
EVALUASI BUTIR SOAL ULANGAN AKHIR SEMESTER GASAL MATA PELAJARAN PENDIDIKAN JASMANI OLAHRAGA DAN KESEHATAN PADA KELAS KHUSUS BAKAT ISTIMEWA OLAHRAGA (BIO) DI SMA NEGERI 4 YOGYAKARTA

Pasca Tri Kaloka dan Sridadi

Universitas Negeri Yogyakarta, Jl. Kolombo No.1, Karangmalang Yogyakarta 55281

email: sridadi@uny.ac.id

Abstract

Educational success could be seen from the achievement of learning achieved by the learners that was obtained through evaluation. The research aimed to determine: (1) the level of difficulty, (2) different power, (3) and determine the effectiveness of the distractors or spread of answer pattern of final test of Odd Semester academic year 2013 / 2014 in SMA N 4 Yogyakarta in the Sport Talented Special Class (Bakat Istimewa Olahraga). This research was descriptive with ex post facto approach. The population was the students of class XI IPS 3 Sport Talented Special Class of 30 students and 2 teachers. The class sample was determined by documentation technique. It covered questions A and B with the answer key. The data were collected by the analysis of question point. Data analysis used ITEMAN. The research result showed that: (1) the level of difficulty of question package A and package B was in the difficult category. Package A difficulty level of the index was ($P < 0.30$) as many as 35 points (72%), while package B was ($P < 0.3$) of 41 points (82%) categorized as difficult. (2) Power difference of Package A and package B points was in the poor category. Different power index of Package A was ($D < 0$) as many as 36 items (42%), while package B was ($D < 0$) as many as 34 items (68%) in the poor category. (3) Effectiveness of distractor or answer pattern spread of Package A of 11 question items was good and 39 question items were not good, while on the package B for 6 question items were good and 44 question items were not good.

Keywords: question point evaluation, lesson of physical education, sports, and health

Abstrak

Keberhasilan pendidikan dapat dilihat dari prestasi belajar yang dicapai peserta didik yang diperoleh melalui kegiatan evaluasi. Penelitian ini untuk mengetahui: (1) tingkat kesukaran, (2) daya beda, (3) serta mengetahui efektifitas distraktor atau penyebaran pola jawaban ulangan akhir semester gasal tahun 2013/2014 di SMA N 4 Yogyakarta di Kelas Khusus Bakat Istimewa Olahraga (BIO). Penelitian ini merupakan penelitian diskriptif dengan pendekatan ex post facto. Populasinya peserta didik kelas XI IPS 3 Bakat Khusus Bakat Istimewa Olahraga sebanyak 30 orang dan 2 guru. Sampel kelas di tentukan dengan teknik dokumentasi. Meliputi paket soal A dan B serta kunci jawaban. Data dikumpulkan dengan analisis butir soal. Analisis data digunakan analisis ITEMAN. Hasil penelitian menunjukkan bahwa: (1) Taraf kesukaran butir soal paket A dan paket B masuk dalam kategori sukar. Indeks taraf kesukaran Paket A ($P < 0,30$) sebanyak 35 butir (72%), sedangkan paket B ($P < 0,3$) sebanyak 41 Butir (82%) masuk kategori sukar. (2) Daya beda butir soal Paket A dan paket B masuk dalam kategori jelek. Indeks daya beda Paket A ($D < 0$) sebanyak 36 butir (42 %), sedangkan paket B ($D < 0$) sebanyak 34 butir (68%) masuk kategori jelek. (3) Efektifitas distraktor atau penyebaran pola jawaban pada paket A sebanyak 11 butir soal yang baik dan 39 butir soal yang tidak baik, sedangkan pada paket B sebanyak 6 butir soal yang baik dan 44 butir soal yang tidak baik.

Kata kunci : evaluasi butir soal, pelajaran pendidikan jasmani olahraga dan kesehatan

PENDAHULUAN

Keberhasilan pendidikan di Indonesia menjadi tanggung jawab bersama antara sekolah, pemerintah dan masyarakat. Keberhasilan pendidikan itu sendiri dapat dilihat dari prestasi belajar yang dicapai peserta didik yang diperoleh melalui kegiatan evaluasi. Evaluasi akan memberikan informasi tingkat pencapaian belajar peserta didik, dan bila dianalisis lebih rinci akan diperoleh informasi tentang kesulitan belajar peserta didik, yaitu konsep-konsep yang belum dikuasai oleh sebagian besar peserta didik. Informasi ini yang harus digunakan pendidik untuk memperbaiki proses belajar mengajar yang nantinya dapat memperbaiki kualitas lulusan. Evaluasi memerlukan data yang akurat, yaitu data yang diperoleh melalui kegiatan pengukuran. Data yang akurat diperoleh apabila alat ukur yang digunakan sah dan handal. Syarat yang tidak terpenuhi dapat menimbulkan kesalahan pengukuran sehingga peserta didik tidak dapat diukur kompetensi yang sebenarnya. Kesalahan dalam evaluasi dapat juga menyebabkan penurunan kualitas pendidikan di Indonesia.

