

---

## Analisis butir tes hasil belajar mata pelajaran PJOK kelas VIII

Sridadi<sup>1\*</sup>, Riky Dwihandaka<sup>1</sup>, Ariyo Bagiastomo<sup>1</sup>

<sup>1</sup>Jurusan Pendidikan Olahraga, Fakultas Ilmu Keolahragaan, Universitas Negeri Yogyakarta, Jl. Colombo No. 1 Yogyakarta

\*Corresponding Author. Email: [sridadi@uny.ac.id](mailto:sridadi@uny.ac.id) , [rikydwihandaka@uny.ac.id](mailto:rikydwihandaka@uny.ac.id), [ariyo.bagiastomo@yahoo.com](mailto:ariyo.bagiastomo@yahoo.com)

---

### Abstrak

Penelitian ini bertujuan untuk memberikan informasi mengenai seberapa baik kualitas butir soal ulangan akhir Mata Pelajaran PJOK Kelas VIII SMP N 1 Ngemplak. Penelitian ini merupakan penelitian deskriptif kuantitatif dengan teknik pengumpulan data menggunakan teknik dokumentasi. Penelitian ini dilakukan dengan cara mengumpulkan data berupa soal ujian, kunci jawaban, dan lembar jawab peserta didik. Selanjutnya data tersebut diinput dan diolah menggunakan program Anates versi 4.09. Subjek penelitian adalah seluruh peserta didik kelas VIII berjumlah 187 peserta didik. Hasil analisis kualitas butir soal menunjukkan bahwa soal ulangan akhir Mata Pelajaran PJOK Kelas VIII SMP N 1 Ngemplak dari total 40 butir soal pilihan ganda, ditinjau dari segi tingkat kesukaran menunjukkan 5 butir soal (12,5%) berkategori sukar, 15 butir soal (37,5%) berkategori sedang, dan 20 butir soal (50%) berkategori mudah. Ditinjau dari daya pembeda, didapatkan 19 butir soal (47,5%) berkategori jelek, 13 butir soal (32,5%) berkategori sedang, dan 8 butir soal (20%) berkategori baik. Ditinjau dari efektivitas pengecoh, didapat sebanyak 5 butir soal (12,5%) berkategori sangat baik, 11 butir soal (27,5%) berkategori baik, 15 butir soal (37,5%) berkategori kurang baik, dan 9 butir soal (22,5%) berkategori tidak baik.

**Kata Kunci:** analisis butir tes, hasil belajar, PJOK

### *The analysis of test item on learning outcomes of physical education subject of class VIII students*

#### Abstract

*This research was aimed to give information about how well the quality of questions in final exam of physical education subject of Class VIII SMP N 1 Ngemplak. This research was a descriptive quantitative research using documentation as data collection techniques. This research was conducted by collecting data in the form of exam questions, key answers, and the students' answer sheets. Next, the data were inputted and processed using the Anates program version 4.09. The research subjects were all students of class VIII totaling 187 students. The results of the quality analysis of the questions showed that the questions in final exam of physical education subject of Class VIII SMP N 1 Ngemplak out of 40 multiple choice questions, viewed from the level of difficulty, 5 items (12.5%) were in difficult category, 15 items (37.5%) were in average category, and 20 items (50%) were in easy category. Viewed from the distinguishing strength, 19 items (47.5%) were categorized as bad, 13 items (32.5%) were categorized average, and 8 items (20%) were categorized good. Viewed from the effectiveness of the deception, it was obtained 5 items (12.5%) were in very good category, 11 items (27.5%) were in good category, 15 items (37.5%) were in poor categories, and 9 items (22.5%) were in not good category.*

**Keywords:** analysis of test item, learning outcomes, physical education

#### PENDAHULUAN

Pendidikan atau *education* merupakan suatu proses transfer falsafah (*philosophy*), sistem nilai (*values*), pengetahuan (*knowledge*) dan kemampuan (*ability*) sebagai bagian

paling penting untuk investasi sumber daya manusia (*investment in human capital*) (Sriundy, 2014, p 1). Pendidikan dijadikan sebagai fondasi utama pembangunan, oleh karena itu pendidikan harus berkualitas jika negara tersebut ingin maju. Pendidikan pada masyarakat yang maju merupakan suatu proses berkelanjutan yang diselenggarakan secara formal, informal, dan nonformal. Pendidikan formal diselenggarakan di sekolah dan dimulai dari pendidikan paling dasar sampai jenjang pendidikan yang paling tinggi. Kegiatan pendidikan di sekolah sering disebut dengan pengajaran, secara umum prosesnya disebut *teaching and learning* atau proses belajar dan mengajar (PBM). Menurut Sriundy (2014, p. 2) secara teoritis kegiatan PBM melibatkan lima komponen penting, yaitu: 1) Tujuan pengajaran, 2) Pendidik atau pengajar, 3) Peserta didik, 4) Bahan ajar atau materi ajar, dan 5) Lingkungan belajar.

Pendidik harus memiliki kecakapan dalam merencanakan sistem pengajaran pada lingkup yang paling kecil (*micro system*). Sriundy (2014, p. 3) menyatakan bahwa kecakapan tersebut meliputi: 1) Kecakapan dalam menetapkan tujuan pengajaran, 2) kecakapan memilih dan menetapkan serta menyiapkan bahan ajar, 3) Kecakapan mendesai perencanaan pengajaran, 4) Kecakapan melaksanakan pengajaran, dan 5) Kecakapan menilai dan membina sistem pengajaran. Masing-masing kecakapan tersebut merupakan komponen yang tidak dapat dipisahkan. Khusus kecakapan poin kelima merupakan kecakapan yang dilaksanakan melalui kegiatan evaluasi terhadap empat komponen penting pendidikan. Dalam mata pelajaran Pendidikan Jasmani Olahraga dan Kesehatan (PJOK) pun demikian. Terdapat berbagai ranah yang perlu dilakukan dalam penilaian dan evaluasi domain kognitif, domain afektif, dan domain psikomotor, dan domain fisik. Keempat domain tersebut merupakan bagian dari tujuan pendidikan jasmani.

