) tingkat kesukaran: 3 butir (10,00%) mudah, 24 butir (80,00%) sedang, dan 3 butir (10,00%) sukar; c) daya beda: 3 butir (10,00%) jelek, 6 butir (20,00%) belum memuaskan, 5 butir (15,67%) memuaskan, dan 16 butir (53,33%) sangat memuaskan; dan efektivitas distraktor: 27 butir (90,00%) distraktor berfungsi mengecoh peserta *try out*, dan 3 butir (10,00%) distraktor tidak berfungsi.
2. Ada 19 butir (63,33%) mencakup 10 soal fisika, 7 soal biologi, 2 soal kimia yang baik berdasarkan teori tes klasik pada soal *try out* UN IPA SMP tahun 2019, sehingga dapat dipakai kemudian hari atau disimpan sebagai bank soal.

B. Saran

Guru dalam menyusun butir soal sebaiknya menguji coba terlebih dahulu, minimal diujikan kepada 30 responden. Dari analisis hasil uji coba tersebut, guru dapat mengklasifikasikan materi yang dianggap sulit oleh siswa. Guru sebaiknya merevisi soal yang belum termasuk dalam kriteria baik dan mengganti distraktor agar berfungsi sebelum diujikan pada skala luas. Selain itu, Dinas Pendidikan dan Olah Raga di tingkat Kabupaten hendaknya memberikan pelatihan penyusunan butir soal yang baik kepada para guru dengan mengundang ahli pengukuran dari BSNP.

UCAPAN TERIMA KASIH

Ucapan terima kasih disampaikan kepada Dikpora Bidang Pendidikan Dasar Kabupaten Tulungagung, MGMP IPA jenjang SMP/MTs Kabupaten Tulungagung, dan Primagama Tulungagung yang telah membantu pengumpulan data penelitian ini melalui dokumentasi Lembar Jawaban Siswa (LJS) peserta *Try Out* UN Tahun 2019 di Kabupaten Tulungagung – Jawa Timur.

DAFTAR RUJUKAN

- Azwar, S. (2013). *Tes Prestasi (Fungsi dan Pengembangan Pengukuran Prestasi Belajar) Edisi II*, Yogyakarta: Pustaka Pelajar Offset.
- BSNP. (2017). *Prosedur Operasional Standar (POS) Penyelenggaraan Ujian Nasional Tahun 2017/2018*. Jakarta: Badan Standar Nasional Pendidikan.
- Budiyanto, M. A. K. (2016). *Sintaks 45 Model Pembelajaran dalam Student Centered Learning (SCL)*. Malang: UMM Press.
- Huda, N. (2015). Analisis Karakteristik Perangkat Soal Ujian Akhir Semester Gasal Matematika Wajib Kelas X di SMA Negeri 9 Yogyakarta. *Prosiding Seminar Nasional Matematika dan Pendidikan Matematika di Universitas Muhammadiyah Purworejo*. Purworejo: UMP.
- Mardapi, D. (2008). *Teknik Penyusunan Instrumen Tes dan Nontes*. Yogyakarta: Mitra Cendikia Press.
- Mardapi, D. (2012). *Pengukuran Penilaian dan Evaluasi Pendidikan*. Yogyakarta: Nuha Medika.
- Oztas, H. (2015). First-Year University Students' Misconceptions about Digestive Tract and Its Function. *TIKM Journal of Education*, 1 (1): 119-123.
- Peraturan Menteri Pendidikan dan Kebudayaan Republik Indonesia. No. 4 Tahun 2018 *Tentang Penilaian Hasil Belajar oleh Satuan Pendidikan dan Penilaian Hasil Belajar oleh Pemerintah*. (Online), <https://bsnp-indonesia.org/wp-content/uploads/2018/02/Permen-4-Tahun-2018.pdf>, diakses pada tanggal 24 September 2019.
- Puspendik. (2017). *Panduan Penulisan Soal 2017 SMP/MTs*. Jakarta: Kemdikbud. (Online), <http://blog.uny.ac.id/suwarnadr/files/2018/05/Pedoman-Penulisan-Soal-SMP-MTs.pdf>, diakses pada tanggal 24 September 2019.
- Reynolds, C. R., Livingston, R. B., & Wilson, V. (2010). *Measurement and assessment in education*. Upper Saddle River: Pearson.

- Setyawarno, D. (2017). Penggunaan Aplikasi Iteman (Item and Test Analysis) Untuk Analisis Butir Soal Pilihan Ganda Berdasarkan Teori Tes Klasik. *JIFP (Jurnal Ilmu Fisika dan Pembelajarannya)*, 1(1): 11-21.
- Sukardi. (2003). *Metodologi Penelitian Pendidikan: Kompetensi dan Praktiknya*. Jakarta: Bumi Aksara.
- Sukardi. (2010). *Evaluasi Pendidikan: Prinsip & Operasionalnya*. Jakarta: Bumi Aksara.
- Surapranata, S. (2004). *Analisis, Validitas, Reliabilitas dan Interpretasi Hasil Tes: Implementasi Kurikulum 2004*. Bandung: Remaja Rosdakarya.
- Sutrisno, Nizaruddin, Yanuar, H.M., & Muhtarom. (2018). *ITEMAN: Item Klasik dan Analisis Uji*. Semarang: CEP-CCDA Universitas PGRI Semarang.
- Wahyuni, T.S. (2010). *Kajian Pemahaman Konsep Materi dan Perubahannya pada Siswa Kelas IX SMP Negeri 1 Tulungagung*. Skripsi, Universitas Negeri Malang.
- Tuysuz, C. (2009). Development of Two-Tier Diagnostic Instrument and Assess Students' Understanding in Chemistry. *Academic Journal*, 4 (6): 626-631.