Pelaksanaan evaluasi sumatif di Sekolah Menengah Atas dilaksanakan dua kali yaitu pada akhir semester satu dan pada akhir semester dua. Tes sumatif pada akhir semester di SMA, yang pada saat ini disebut uji kompetensi. Ulangan akhir semester merupakan suatu bentuk *assessment* kepada peserta didik yang digunakan oleh institusi pendidikan di setiap jenjang pendidikan. Ulangan akhir semester mempunyai fungsi untuk memberi gambaran tentang tingkat penguasaan peserta didik terhadap materi pelajaran selama satu tahun pelajaran, dan sebagai laporan kepada orang tua peserta didik, serta dapat dijadikan bahan pengambilan keputusan untuk promosi kelas. Selain itu, ulangan akhir semester yang dilaksanakan secara bersama-sama oleh SMA-SMA di suatu daerah dapat juga digunakan untuk mengetahui sejauh mana keberhasilan proses pembelajaran di suatu sekolah apabila dibandingkan dengan sekolah yang lain. Melihat dari tujuan dan manfaat maka ulangan akhir semester perlu dilaksanakan secara rutin setiap tahunnya oleh institusi pendidikan pada setiap jenjang pendidikan.

Pelaksanaan ulangan akhir semester SMA di propinsi DIY berbeda di kota dan setiap kabupaten. Observasi awal di lapangan menunjukkan bahwa ulangan akhir semester gasal SMA Negeri 4 pada tahun ajaran 2013/2014 di kota Yogyakarta dilaksanakan oleh sekolah dengan pembuat soal adalah sekolah. Penyusun soal ulangan akhir semester gasal SMA Negeri 4 Yogyakarta pada tahun ajaran 2013/2014 adalah guru mata pelajaran di sekolah. Soal yang digunakan dalam ulangan akhir semester akan berbeda apabila penyusun soal adalah guru mata pelajaran di sekolah masing-masing sehingga daya serap, tingkat kesukaran, daya beda butir, dan fungsi distraktor akan berbeda. Ulangan akhir semester SMA pada hakikatnya merupakan evaluasi hasil belajar yang dilaksanakan oleh pendidik. Hal-hal yang berkaitan dengan evaluasi belajar, seperti: pengertian, fungsi dan tujuan, teknik dan alat evaluasi, jenis tes prestasi hasil belajar, analisis tes prestasi hasil belajar dan penyusunan soal ulangan, akan diuraikan sebagai berikut:

Soal ulangan akhir semester biasanya menggunakan pilihan ganda, *essay*, soal menjodohkan dan mengarang. Perlu hal itu dianalisis mengenai soal apa yang terbaik digunakan sehingga mampu untuk mengukur kemampuan peserta didik dalam waktu yang tidak lama dan menghemat dana dalam pelaksanaannya. Hasil suatu pengukuran atau ulangan harus memiliki kesalahan yang sekecil mungkin. Tingkat kesalahan ini berkaitan dengan kehandalan alat ukur. Kesalahan pengukuran ada yang bersifat acak dan ada yang sistematis. Kesalahan acak disebabkan oleh kondisi fisik dan mental peserta tes dan penyusun tes maupun pengawas yang bervariasi. Pengawas ulangan akhir semester biasanya masih belum terlalu memperhatikan hal ini. Para pengawas sering kali berjalan berkeliling untuk melihat hasil kerja peserta didik dalam mengawasi ulangan akhir semester, mengumumkan bahwa waktu ulangan sudah hampir habis dan sebagainya. Mereka tidak mengetahui bahwa apa yang mereka lakukan menambah kecemasan peserta tes dalam mengerjakan ulangan sehingga dapat menyebabkan hasil ulangan tidak menunjukkan kemampuan yang sebenarnya.

Kenyataan di lapangan juga menunjukkan bahwa sekolah maupun MKKS belum melakukan analisis terhadap butir-butir tes yang digunakan dalam ulangan akhir semester SMA Negeri 4 Yogyakarta khususnya kelas olahraga. Hal ini dapat menyebabkan informasi yang didapatkan dari hasil tes tidak akurat dan objektif. Saifuddin Azwar (2006: 2) mengatakan bahwa sifat suatu instrumen ukur yang tidak reliabel atau tidak valid akan memberikan informasi yang tidak akurat mengenai keadaan subjek atau individu yang dikenai tes itu. Informasi yang keliru apabila digunakan sebagai dasar pertimbangan dalam pengambilan suatu keputusan maka tidak akan menunjukkan keputusan yang tepat. Pernyataan di atas menunjukkan pentingnya validitas dan variabilitas suatu alat ukur tes, dari tes yang tidak valid dan tidak reliabel akan di dapatkan informasi yang salah sehingga apabila informasi itu secara tidak langsung digunakan sebagai dasar untuk pengambilan keputusan maka akan banyak pihak yang dirugikan. Validitas, jika dikaitkan dengan bidang psikologi, dapat dijumpai dalam tiga konteks yaitu validitas penelitian, validitas soal dan validitas alat ukur. Validitas penelitian merupakan derajat kesesuaian hasil penelitian dengan keadaan sebenarnya. Validitas soal berkaitan dengan kesesuaian antara suatu soal dengan soal lain. Sedangkan validitas alat ukur merujuk pada kecermatan ukurnya suatu tes (Sumadi Suryabrata, 2004: 40).

Ulangan akhir semester membutuhkan dana yang besar untuk pembiayaan pelaksanaannya. Dana digunakan untuk pembuatan butir soal oleh guru-guru terpilih, penggandaan paket soal selain itu sekolah harus mengeluarkan dana untuk membeli soal dari MKKS. Standar Penilaian Pendidikan sebagai proses pengumpulan dan pengolahan informasi untuk mengatur pencapaian hasil belajar peserta didik mencakup: Penilaian otentik, penilaian diri, penilaian berbasis portofolio, ulangan harian, ulangan tengah semester, ulangan akhir semester, ujian tingkat kompetensi, ujian mutu tingkat kompetensi ujian nasional dan ujian sekolah/madrasah.