Evaluasi mengacu pada proses sistematis untuk menentukan tujuan pembelajaran yang dicapai oleh siswa (Gronlund, 1982, p. 5-6). Pemenuhan tujuan pembelajaran dilakukan dengan pengukuran. Dengan mengukur dan mengevaluasi, guru dapat mendiagnosis kekuatan dan kelemahan siswa kemudian mengambil tindakan untuk kemajuan dan peningkatan siswa. Jika efektif, pengukuran dan evaluasi dapat meningkatkan situasi pembelajaran. Tanpa evaluasi dan pengukuran, tidak mungkin untuk mengetahui kebutuhan dan kemampuan siswa (Tshabalala & Ncube, 2014, p. 141). Karena itulah penilaian dan evaluasi sangat diperlukan untuk mengetahui apakah pembelajaran sudah mencapai tujuan pendidikan. Edwind Wandt dan Gerald W. Brown (dalam Sudijono, 2009, p. 1) mengemukakan bahwa, *Evaluation refers to the act or process to determining the value of something*. Menurut definisi ini, istilah evaluasi itu merujuk kepada atau mengandung pengertian: suatu tindakan atau suatu proses untuk menentukan nilai dari sesuatu.

Sudijono (2009, p. 1) memaparkan bahwa apabila definisi yang dikemukakan oleh Edwin Wandt dan Gerald W. Brown itu untuk memberikan definisi tentang evaluasi pendidikan, evaluasi pendidikan dapat diberi pengertian sebagai kegiatan atau proses penentuan nilai pendidikan sehingga dapat diketahui mutu atau hasil-hasilnya. Dalam pembelajaran PJOK di sekolah, semua ranah perlu diperhatikan dan mendapat porsi yang sama besar. Akan tetapi, pada kenyataannya persentase ranah psikomotor selalu ditonjolkan dalam evaluasi pembelajaran PJOK. Padahal penilaian dalam pembelajaran pendidikan jasmani meliputi empat ranah. Ranah kognitif juga perlu memerlukan perhatian khusus dari guru karena ranah tersebut dapat mengukur perkembangan kemampuan pemahaman peserta didik dalam memahami suatu materi pembelajaran.

Penilaian ranah kognitif penting dilakukan untuk mengetahui kemampuan peserta didik dari segi pengetahuan. Guru dapat mengadakan penilaian ranah kognitif untuk mengetahui sejauh mana proses pembelajaran berhasil dan sejauh mana pemahaman peserta didik terhadap materi yang telah diajarkan oleh guru selama kurun waktu tertentu. Alat evaluasi dalam pendidikan yang digunakan untuk mengumpulkan data dapat berupa tes atau nontes (Purwanto, 2011, p. 56). Alat evaluasi hasil belajar, teknik tes merupakan teknik yang paling sering digunakan oleh guru untuk mengadakan penilaian ranah kognitif sehingga dapat diketahui sejauh mana peserta didik menguasai dan memahami materi. Dengan tes inilah, guru biasanya melakukan evaluasi. Tes hasil belajar digunakan sebagai pertimbangan guru

untuk menentukan apakah siswa naik kelas berikutnya atau tidak (Kusumawati & Hadi, 2018, p. 1). Selain itu, hasil tes ini digunakan sebagai evaluasi bagi pendidik dalam pelaksanaan proses pembelajaran. Item tes akhir yang baik akan memberikan hasil pengukuran yang baik (Mardapi, 2012, p. 27).

Sudijono (2011, p. 67-73) mengemukakan bahwa tes merupakan cara atau prosedur yang perlu ditempuh dalam rangka pengukuran dan penilaian di bidang pendidikan, yang berbentuk pemberian tugas atau serangkaian tugas (baik berupa pertanyaan yang harus dijawab, atau perintah yang harus dikerjakan oleh testee) sehingga dapat dihasilkan nilai yang melambangkan tingkah laku atau prestasi testee; nilai mana dapat dibandingkan dengan nilai-nilai yang dicapai oleh testee lainnya, atau dibandingkan dengan nilai standar tertentu. Evaluasi yang dilakukan dapat berjalan secara maksimal dan mengenai target apabila pengukuran yang dilakukan sesuai dengan kompetensi yang akan diukur. Oleh karena itu, guru harus mengetahui langkah-langkah penyusunan tes dalam ranah kognitif sesuai dengan materi yang telah diberikan kepada peserta didik. Langkah pertama dalam penyusunan tes kognitif yaitu menyusun kisi-kisi materi yang akan diujikan. Dalam hal ini, guru harus membuat ringkasan garis besar mengenai materi yang telah diajarkan kepada peserta didik. Hal ini dilakukan agar instrumen tes yang diberikan dapat mencakup materi yang penting dari pelajaran yang telah dilakukan sebelumnya sehingga benar-benar dapat mengukur apa yang seharusnya diukur. Dengan kisi-kisi pula, peserta didik akan benar-benar siap dengan tes yang akan dilakukan karena mereka dapat mempersempit wilayah belajar sehingga kemampuan mereka dapat benar-benar terukur.

Guru harus bisa memilih tipe tes yang tepat dan membuat soal tes berdasarkan kisi-kisi tersebut. Dalam ranah kognitif, guru dapat memilih item tes pilihan berganda. Dalam tes pilihan berganda, seorang guru harus bisa membuat soal yang baik. Soal tersebut tidak boleh asal dibuat, melainkan harus benar-benar dapat memberikan hasil yang mencerminkan kemampuan peserta didik yang sebenarnya. Tes hasil belajar kognitif umumnya terdiri atas item-item yang berkaitan dengan kecakapan siswa, yaitu sebagai prediksi tingkat pencapaian siswa terhadap tujuan-tujuan belajar yang sudah ditetapkan. Oleh karena itu kualitas tes sangat tergantung pada kualitas item-item yang menyusunnya. Item tes yang berkualitas tinggi lebih bermanfaat walaupun jumlah itemnya sedikit, daripada item yang banyak tetapi berkualitas rendah. Menurut Azwar dalam Sriundy (2014, p.160) item tes kualitas rendah akan menurunkan fungsi ukur tes, bahkan akan memberikan gambaran hasil yang menyesatkan. Tes yang sudah disusun berdasarkan teori dan cara kerja standar tertentu, belum tentu secara empirik mampu berfungsi dengan baik di lapangan. Oleh karena itu instrumen tes harus tetap diuji melalui prosedur ujicoba. Hasil ujicoba dapat digunakan untuk menilai kualitas tes dan sebagai pertimbangan untuk melakukan perbaikan.