Penggunaan soal-soal pilihan ganda dalam ulangan akan memberikan efisiensi dalam pelaksanaan ulangan karena dapat menghemat

waktu bagi peserta didik untuk mengerjakan soal, mencakup materi pelajaran yang luas dan pengkoreksian hasil jawaban dapat dilaksanakan dengan cepat. Pengawas juga perlu memperhatikan faktor kecermatan, yaitu dalam mengawasi ulangan sehingga kecurangan dapat dihindari dan hal-hal yang mengganggu peserta didik dalam mengerjakan soal dapat diminimalisir. Panitia ulangan juga perlu untuk memperhatikan masalah ketepatan waktu dalam beberapa hal, yaitu ketepatan waktu dalam pengerjaan soal oleh peserta didik, ketepatan waktu untuk datang bagi panitia dan peserta ulangan serta ketepatan waktu dalam memberikan nilai pada peserta didik sebelum rapat dibagikan.

Faktor ekonomi juga perlu diperhatikan, yaitu besarnya dana yang dikeluarkan untuk pelaksanaan ulangan benar-benar sudah digunakan sebagaimana mestinya. Hasil ulangan akhir semester harus dapat digunakan sebagai perbaikan kegiatan pembelajaran di sekolah sehingga aspek produktifitas dapat tercapai. Jadi hasil analisis akan menunjukkan komponen sistem ulangan akhir semester, mana yang belum berfungsi sebagaimana mestinya sehingga dapat dilakukan perbaikan pada sistem ulangan akhir semester tersebut. Djemari Mardapi (2004: 12) mengatakan bahwa usaha peningkatan kualitas pendidikan dapat ditempuh melalui peningkatan kualitas pengajaran dan kualitas evaluasinya. Oleh karena itu, setiap usaha memperbaiki kualitas pendidikan harus mencakup usaha untuk semakin menyempurnakan sistem evaluasi yang digunakan. Sistem evaluasi yang baik akan mendorong pendidik mengajar lebih baik dan peserta didik belajar lebih baik. Berangkat dari fungsi tes sebagai alat ukur, maka sebuah tes baru dapat dianggap berhasil menjalankan fungsinya jika ia mampu memberikan informasi yang sesuai dengan keadaan sebenarnya dari objek yang diukur. Tes yang tidak mampu memberikan informasi yang diinginkan tidak lebih dari sampah. Oleh karena itu, sebelum digunakan tes hasil belajar harus dianalisis terlebih dahulu sehingga dapat memberikan hasil yang memuaskan.

Berdasar pada pendapat tersebut dan fakta bahwa belum pernah dilakukannya analisis terhadap materi ujian tulis untuk materi pembelajaran penjasorkes di SMAN 4 Kota Yogyakarta terutama

pada kelas khusus olahraga, maka tulisan ini akan membahas tentang hal tersebut berdasar pada hasil penelitian yang telah dilakukan.

Adapun pembahasan khusus yang akan dituliskan dalam artikel ini adalah mencakup, (a) gambaran tingkat kesukaran soal ulangan akhir semester gasal SMA Negeri 4 Yogyakarta pada kelas olahraga tahun ajaran 2013/2014, (b) gambaran tingkat daya beda soal ulangan akhir semester gasaltahun 2013/2014 di SMA Negeri 4 Yogyakarta pada kelas bakat istimewa olahraga, dan (c) gambaran efektifitas distraktor atau penyebaran pola jawaban soal ulangan akhir semester gasal tahun 2013/2014 di SMA Negeri 4 Yogyakarta pada kelas bakat istimewa olahraga. Hasil dari pembahasan ini diharapkan guru memperoleh informasi mengenai tingkat kesukaran, daya beda butir dan penyebaran soal jawaban ulangan akhirsemester gasal kelas olahraga di SMA Negeri 4 Yogyakarta sehingga dapat digunakan untuk perbaikan dimasa yang akan datang. Peserta didik diharapkan dapat meningkatkan belajarnya setelah mengetahui tingkat kesukaran, tingkat daya beda butir dan efektifitas distraktor. Guru dapat memperbaiki kualitas soal setelah mengetahui tingkat kesukaran, daya beda dan distraktor dari soal yang dibuat. Mahasiswa dapat memperluas khasanah ilmu pengetahuan pada umumnya, serta dapat digunakan referensi bagi mahasiswa yang menekuni bidang pendidikan. Sumbangan empiris bagi SMA Negeri 4 Yogyakarta untuk melakukan perbaikan dalam pelaksanaan ulangan akhir semester gasal pada tahun-tahun ajaran yang akan datang.