Analisis kualitas butir soal adalah suatu prosedur yang sistematis, yang akan memberikan informasi-informasi yang sangat khusus terhadap butir tes yang kita susun (Arikunto, 2013, p. 220). Surapranata (2009, p. 1) menyatakan analisis kualitas soal dilakukan untuk mengetahui berfungsi tidaknya sebuah soal, analisis kualitas butir soal merupakan kegiatan menganalisis tiap-tiap butir soal secara mendetail menggunakan metode pengujian tertentu. Analisis kualitas butir soal adalah kegiatan yang dilakukan untuk mengidentifikasi soal-soal baik, kurang baik dan soal jelek dan memperoleh petunjuk untuk melakukan perbaikan (Daryanto, 2008, p. 179). Analisis butir soal adalah suatu usaha untuk mengetahui mana butir soal yang baik, layak, dan dapat digunakan maupun butir soal yang kurang baik dan harus direvisi agar dapat menjadi soal yang benar-benar berkualitas baik. Supardi (2016, p. 82) menyatakan bahwa bahan-bahan penunjang analisis butir soal secara kualitatif adalah kisi-kisi pembuatan soal, kurikulum acuan yang digunakan, buku sumber, kamus bahasa Indonesia, dan pedoman analisis kualitas butir soal objektif maupun subjektif.

Tujuan analisis butir soal juga untuk membantu meningkatkan tes melalui revisi atau membuang soal yang tidak efektif, serta untuk mengetahui informasi diagnostik pada peserta didik apakah mereka sudah/belum memahami materi yang telah diajarkan (Aiken, 1994, p. 63). Hasil analisis akan menunjukkan seberapa baik kualitas suatu soal. Soal yang berkualitas dapat diketahui dari tingkat kesukaran soal, daya pembeda, dan efektivitas

pengecoh soal. Lewat hasil analisis, guru dapat memberikan evaluasi terhadap soal tersebut. Analisis butir soal dilaksanakan sebagai salah satu tindak lanjut perbaikan kualitas suatu tes. Bila kualitas butir soal baik, evaluasi yang dilakukan akan berjalan maksimal. Evaluasi yang berjalan maksimal akan membuat proses belajar mengajar semakin baik dan meningkatkan kemampuan peserta didik dari tahun ke tahun.

Anastasi dan Urbin dalam Kusaeri dan Suprananto (2012, p. 164) menyatakan bahwa kegiatan analisis butir soal memiliki banyak manfaat, di antaranya: 1) dapat membantu pengguna tes dalam mengevaluasi kualitas tes yang digunakan; 2) relevan bagi penyusunan tes informal seperti tes yang disiapkan guru untuk siswa di kelas; 3) mendukung penulisan butir soal yang efektif; 4) secara materi dapat memperbaiki tes di kelas; 5) meningkatkan validitas dan reliabilitas soal. Kusaeri dan Suprananto (2012, p.164) menyatakan beberapa manfaat lain yang diperoleh pendidik dalam melakukan kegiatan analisis kualitas butir soal antara lain:

1. Menentukan soal-soal yang cacat atau tidak berfungsi dengan baik,
2. Meningkatkan butir soal melalui tiga komponen analisis, yaitu tingkat kesukaran, daya pembeda, dan pengecoh soal,
3. Merevisi soal yang tidak relevan dengan materi yang diajarkan, ditandai dengan banyaknya anak yang tidak dapat menjawab butir soal tertentu.

### **Tingkat Kesukaran**

Tingkat kesukaran sebuah item tes dinyatakan dengan besaran indeks yang biasa disebut dengan indeks kesukaran item. Tingkat kesukaran disimbolkan dengan huruf  $p$ , yang merupakan rasio antara jumlah peserta tes dengan banyaknya peserta tes yang berhasil menjawab item tersebut dengan benar (Sriundy, 2014, p. 161). Oleh karena itu tingkat kesukaran item tes sesungguhnya merupakan proporsi jawaban benar yang didapat dari hasil bagi jawaban benar dengan seluruh item dan dikalikan 100%. Bermutu atau tidaknya butir-butir item tes hasil belajar pertama-tama dapat diketahui dari derajat kesukaran atau taraf kesulitan yang dimiliki oleh masing-masing butir item tersebut. Butir-butir item tes hasil belajar dapat dinyatakan sebagai butir-butir item yang baik, apabila butir-butir item tersebut tidak terlalu sukar dan tidak pula terlalu mudah dengan kata lain derajat kesukaran itu adalah sedang atau cukup (Sudijono, 2011, p. 370).

Arikunto (2013, p. 210) memaparkan bahwa kriteria yang digunakan sebagai tolak ukur tingkat kesukaran suatu soal adalah semakin kecil indeks yang diperoleh, makin sulit soal tersebut sebaliknya, semakin besar indeks yang diperoleh, semakin mudah soal tersebut. Butir soal yang baik memiliki indeks kesukaran antara 30%-70%. Butir soal yang memiliki tingkat kesukaran di bawah 30% dikatakan sukar. Butir soal yang memiliki tingkat kesukaran 30%-70% dikatakan sedang, sedangkan tingkat kesukaran di atas 70% dikatakan terlalu mudah.