TINJAUAN PUSTAKA

Hakikat Evaluasi

Evaluasi berasal dari kata *evaluation* (bahasa Inggris). Kata tersebut diserap ke dalam perbendaharaan istilah bahasa Indonesia dengan tujuan mempertahankan kata aslinya dengan sedikit penyesuaian lafal Indonesia menjadi "evaluasi". Suharsimi Arikunto (2013: 1), menjelaskan bahwa pengertian evaluasi adalah kegiatan untuk mengumpulkan informasi tentang bekerjanya sesuatu, yang selanjutnya informasi tersebut digunakan untuk menentukan alternatif yang tepat dalam mengambil sebuah keputusan. Evaluasi

menurut Arikunto (2013:3) mengemukakan bahwa evaluasi meliputi kedua langkah di atas, yakni mengukur dan menilai. Melalui evaluasi akan dapat di ketahui tentang hasil belakar mengajar yang dilakukan oleh guru dan siswa. Dengan demikian hasil belajar siswa dapat dipantau secara kontinyu, sehingga dapat diketahui materi mana yang telah dikuasai siswa dan materi mana yang belum dikuasai.

Evaluasi yang dilaksanakan oleh pendidik menurut Djemari Mardapi (2004: 12) dapat digolongkan menjadi dua, yaitu formatif dan sumatif. Informasi yang didapatkan dari penilaian formatif digunakan untuk menyesuaikan proses mengajar dan proses pembelajaran dengan kebutuhan peserta didik. Guru dapat menggunakan informasi dari penilaian formatif untuk mengambil tindakan yang dianggap perlu seperti *reteaching*, mencoba pendekatan alternatif terhadap peserta didik, atau menawarkan cara-cara lain untuk praktek apabila guru mengetahui bila peserta didik mendapatkan kesulitan. Evaluasi formatif bertujuan untuk memperbaiki cara atau strategi mengajar, sehingga hasilnya tidak digunakan untuk menentukan nilai peserta didik sedangkan evaluasi sumatif bertujuan untuk menentukan keberhasilan belajar peserta didik, sehingga hasilnya berupa nilai yang diperoleh peserta didik.

Validitas Tes

Penelitian ini membahas tentang karakteristik butir, oleh karena itu jenis validitas yang menjadi pembahasan utama adalah validitas tes. menurut Allen & Yen (1979: 95) validitas tes dapat dibagi kedalam tiga kelompok utama yaitu: a. validitas isi (*content validity*), b. validitas konstruk (*construct validity*) dan c. validitas kriteria (*criterion related validity*). Validitas isi menunjuk pada sejauhmana isi perangkat soal tersebut mengukur apa yang seharusnya diukur. Djemari Mardapi (2004: 26), dalam kaitannya dengan kegiatan pembelajaran, validasi isi suatu tes harus menjawab pertanyaan "sejauh mana *item test* itu mencakup keseluruhan situasi yang ingin diukur oleh tes tersebut."

Allen & Yen (1979: 95) membagi validitas isi kedalam dua kelompok yaitu *face validity* (validitas muka) dan *logical validity* (validitas logis). Validitas muka dapat dicapai jika tampilan tes tersebut

telah meyakinkan untuk mengungkap atribut yang hendak diukur. Adapun validitas logis menunjukkan sejauhmana isi tes mengungkapkan representasi dari ciri-ciri atribut yang hendak diukur. Validitas konstruk merujuk pada sejauhmana suatu tes mengukur suatu konstruk teoretik atau *trait* yang hendak diukur (Allen & Yen, 1979: 108) konstruk dalam pengertian ini adalah berkaitan dengan aspek-aspek psikologi seseorang khususnya aspek kognitif, afektif dan psikomotor.

Validitas kriteria merupakan validitas yang disusun berdasarkan kriteria yang telah ada sebelumnya. Kesahihan alat ukur, dalam validitas kriteria, dilihat dari sejauhmana hasil pengukuran tersebut sama dengan hasil pengukuran alat lain yang dijadikan kriteria. Validitas kriteria dibedakan menjadi dua macam yaitu berdasarkan kapan kriteria itu dapat dimanfaatkan. Jika dimanfaatkan dalam waktu dekat maka disebut validitas konkurent (*concurrent validity*) dan jika dimanfaatkan diwaktu akan datang disebut validitas prediktif (*predictive validity*) (Ebel & Frisbie, 1986: 95).

Analisis Butir Tes

Menurut Ngalim Purwanto (2010:118-120) Analisis soal tes ialah mencari soal tes mana yang baik dan mana yang tidak baik, dan mengapa soal itu dikatakan baik atau tidak baik. Dengan mengetahui soal-soal yang tidak baik itu selanjutnya kita dapat mencari kemungkinan sebab-sebab mengapa soal itu tidak baik. Dengan membuat analisis soal, sedikitnya kita dapat mengetahui tiga hal penting yang dapat di peroleh dari tiap soal, yaitu;

- a. Sampai di mana tingkat atau taraf kesukaran soal itu (*defficulty level of an item*)
- b. Apakah soal itu mempunyai daya beda (*discriminating power*) sehingga dapat membedakan kelompok peserta didik yang pandai dengan kelompok peserta didik yang bodoh
- c. Apakah semua alternatif jawaban (*options*) menarik jawaban-jawaban, ataukah ada yang demikian tidak menarik sehingga tidak perlu dimasukkan ke dalam soal.

Menurut Ngalim Purwanto (2010:119) untuk menghitung taraf kesukaran dan daya pembeda tiap soal dari suatu tes, kita perlu terlebih dahulu

mengelompokkan hasil tes tersebut menjadi tiga kelompok berdasarkan peringkat dari keseluruhan skor yang kita peroleh. Ketiga kelompok yang dimaksud ialah, (a) kelompok pandai atau *upper group* (25% dari peringkat bagian atas), (b) kelompok kurang atau *lower group* (25% dari peringkat bagian bawah), (c) kelompok sedang atau *middle group* (50% dari peringkat bagian tengah).

METODE PENELITIAN

Desain penelitian ini adalah *ex post facto* sering disebut *after the fact*. Dalam pengertian yang lebih khusus, (Sukardi, 2004:165) menguraikan bahwa penelitian *ex post facto* adalah penelitian yang dilakukan sesudah perbedaan-perbedaan dalam variabel bebas. Populasi pada penelitian ini adalah siswa dan guru olahraga di SMAN 4 Yogyakarta. Guru di SMA Negeri 4 yang menjadi sumber informasi sebanyak dua orang guru olahraga, dan staf SMA Negeri 4 yang meliputi guru BP dan wakil kepala sekolah urusan kesiswaan serta wakil kepala sekolah urusan kurikulum serta kepala sekolah SMA Negeri 4. Dua guru olahraga itu sekaligus sebagai penyusun soal ulangan. Teknik pengumpulan data yang digunakan untuk dalam penelitian ini yaitu: dokumentasi, penelusuran terhadap dokumen-dokumen penting yang berhubungan dengan pelaksanaan ulangan akhir semester gasal di SMA N 4 Yogyakarta dilakukan untuk lebih melengkapi data yang diperlukan untuk menjawab permasalahan penelitian, atau memperkuat data yang diperoleh melalui teknik yang lain. Dokumen yang penting untuk penelitian ini adalah paket soal dan kunci jawaban, dengan ini maka dapat dilakukan analisis butir soal untuk mengetahui kualitas soal ulangan akhir semester ini. Instrumen yang digunakan dalam penelitian ini merupakan instrumen yang dibuat sesuai dengan komponen pelaksanaan ulangan akhir semester gasal SMA Negeri 4 Yogyakarta tahun ajaran 2013/2014 di kota Yogyakarta pada tahap evaluasi yang diteliti. Instrumen penelitian adalah dokumentasi paket soal dan kunci jawaban kelas XI kelas khusus berbakat istimewa olahraga SMA Negeri 4 yogyakarta semester gasal. Validitas instrumen yang digunakan dalam penelitian ini adalah validitas isi (*content validity*). Validitas isi dalam penelitian

ini dilakukan dengan cara uji validitas berdasarkan penilaian/pertimbangan para ahli (*expert judgement*). Cara ini dilakukan dengan meminta pertimbangan kepada pembimbing untuk diperiksa dan dievaluasi secara sistematis, apakah butir-butir instrumen telah mewakili apa yang hendak diukur.

Data yang diperoleh dari penelitian ini dianalisis dengan menggunakan teknik analisis deskriptif untuk mengetahui bagaimana sistem ulangan akhirsemester gasal di SMA Negeri 4 Yogyakarta. Analisis perangkat tes pilihan ganda dilakukan dengan menggunakan bantuan program *Item and Test Analysis* (ITEMAN) versi 3.00. Hal ini dilakukan karena karakteristik serta kualitas secara empirik dari butir-soal dapat diketahui dengan menggunakan program ini. Hasil program ITEMAN adalah karakteristik butir dan perangkat tes secara statistik. Statistik butir tes tersebut adalah (1) tingkat kesukaran, (2) daya beda, (3) efektifitas distraktor. Secara empiris, kualitas butir tes yang diinginkan adalah butir yang memiliki tingkat kesukaran antara 0,3 sampai 0,7. Daya beda yang diinginkan adalah $\geq 0,3$ dikarenakan daya beda akan jelek jika indeksny di bawah 0,3. Adapun efektifitas distraktor berfungsi dengan baik jika dipilih oleh paling sedikit 2% untuk lima pilihan jawaban. Berikut adalah kriteria tingkat kesukaran butir :

Tabel 1. Tingkat Kesukaran Butir

Proportion Correct (p)	Kategori Soal
$p > 0,71$	Mudah
$0,31 \leq p \leq 0,70$	Sedang
$p < 0,30$	Sukar

Sumber : Suharsimi arikunto (2013:222)

Tabel pengkategorian daya beda butir soal disajikan sebagai berikut:

Tabel 2. Kategori Daya Beda Butir

Daya Beda	Kategori
0,00 – 0,20	Jelek
0,21 – 0,40	Cukup
0,41 – 0,70	Baik
0,71 – 1,00	Baik Sekali

Sumber : Suharsimi Arikunto (2013 : 232)

Adapun kriteria yang di tetapkan untuk menentukan baik tidaknya efektifitas distraktor atau penyebaran soal adalah sebagai berikut :

1. Baik apabila efektifitas distraktor atau penyebaran soal dipilih oleh 5% atau lebih oleh pengikut tes atau peserta didik.
2. Tidak Baik apabila efektifitas distraktor atau penyebaran soal dipilih kurang dari 5% oleh pengikut tes atau peserta didik.

HASIL DAN PEMBAHASAN PENELITIAN

Hasil

Pelaksanaan ulangan akhir semester gasal SMAN 4 Yogyakarta di Yogyakarta tahun ajaran 2013/2014, sepenuhnya dilaksanakan oleh pihak sekolah, tidak dilaksanakan oleh MKKS ataupun Dinas Pendidikan. Hal ini disikapi dengan baik oleh beberapa guru dengan asumsi bahwa yang mengetahui sejauh mana kompetensi yang telah diajarkan kepada peserta didik adalah guru di sekolah tersebut. Mereka berpendapat bahwa sangat dimungkinkan terdapat perbedaan antara kompetensi yang telah diajarkan antara satu SMA dengan SMA yang lain, sehingga soal yang disusun untuk ulangan akhir semester gasal tahun 2013/2014, berbeda satu dengan yang lainnya.