Sudaryono (2012, p. 117) menyatakan tingkat kesukaran butir soal memiliki 2 kegunaan, yaitu: 1) Kegunaan bagi pendidik, meliputi: a) Sebagai pengenalan konsep terhadap pembelajaran ulang dan memberi masukan kepada peserta didik tentang hasil belajar mereka, b) Memperoleh informasi tentang penekanan kurikulum atau mencurigai butir soal yang bias. 2) Kegunaan bagi pengujian dan pengajaran, meliputi: a) Pengenalan konsep yang diperlukan untuk diajarkan ulang, b) Tanda-tanda terhadap kelebihan dan kelemahan pada kurikulum sekolah, c) Memberi masukan kepada peserta didik, d) Tanda-tanda kemungkinan adanya butir soal yang bias, e) Merakit tes yang memiliki ketepatan daya soal.

### **Daya Beda**

Menurut Anastasi dan Urbina dalam Purwanto (2011, p. 102), daya beda berhubungan dengan derajat kemampuan butir membedakan dengan baik perilaku pengambil tes dalam tes yang dikembangkan. Soal dapat dikatakan mempunyai daya pembeda jika soal tersebut dapat dijawab oleh siswa berkemampuan tinggi dan tidak dapat dijawab oleh siswa berkemampuan rendah. Jika suatu soal dapat dijawab oleh siswa pintar maupun kurang, berarti soal tersebut tidak mempunyai daya beda, demikian juga jika soal tersebut tidak

dapat dijawab oleh siswa pintar dan siswa kurang, berarti soal tersebut tidak baik sebab tidak mempunyai daya pembeda. Daya pembeda dapat diketahui dengan melihat besar kecilnya angka indeks diskriminasi item.

Angka indeks diskriminasi item adalah sebuah angka atau bilangan yang menunjukkan besar kecilnya daya (*discriminatory power*) yang dimiliki oleh sebutir item. *Discriminatory power* pada dasarnya dihitung atas dasar pembagian testee ke dalam dua kelompok, yaitu kelompok atas (*the higher group*) yang tergolong pandai dan kelompok bawah (*the lower group*) yaitu kelompok testee yang tergolong bodoh (Sudijono, 2011, p.387). Arikunto (2013, p. 226-232) memaparkan bahwa daya pembeda soal adalah kemampuan sesuatu soal untuk membedakan antara peserta didik yang pandai (berkemampuan tinggi) dengan peserta didik yang bodoh (berkemampuan rendah). Butir-butir soal yang baik adalah butir-butir soal yang mempunyai indeks diskriminasi 0,4 sampai dengan 0,7.

### Efektivitas Pengecoh

Alternatif jawaban atau pilihan pengecoh yang ada pada tes pilihan ganda memegang peranan penting dalam memilah peserta didik atau siswa yang kompeten dengan baik dan yang tidak. Sebuah item tes dikatakan baik, apabila pilihan pengecoh dapat berfungsi. Fungsi tersebut ditunjukkan dengan banyaknya peserta didik atau siswa yang memilih pilihan pengecoh tersebut, tentunya pilihan tersebut dilakukan oleh siswa yang tidak tuntas (berkemampuan rendah). Efektivitas fungsi pengecoh suatu item dapat diketahui dengan melakukan analisis terhadap distribusi jawaban peserta didik atau siswa pada setiap pilihan pengecoh yang disediakan. Tujuan analisis tersebut adalah untuk mengetahui pengecoh manakah yang memiliki fungsi pengecoh yang baik, yang harus dipertahankan, dan pengecoh yang mana yang fungsinya tidak baik, untuk kemudian diganti (Sriundy, 2014, p. 166).

Arikunto (2013, p. 233) memaparkan bahwa sebuah distraktor (pengecoh) dapat dikatakan berfungsi dengan baik apabila distraktor tes tersebut mempunyai daya tarik yang besar bagi pengikut-pengikut tes yang kurang memahami konsep atau kurang menguasai bahan. Butir soal yang baik pengecohnya akan dipilih secara merata oleh peserta didik yang menjawab salah. Sebaliknya, soal yang kurang baik pengecohnya akan dipilih tidak merata. Zulaiha (2008, p. 18) mengemukakan bahwa menentukan efektivitas pengecoh menggunakan *Prop. Endorsing* yaitu penyebaran pilihan jawaban (*option*) atau persentase peserta didik yang menjawab pilihan jawaban.

Berdasarkan latar belakang masalah di atas, dapat diketahui bahwa evaluasi sangat penting dalam pembelajaran di kelas. Guru harus dapat membuat soal yang benar-benar mengevaluasi pembelajaran. Akan tetapi, tidak semua guru melakukan analisis butir soal sebelum soal tersebut diberikan kepada peserta didik. Karena itulah, peneliti tertarik untuk mengangkat permasalahan tersebut sebagai fokus utama dalam penelitian ini. Berdasarkan pengamatan penulis, di SMP N 1 Ngemplak terdapat 6 kelas untuk tingkat kelas VIII, yakni kelas VIII A, VIII B, VIII C, VIII D, VIII E, dan VIII F. Terdapat dua guru yang mengampu Mata Pelajaran PJOK. Guru Mata Pelajaran PJOK membuat sendiri instrumen penilaian ulangan akhir semester. Ulangan Akhir Semester dilakukan untuk mengetahui sejauh mana kemampuan peserta didik dalam menguasai materi pembelajaran Pendidikan Jasmani, Olahraga, dan Kesehatan dalam bentuk teori selama satu semester. Mengingat pentingnya tes Ulangan Akhir Semester, dalam pelaksanaannya diperlukan soal yang berkualitas sehingga dapat memberikan gambaran mengenai kelebihan dan kekurangan peserta didik dalam ranah kognitif. Kedua guru mata pelajaran PJOK tersebut belum melakukan analisis butir soal secara mendalam. Oleh karena itu, kualitas butir soal Ulangan Akhir Semester yang dibuat oleh guru PJOK belum diketahui bila ditinjau dari segi tingkat kesukaran, daya pembeda, dan efektivitas pengecoh untuk soal pilihan ganda. Keterbatasan aplikasi menjadi kendala guru PJOK dalam menganalisis kualitas soal. Berdasarkan uraian tersebut, peneliti bermaksud untuk melakukan penelitian dengan judul "Analisis Butir Tes Hasil Belajar Mata Pelajaran PJOK Kelas VIII".

## **METODE**

### **Jenis Penelitian**

Penelitian ini merupakan penelitian deskriptif kuantitatif dengan teknik pengumpulan data menggunakan teknik dokumentasi. Penelitian ini dilakukan dengan cara mengumpulkan data berupa soal ujian, kunci jawaban, dan lembar jawab peserta didik. Selanjutnya data tersebut diinput dan diolah menggunakan program Anates versi 4.09.

### **Waktu dan Tempat Penelitian**

Penelitian dilaksanakan pada bulan Juli sampai dengan Agustus 2018 di SMP N 1 Ngemplak Sleman Yogyakarta.

### **Subjek Penelitian**

Subjek penelitian ini adalah seluruh peserta didik kelas VIII di SMP N 1 Ngemplak sebanyak 187 peserta didik yang terbagi dalam enam kelas.

### **Teknik Pengumpulan Data**

Teknik pengumpulan data dalam penelitian ini menggunakan teknik dokumentasi dan teknik observasi. Menurut Sugiyono, (2010, p. 82) dokumen merupakan catatan peristiwa yang sudah berlalu. Dokumen bisa berbentuk tulisan, gambar, atau karya-karya monumental dari seseorang. Menurut Sudijono, (2011, p. 76) observasi adalah cara menghimpun bahan-bahan keterangan (= data) yang dilakukan dengan pengamatan dan pencatatan secara sistematis terhadap fenomena-fenomena yang sedang dijadikan sasaran pengamatan.

Teknik dokumentasi dalam penelitian ini digunakan untuk mendokumentasikan soal ulangan akhir Mata Pelajaran PJOK peserta didik kelas VIII SMP N 1 Ngemplak, Kabupaten Sleman, Provinsi Daerah Istimewa Yogyakarta dan lembar jawaban peserta didik. Sedangkan teknik observasi digunakan untuk pengamatan dan observasi lebih jauh pada lembar soal ulangan akhir semester genap dan lembar jawaban peserta didik yang lebih lanjut akan dapat membantu proses analisis data.

### **Teknik Analisis Data**

Soal ulangan akhir Kelas VIII SMP N 1 Ngemplak, Kabupaten Sleman, Provinsi Daerah Istimewa Yogyakarta yang berbentuk pilihan ganda dianalisis menggunakan analisis butir soal. Sedangkan analisis kuantitatif meliputi validitas, tingkat kesukaran, daya beda, dan pengecoh. Analisis data secara kuantitatif dalam penelitian ini dilakukan dengan menggunakan bantuan dari program komputer, yaitu Anates versi 4.09. Anates merupakan perangkat lunak (*software*) yang dibuat melalui bahasa pemrograman komputer yang diciptakan khusus untuk analisis statistik butir soal dan tes. Peneliti dalam analisis data menggunakan program komputer yang khusus untuk menganalisis butir soal untuk mencari validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan keefektifan penggunaan distraktor. Program ini dipilih karena dalam penggunaannya sangat efektif dalam menganalisis kualitas butir soal pilihan ganda, sederhana dan mudah pengoperasiannya dibandingkan dengan program lain.

Ada tiga tahap analisis butir soal menggunakan program Anates yang terdiri atas: memasukkan data; analisis data; dan hasil analisis.

1. Memasukkan data (*Input Data*): 1) Buka program Anates Pilihan Ganda; 2) Pada kolom file, pilih opsi "Buat File Baru"; 3) Kemudian muncul tampilan "Informasi Jawaban Subjek" yang berisikan tiga kolom jawaban, yaitu: Jumlah subjek (jumlah peserta didik yang akan dianalisis), jumlah butir soal (jumlah soal yang akan dianalisis), dan jumlah pilihan jawaban. Isikan sesuai dengan data yang ada. Kemudian klik "OK"; 4) Kemudian muncul tabel data yang masih kosong. Terdiri dari beberapa kolom, yaitu: Nomor urut (nomor urut peserta didik), nomor subyek (nama peserta didik), nomor butir baru (nomor soal), nomor butir asli (nomor soal); 5) Di baris pertama, terdapat keterangan "Nama subyek I kunci". Isikan kunci jawaban di baris pertama sesuai dengan nomor soal; 6) Di baris kedua isikan nama peserta didik dan

- jawaban peserta didik. Begitu seterusnya hingga semua data dimasukkan; 7) Setelah semua data berhasil dimasukkan, klik "SIMPAN" diopsi paling atas di atas tabel.
2. Analisis Data; 1) Bila file sudah tersimpan, klik opsi "Baca File". Pilih file input yang sudah tersimpan; 2) Kemudian klik opsi "Kembali ke Menu Utama". Di kolom bagian penyekoran, pilih opsi "Olah Semua Otomatis". Kemudian akan muncul hasil analisis data dilihat dari reliabilitas, daya pembeda, tingkat kesukaran, efektivitas pengecoh, dan validitas; 3) Variabel Penelitian Analisis butir soal merupakan pengkajian butir-butir soal yang bertujuan untuk memperoleh butir soal yang baik. Soal yang baik dapat ditinjau dari beberapa aspek, yaitu: (a). Tingkat Kesukaran, Tingkat kesukaran adalah banyaknya peserta didik yang menjawab benar untuk suatu butir soal, yang diperoleh dengan cara menentukan banyaknya siswa yang menjawab soal itu dengan benar terhadap jumlah seluruh siswa. Semakin homogen alternatif jawaban suatu butir soal, semakin tinggi tingkat kesukaran suatu soal; (b) Daya Pembeda, Daya pembeda adalah kemampuan suatu soal untuk membedakan peserta didik yang mampu dengan siswa yang kurang mampu dalam mengerjakan soal. Daya pembeda soal akan mengkaji soal-soal tes dari segi kemampuan tes tersebut dalam mengelompokkan peserta didik yang masuk dalam kategori prestasi tinggi maupun rendah; (c) Efektivitas Pengecoh, Efektifitas pengecoh adalah distribusi testee dalam menentukan pilihan jawaban pada soal berbentuk pilihan ganda. Pola jawaban soal dapat menentukan apakah pengecoh berfungsi dengan baik atau tidak. Semakin homogen alternatif jawaban, semakin baik efektivitas pengecoh.
  3. Hasil Analisis  
Hasil akhir dari analisis menggunakan Program Anates adalah status atau kondisi soal yang ditinjau dari: a) Tingkat kesukaran, b) Daya pembeda, dan c) Efektivitas pengecoh.