1. Taraf Kesukaran Butir Soal

a. Paket A

Berikut ini pemaparan hasil analisis butir soal dengan menggunakan Program ITEMAN (tm) version 3.00 :

Tabel 3. Tingkat Kesukaran Butir Soal Pendidikan Jasmani Olahraga dan Kesehatan KELAS XI IPS 3 Paket A SMA N 4 Yogyakarta tahun 2013/2014

Kategori	Jumlah (Butir)	Persentase	Nomor Butir
Mudah $p > 0,70$	0	0%	
Sedang $0,31 \leq p \leq 0,70$	9	18%	5, 20, 24, 31, 34, 38, 41, 43, 46
Sukar $p < 0,30$	41	82%	1, 2, 3, 4, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 21, 22, 23, 25, 26, 27, 28, 29, 30, 32, 33, 35, 36, 37, 39, 40, 41, 42, 44, 45, 47, 48, 49, 50
Total	50	100%	

b. Paket B

Berikut ini pemaparan hasil analisis butir soal dengan menggunakan Program ITEMAN (tm) version 3.00 :

Tabel 4. Tingkat Kesukaran Butir Soal Ulangan Akhir Semester, Pendidikan Jasmani Olahraga dan Kesehatan KELAS XI IPS 3 Paket B SMA N 4 Yogyakarta tahun 2013/2014

Kategori	Jumlah (Butir)	Persentase	Nomor Butir
Mudah P > 0,70	1	2%	29
Sedang 0,31 ≤ P ≤ 0,70	14	28%	20, 21, 22, 23, 24, 25, 26, 27, 31, 34, 38, 39, 42, 43
Sukar P < 0,30	35	70%	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 28, 30, 32, 33, 35, 36, 37, 40, 41, 44, 45, 46, 47, 48, 49, 50
Total	50	100%	

2. Daya Beda Butir Soal

a. Paket A

Berikut ini pemaparan hasil analisis butir soal dengan menggunakan Program ITEMAN (tm) version 3.00 :

Tabel 5. Kategori Daya Beda Butir Soal Pendidikan Jasmani Olahraga dan Kesehatan KELAS XI IPS 3 Paket A SMA N 4 Yogyakarta tahun 2013/2014

Kategori	Jumlah (Butir)	Persentase	Nomor Butir
Jelek D : 0,00 – 0,20	36	42%	1, 2, 3, 4, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 22, 24, 26, 27, 28, 32, 33, 35, 36, 37, 39, 40, 44, 45, 47, 48, 49, 50
Cukup D : 0,21 – 0,40	0	0%	-
Baik D : 0,41 - 70	1	2%	23
Baik Sekali D : 0,71 – 1,00	13	26%	5, 20, 21, 25, 29, 30, 31, 34, 38, 41, 42, 43, 46
Total	50	100%	

Kategori terbanyak pada butir soal mata pelajaran pendidikan jasmani olahraga dan kesehatan SMAN 4 Yogyakarta pada semester gasal tahun ajaran 2013/2014 paket A adalah pada kategori jelek, hal ini menunjukkan bahwa peserta tes yang masuk golongan atas banyak menjawab salah dan peserta tes golongan bawah banyak menjawab salah.

b. Paket B

Berikut ini pemaparan hasil analisis butir soal dengan menggunakan Program ITEMAN (tm) version 3.00 :

Tabel 6. Kategori Daya Beda Butir Soal Ulangan Akhir Semester Pendidikan Jasmani Olahraga dan Kesehatan KELAS XI IPS 3 Paket B SMA N 4 Yogyakarta tahun 2013/2014

Kategori	Jumlah (Butir)	Persentase	Nomor Butir
Jelek D : 0,00 – 0,20	34	68%	2, 3, 4, 6, 7, 8, 9, 10, 11, 12, 14, 15, 16, 17, 18, 19, 20, 24, 26, 27, 28, 29, 30, 32, 33, 34, 36, 37, 39, 40, 44, 45, 46, 47
Buruk D : 0,21-0,40	3	6%	21, 25, 31
Tidak Baik D : 0,41 – 0,70	6	12%	1, 38, 41, 42, 43, 48
Baik D : 0,71 – 1,00	7	14%	5, 13, 22, 23, 35, 49, 50
Total	50	100%	

Kategori terbanyak pada butir soal mata pelajaran pendidikan jasmani olahraga dan kesehatan SMA N 4 Yogyakarta Paket B adalah pada kategori jelek, hal ini menunjukkan bahwa peserta tes yang masuk golongan atas banyak menjawab salah dan peserta tes golongan bawah banyak juga menjawab salah.

3. Efektifitas Distraktor atau penyebaran pola jawaban

a. Paket A

Karakteristik butir yang baik juga ditentukan oleh efektifitas distraktor (pengecoh). Penelaahan efektifitas distraktor dilakukan untuk mengetahui apakah distraktor-distraktor sudah berfungsi sebagaimana mestinya, dalam arti apakah pengecoh tersebut telah dipilih oleh peserta dari kelompok berkemampuan rendah dan tidak dipilih atau hanya sedikit dari kelompok berkemampuan tinggi. Sebuah distraktor dikatakan baik jika direspon oleh minimal 5% peserta dan indeks poin biserialnya negatif selain kunci jawaban.