## HASIL DAN PEMBAHASAN

### Hasil Penelitian

Penelitian ini dilaksanakan untuk mengevaluasi butir soal ulangan akhir Mata Pelajaran Pendidikan Jasmani Olahraga dan Kesehatan Kelas VIII SMP N 1 Ngemplak dilihat dari segi validitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecoh. Instrumen yang digunakan untuk menganalisis kualitas butir soal antara lain: Soal ujian akhir semester, kunci jawaban, dan lembar jawaban peserta didik yang mengikuti ujian. Jumlah soal pilihan ganda yang dianalisis adalah sebanyak 40 butir soal dengan populasi sebanyak 187 peserta didik yang mengikuti ujian akhir semester genap. Pengolahan instrumen yang ada yaitu dengan metode dokumentasi. Instrumen berupa soal, kunci jawaban, dan lembar jawaban peserta didik dianalisis menggunakan Program Anates versi 4.09 sehingga dapat diketahui kualitas butir soal yang dilihat dari segi tingkat kesukaran, daya pembeda, dan efektivitas pengecoh. Adapun hasil analisis butir soal memperoleh hasil sebagai berikut:

### Tingkat Kesukaran

Berdasarkan hasil analisis butir soal pilihan ganda ulangan akhir Mata Pelajaran Pendidikan Jasmani Olahraga dan Kesehatan Kelas VIII SMP N 1 Ngemplak melalui program Anates 4.09 diperoleh hasil bahwa dari keseluruhan 40 soal pilihan ganda yang dianalisis, menunjukkan sebanyak 20 soal atau dapat dikatakan 50% soal dalam kategori mudah, 15 soal atau 37,5% soal dalam kategori sedang, dan 5 soal atau 12,5% soal dalam kategori sukar.

Tabel 1. Distribusi Tingkat Kesukaran Butir Soal

No.	Kategori	No. Soal	Jumlah	Persentase
1	Sukar	2, 21, 27, 30, 39	5	12,5%
2	Sedang	1, 3, 6, 9, 10, 13, 14, 15, 20, 26, 28, 29, 31, 36, 37	15	37,5%
3	Mudah	4, 5, 7, 8, 11, 12, 16, 17, 18, 19, 22, 23, 24, 25, 32, 33, 34, 35, 38, 40	20	50%

Kemudian bila dilihat dalam diagram lingkaran hasil analisis butir soal pilihan ganda Mata Pelajaran Pendidikan Jasmani Olahraga dan Kesehatan Kelas VIII SMP N 1 Ngemplak adalah sebagai berikut:



Gambar 1. Diagram Pie Persentase Tingkat Kesukaran Butir Soal

#### Daya Pembeda

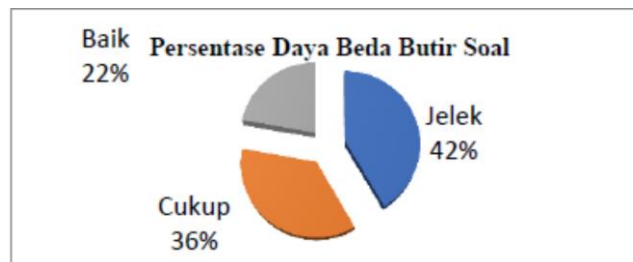
Berdasarkan dari hasil analisis butir soal pilihan ganda ulangan akhir Mata Pelajaran Pendidikan Jasmani Olahraga dan Kesehatan Kelas VIII SMP N 1 Ngemplak melalui program Anates 4.09 maka dapat diperoleh hasil bahwa dari keseluruhan butir sebanyak 40 soal pilihan ganda yang diujikan, terdapat butir soal yang memiliki daya pembeda jelek berjumlah 19 butir soal atau sebanyak 47,5%, butir soal yang memiliki daya pembeda cukup berjumlah 13 butir soal atau sebanyak 32,5%, butir soal yang memiliki daya pembeda baik berjumlah 8 butir soal atau sebanyak 20%.



**Tabel 5. Distribusi Daya Pembeda Butir Soal**

No	Daya Pembeda	Soal Pilihan Ganda		
		No. Butir	Jumlah	Persentase
1	<0,20 Jelek	5, 9, 10, 13, 16, 17, 18, 19, 21, 22, 23, 24, 25, 27, 32, 33, 34, 35, 39	19	47,5%
2	0,21 – 0,4 0 Sedang	1, 2, 3, 4, 6, 7, 8, 9, 11, 14, 15, 31, 37	13	32,5%
3	0,41 - 0,7 0 Baik	12, 20, 26, 28, 29, 30, 36, 38	8	20%

Kemudian apabila hasil analisis butir soal pilihan ganda Mata Pelajaran Pendidikan Jasmani Olahraga dan Kesehatan Kelas VIII SMP N 1 Ngemplak adalah sebagai berikut:



**Gambar 3. Diagram Pie Persentase Daya Beda Butir Soal**

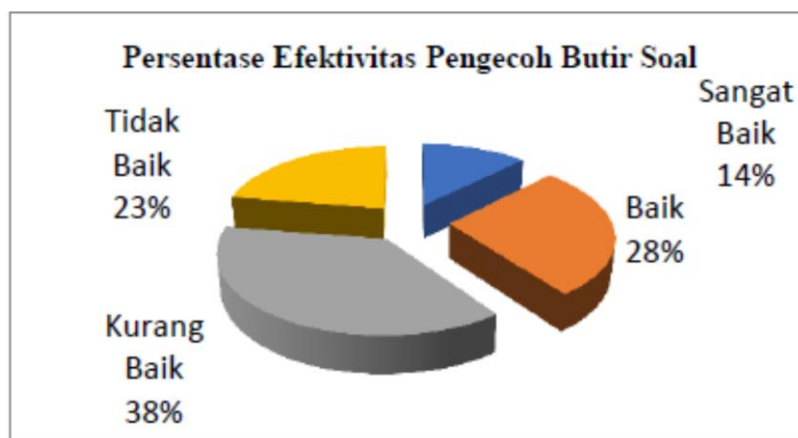
### Efektivitas Pengecoh/Distraktion

Efektivitas pengecoh dianalisis menggunakan program Anates Versi 4.09 soal pilihan ganda Mata Pelajaran Pendidikan Jasmani Olahraga dan Kesehatan Kelas VIII SMP N 1 Ngemplak dikatakan baik apabila dipilih >5% dari jumlah peserta didik. Dalam menginterpretasikan Efektivitas Pengecoh setiap butir soal digunakan kriteria pada skala Likert sebagai berikut: 1) Efektivitas pengecoh dikatakan sangat baik apabila ketiga pengecoh berfungsi; 2) Efektivitas pengecoh dikatakan baik apabila terdapat dua pengecoh yang berfungsi; 3) Efektivitas pengecoh dikatakan cukup baik apabila terdapat satu pengecoh yang berfungsi; 4) Efektivitas pengecoh dikatakan tidak baik apabila semua pengecoh tidak berfungsi. Dilihat dari efektivitas pengecoh dari data yang telah dianalisis didapatkan hasil bahwa dari 40 butir soal, terdapat 5 butir soal atau 12,5% soal dalam kategori sangat baik, 11 butir soal atau 27,5% soal dalam kategori baik, 15 butir soal atau 37,5% dalam kategori kurang baik, dan 9 soal atau 22,5% dalam kategori tidak baik. Distribusi berdasarkan indeks efektivitas pengecoh adalah sebagai berikut:

**Tabel 6. Distribusi Efektivitas Pengecoh Butir Soal**

No.	Efektivitas Pengecoh	Butir Soal	Jumlah	Persentase
1	Sangat Baik	21, 25, 32, 33, 37	5	12,5%
2	Baik	4, 7, 8, 16, 17, 19, 20, 24, 27, 34, 35	11	27,5%
3	Kurang Baik	1, 3, 5, 10, 12, 13, 15, 18, 22, 28, 29, 30, 31, 36, 39	15	37,5%
4	Tidak Baik	2, 6, 9, 11, 14, 23, 26, 38, 40	9	22,5%

Kemudian apabila soal pilihan ganda Mata Pelajaran Pendidikan Jasmani Olahraga dan Kesehatan Kelas VIII SMP N 1 Ngemplak dilihat menggunakan diagram lingkaran adalah sebagai berikut:



**Gambar 4. Diagram Pie Persentase Efektivitas Pengecoh Butir Soal**

### Pembahasan Tingkat Kesukaran

Arikunto (2013, p. 210) memaparkan bahwa kriteria yang digunakan sebagai tolak ukur tingkat kesukaran suatu soal adalah semakin kecil indeks yang diperoleh, makin sulit soal tersebut sebaliknya, semakin besar indeks yang diperoleh, semakin mudah soal tersebut. Butir soal yang baik memiliki indeks kesukaran antara 30%-70%. Butir soal yang memiliki tingkat kesukaran di bawah 30% dikatakan sukar. Butir soal yang memiliki tingkat kesukaran 30%-70% dikatakan sedang, sedangkan tingkat kesukaran di atas 70% dikatakan terlalu mudah. Berdasarkan analisis tingkat kesukaran yang telah dilakukan menggunakan program Anates 4.09 memperoleh hasil bahwa dari 40 butir soal pilihan ganda, terdapat 5 soal berkategori sukar yang terdapat pada nomor 2, 21, 27, 30, 39 atau 12,5% dari total soal. 15 soal berkategori sedang yang terdapat pada nomor 1, 3, 6, 9, 10, 13, 14, 15, 20, 26,

28, 29, 31, 36, 37 atau 37,5% dari total soal. Dan 20 soal berkategori mudah yang terdapat pada nomor 4, 5, 7, 8, 11, 12, 16, 17, 18, 19, 22, 23, 24, 25, 32, 33, 34, 35, 38, 40 atau 50% dari total soal. Hasil persentase yang diperoleh menunjukkan bahwa tingkat kesukaran soal 70%- 100% sebanyak 20 butir (50%), tingkat kesukaran soal antara 30%-70% sebanyak 15 butir (37,5%), dan tingkat kesukaran soal antara 0-30% sebanyak 5 butir (12,5%). Anas Sudijono (2015: 373) mengemukakan bahwa apabila  $P < 0,30$  maka interpretasinya terlalu sukar, apabila  $P > 0,30-0,70$  interpretasinya sedang, sedangkan  $P > 0,70$  interpretasinya terlalu mudah. Hasil tersebut menunjukkan bahwa tingkat kesukaran soal memiliki mean  $P = 63,1\%$ . Sehingga dapat dikatakan bahwa tingkat kesukaran soal pilihan ganda Mata Pelajaran Pendidikan Jasmani Olahraga dan Kesehatan Kelas VIII SMP N 1 Ngemplak masuk dalam mean  $P$  antara 30%-70% yang dapat diinterpretasikan memiliki tingkat kesukaran soal yang sedang.