Hasil Iteman menunjukkan bahwa butir-butir soal pendidikan jasmani olahraga dan kesehatan pada ulangan akhir semester

gasal tahun ajaran 2013/2014 SMA Negeri 4 Yogyakarta Kelas XI IPS 3 Paket A yang memiliki distraktor yang baik berjumlah 11 butir soal yaitu butir soal nomor 5,20,21,29,30,31,34,38,41,43,46. Sisanya adalah butir soal yang setidak-tidaknya memiliki satu distraktor yang tidak berfungsi dengan baik

b. Paket B

Hasil Iteman dapat dilihat bahwa butir-butir soal pendidikan jasmani olahraga dan kesehatan SMA Negeri 4 XI IPS 3 Paket B yang memiliki distraktor yang baik berjumlah 6 butir soal yaitu butir soal nomor 1,5,13,22,23,35 . Sisanya adalah butir soal yang setidak-tidaknya memiliki satu distraktor yang tidak berfungsi dengan baik.

4. Karakteristik Butir Soal Ulangan akhir Semester Gasal SMA Negeri 4 Yogyakarta

Untuk mengetahui kualitas terbaik butir soal di suatu kelas dan kualitas terbaik perangkat soal yang sama antar kelas, disajikan perbandingan sebagai berikut:

Tabel 7. Perbandingan Hasil Analisis Butir Soal Antar Paket A dengan Paket B Kelas XI IPS 3 SMA Negeri 4

Mapel	Persentase Kategori Tingkat Kesukaran			Persentase Kategori Daya Beda				Distraktor atau pola jawaban soal	
	Mudah	Sedang	Sukar	Jelek	Cukup	Baik	Baik sekali	Baik	Tidak Baik
Paket A	0 %	18 %	82 %	42 %	2 %	0 %	26 %	11	39
Paket B	2 %	28 %	70 %	68 %	6 %	12 %	14 %	6	44

Perbandingan mata pelajaran Pendidikan Jasmani Olahraga dan Kesehatan Kelas XI IPS 3 paket A dengan paket B menunjukkan bahwa pada parameter tingkat kesukaran butir soal yang sedang pada soal paket B lebih banyak daripada butir soal Paket A (28% : 18%). Sedangkan pada parameter daya beda, persentase butir soal paket A yang mempunyai daya beda baik lebih banyak daripada butir soal paket B (0% : 12%). Paket soal ulangan semester di kelas XI IPS 3 SMA Negeri 4 memiliki distraktor atau penyebaran soal paket A mempunyai 11 soal yang baik, sedangkan pada paket B mempunyai 6 soal yang baik.

Pembahasan Penelitian

1. Karakteristik Butir Soal Ulangan Akhir Semester Gasal SMA N 4 Yogyakarta Tahun Ajaran 2013/2014

a. Taraf kesukaran pada Soal Pendidikan Jasmani Olahraga dan Kesehatan Kelas XI IPS 3 Paket A Kelas bakat olahraga SMA Negeri 4 menunjukkan bahwa 82% butir soal berada pada kategori soal yang sukar, kemudian 18 % butir soal berada pada kategori sedang. Paket B Kelas bakat istimewa olahraga SMA Negeri 4 Yogyakarta, dimana sebagian besar (70%) butir soal tingkat kesukarannya berada pada kategori sukar. Menurut Ngalim Purwanto (2013 : 130) soal ini termasuk soal yang sukar, karena jumlah butir soal yang sukar mencapai 41 soal. Dalam hal ini ada dua keterangan yang mungkin untuk data analisis soal tersebut : (1) soal tersebut bersifat *ambiguous* (mempunyai dua arti), terutama bagi siswa-siswa yang mengetahui banyak hal tentang hal yang ditanyakan, atau (2) para siswa belum pernah mempelajari hal-hal seperti yang ditanyakan.

b. Daya beda pada Soal Pendidikan Jasmani Olahraga dan Kesehatan Kelas XI IPS 3 Paket A Kelas bakat olahraga SMA Negeri 4 menunjukkan sebagian besar 42% butir soal, daya bedanya berada pada kategori jelek, kemudian dari lima puluh butir soal hanya ada tiga belas soal yang berada pada kategori butir yang baik sekali, sedangkan belum ada soal yang berada pada kategori butir yang cukup. Paket B Kelas bakat istimewa olahraga SMA Negeri 4 Yogyakarta, sebagian besar (68%) butir soal berada pada kategori yang jelek, kemudian dari lima puluh butir soal hanya ada 7 yang yang berada pada kategori baik sekali, sedangkan kategori baik ada 6 soal dan kategori baik ada 3 soal. Menurut Suharsimi arikunto (2013:233) Nilai-nilai dengan klasifikasi daya pembeda dianjurkan oleh para penulis soal adalah antara 0-30 dan 0,70 dengan kata lain antara kategori cukup – baik. Oleh karena itu, soal pada paket A dan paket B termasuk mempunyai daya beda yang jelek.

c. Distraktor atau pola soal jawaban pada Soal Pendidikan Jasmani Olahraga dan Kesehatan Kelas XI IPS 3 Paket A Kelas bakat olahraga SMA Negeri 4 menunjukkan bahwa distraktor atau penyebaran soal yaitu 11 soal yang baik, sedangkan pada paket B mempunyai 6 soal yang baik, sisanya adalah soal yang tidak baik. Menurut Suharsimi Arikunto (2013:234) Suatu distraktor di katakan berfungsi baik jika paling sedikit dipilih oleh 5% pengikut tes, oleh karena itu soal yang terdapat pada paket A hanya memiliki 11 soal yang baik dan pada paket B hanya 6 soal baik.

KESIMPULAN DAN SARAN

Kesimpulan

Kesimpulan yang didapatkan dari penelitian ini adalah sebagai berikut:

1. Tingkat kesukaran butir soal pada soal paket B lebih sulit daripada tingkat kesukaran butir soal paket A. Jadi tingkat kesukaran butir soal mata pelajaran Pendidikan Jasmani Olahraga dan Kesehatan Kelas Bakat Khusus Istimewa di SMA N 4 pada ulangan akhir semester gasal adalah sukar. Karena $P < 0,30$ atau indeks kesukaran $< 0,30$ termasuk kategori sukar, sedang tingkat kesukaran butir soal mata pelajaran Pendidikan Jasmani Olahraga dan Kesehatan kelas IPS 3 Paket B presentase kesukarannya adalah 82 % atau jumlah soal yang sukar mencapai 41 butir soal sedangkan pada paket soal A tingkat kesukarannya juga termasuk sukar, karena jumlah soal yang sukar mencapai 35 butir atau 70 %. Hingga keduanya termasuk kategori sukar.
2. Soal paket A mata pelajaran Pendidikan Jasmani dan Olahraga dan Kesehatan pada ulangan akhir semester gasal tahun 2013/2014 mempunyai daya beda yang termasuk kategori jelek karena $D < 0$ atau indeks daya beda < 0 , dan jumlah butir soal yang sangat buruk mencapai 36 butir atau 42 % . Sedang daya beda pada soal paket B juga termasuk sangat buruk karena $D < 0$, dan jumlah butir soal yang sangat buruk 34 butir soal dan prosentasenya mencapai 68%. Sehingga ulangan akhir semester gasal pada

mata pelajaran Pendidikan Jasmani dan Olahraga dan Kesehatan di SMA N 4 tahun 2013/2014 termasuk jelek .

3. Efektifitas distraktor atau penyebaran pola jawaban akhir semester gasal pada tahun 2013/2014 di SMA 4 Yogyakarta pada kelas bakat istimewa olahraga Paket A memiliki 11 butir soal yang baik, sedangkan pada Paket B terdapat 6 butir soal yang baik dari semua soal yang berjumlah 50 butir soal.

Saran

1. Tingkat kesukaran soal ulangan akhir semester gasal tahun 2013/2014 di SMAN 4 Yogyakarta Kelas bakat istimewa mata pelajaran pendidikan jasmani olahraga dan kesehatan termasuk kategori sukar, maka guru harus berusaha mencapai kriteria ketuntasan minimal (KKM) yang menurut kurikulum yang berlaku adalah 75 hingga peserta didik harus dapat mengerjakan soal sejumlah minimal 75% dari seluruh butir soal.
2. Karena daya beda soal ulangan akhir semester gasal yang jelek itu, maka guru harus berusaha agar daya beda peserta didik tidak terlalu jauh antara peserta didik yang pandai dan peserta didik yang kurang
3. Efektifitas distraktor atau pola jawaban pada paket A baru ada 11 butir soal yang baik, sedang paket B baru ada 6 soal yang baik dari jumlah 50 butir soal. Sehingga seharusnya perlu ada soal yang direvisi yaitu soal yang belum baik dalam paket A ada 39 soal dan pada paket B yang perlu direvisi ada 44 butir soal
4. Guru agar menganalisis butir ulangan akhir semester, serta dapat menyimpan butir-butir yang baik ke dalam bank soal. Guru akan memiliki bank soal yang baik dan variatif dapat digunakan untuk menyusun paket soal ulangan pada kesempatan ulangan yang lain.

DAFTAR PUSTAKA

- Allen, M. J. & Yen, W. M. (1979). *Introduction To Measurement Theory*. Monterey, California: Brookd/Cole Publishing Company.
- Djemari Mardapi. (2004). Pengembangan Sistem Penilaian Berbasis Kompetensi. Makalah. Disampaikan pada Seminar Himpunan Evaluasi

- Pendidikan Indonesia (HEPI) "Rekayasa Sistem Penilaian untuk Meningkatkan Kualitas Pendidikan". Yogyakarta.
- Ebel, R.L (1972). *Essentials of Educational Measurement*. (3rd, ed) Englewood Cliffts, NJ: Prentice Hall Inc.
- Saifuddin Azwar. (2006). *Reliabilitas dan Validitas*. Yogyakarta: Pustaka Pelajar Offset.
- Suharsimi Arikunto. (2013). *Dasar-dasar Evaluasi Pendidikan*. Jakarta: Bumi Aksara.
- Ngalim Purwanto. (2010). *Prinsip-prinsip dan teknik Evaluasi Pengajaran*. Bandung: PT Remaja Rosdakarya.
- Sumadi Suryabrata. (2004). *Pengembangan Alat Ukur Psikologi*. Yogyakarta: Andi.
- Sukardi. (2004). *Metodologi Penelitian Pendidikan*. Jakarta: Bumi Aksara.