### Daya Pembeda

Arikunto (2013, p. 226-232) memaparkan bahwa daya pembeda soal adalah kemampuan sesuatu soal untuk membedakan antara peserta didik yang pandai (berkemampuan tinggi) dengan peserta didik yang bodoh (berkemampuan rendah). Butir-butir soal yang baik adalah butir-butir soal yang mempunyai indeks diskriminasi 0,4 sampai dengan 0,7. Berdasarkan hasil penelitian yang telah dianalisis menggunakan program Anates versi 4.09, menunjukkan bahwa dari 40 butir soal, sebanyak 19 butir termasuk dalam kategori soal jelek yang terdapat pada nomor 5, 9, 10, 13, 16, 17, 18, 19, 21, 22, 23, 24, 25, 27, 32, 33, 34, 35, 39, sebanyak 13 butir soal termasuk dalam kategori soal sedang yang terdapat pada nomor 1, 2, 3, 4, 6, 7, 8, 9, 11, 14, 15, 31, 37, dan sebanyak 8 butir soal termasuk dalam kategori baik yang terdapat pada nomor 12, 20, 26, 28, 29, 30, 36, 38. Sudijono (2015, p. 389) mengemukakan apabila besarnya  $D$  kurang dari 0,20 maka soal dikatakan lemah sekali/jelek, apabila nilai  $D$  di antara 0,20-0,40 maka dikategorikan sedang/cukup, apabila nilai  $D$  di antara 0,40-0,70 memiliki kategori baik, dan apabila nilai  $D$  di antara 0,70-1,0 dikategorikan sangat baik. Maka menurut hasil analisis, diketahui bahwa soal yang memiliki nilai  $D$  sebesar  $< 0,20$  sebanyak 19 butir soal (47,5%), soal yang memiliki nilai  $D$  sebesar 0,20-0,40 sebanyak 13 butir soal (32,5%), dan soal yang memiliki nilai  $D$  sebesar 0,40-0,70 sebanyak 8 butir soal (20%).

### Efektivitas Pengecoh/Distraktion

Arikunto (2013, p. 233) memaparkan bahwa sebuah distraktor (pengecoh) dapat dikatakan berfungsi dengan baik apabila distraktor tes tersebut mempunyai daya tarik yang besar bagi pengikut-pengikut tes yang kurang memahami konsep atau kurang menguasai bahan. Butir soal yang baik pengecohnya akan dipilih secara merata oleh peserta didik yang menjawab salah. Sebaliknya, soal yang kurang baik pengecohnya akan dipilih tidak merata. Kriteria efektivitas pengecoh adalah sebagai berikut: Zulaiha (2008, p. 18) mengemukakan bahwa menentukan efektivitas pengecoh menggunakan *Prop. Endorsing* yaitu penyebaran pilihan jawaban (*option*) atau persentase peserta didik yang menjawab pilihan jawaban. Harapannya adalah pengecoh (selain kunci) dapat dipilih oleh sedikitnya 0,0025 (2,5%). Menurut hasil data yang telah dianalisis menggunakan program Anates versi 4.09, fungsi distraktor yang berfungsi sangat baik atau 3 pengecoh berfungsi dengan efektif berjumlah 5 butir soal yang terdapat pada nomor 21, 25, 32, 33, 37, distraktor yang berfungsi baik atau 2 pengecoh berfungsi dengan efektif berjumlah 11 butir soal yang terdapat pada nomor 4, 7, 8, 16, 17, 19, 20, 24, 27, 34, 35, distraktor yang berfungsi kurang baik atau hanya 1 pengecoh yang berfungsi dengan efektif berjumlah 15 butir soal yang terdapat pada nomor 1, 3, 5, 10, 12, 13, 15, 18, 22, 28, 29, 30, 31, 36, 39, dan distraktor yang jelek atau tidak ada pengecoh yang berfungsi berjumlah 9 butir soal yang terdapat pada nomor 2, 6, 9, 11, 14, 23, 26, 38, 40. Dari hasil tersebut dapat diketahui bahwa dari 40 soal, sebanyak 12,5% soal memiliki efektivitas pengecoh yang sangat baik, 27,5% memiliki efektivitas pengecoh yang baik, 37,5% memiliki efektivitas pengecoh yang cukup baik, dan 22,5% soal memiliki efektivitas pengecoh yang jelek. Pengecoh yang cukup baik dan jelek harus diperbaiki.

Berdasarkan hasil keseluruhan analisis soal pilihan ganda Mata Pelajaran Pendidikan Jasmani Olahraga dan Kesehatan Kelas VIII SMP N 1 Ngemplak dengan menggunakan program Anates Versi 4.09, diperoleh hasil bahwa dari 40 butir soal pilihan ganda Mata Pelajaran Pendidikan Jasmani Olahraga dan Kesehatan Kelas VIII SMP N 1 Ngemplak, dilihat dari tingkat kesukaran adalah 5 soal (12,5%) berkategori sukar, 15 soal (37,5%) berkategori sedang, dan 20 soal (50%) berkategori mudah. Ditinjau dari daya pembeda, sebanyak 8 soal (20%) berkategori baik, 13 soal (32,5%) berkategori sedang, dan 19 soal (47,5%) berkategori jelek. Ditinjau dari efektivitas pengecoh, sebanyak 5 soal (12,5%) berkategori sangat baik, 11 soal (27,5%) berkategori baik, 15 soal (37,5%) berkategori kurang baik, dan 9 soal (22,5%) berkategori tidak baik. Untuk soal yang layak masuk dalam bank soal adalah soal dengan kategori kualitas sangat baik, sedangkan untuk soal yang masuk dalam kategori baik dan sedang sebaiknya direvisi sesuai dengan kekurangan soal masing-masing dilihat dari tiga kriteria kualitas soal yang baik yang meliputi tingkat kesukaran, daya pembeda, dan efektivitas pengecoh. Sedangkan untuk soal yang memiliki kriteria tidak baik dan sangat tidak baik sebaiknya soal tersebut dibuang atau digantikan dengan soal yang baru karena soal dalam kategori ini tidak layak untuk dijadikan instrumen evaluasi pembelajaran.

## SIMPULAN

Berdasarkan hasil penelitian dan pembahasan analisis butir soal yang meliputi segi validitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecoh pada soal pilihan ganda ujian akhir Mata Pelajaran Pendidikan Jasmani Olahraga dan Kesehatan Kelas VIII SMP N 1 Ngemplak maka dapat disimpulkan bahwa dari total 40 butir soal pilihan ganda ditinjau dari tingkat kesukaran soal menunjukkan sebanyak 20 soal (50%) dalam kategori mudah, 15 soal (37,5%) dalam kategori sedang, dan 5 soal (12,5%) dalam kategori sukar. Ditinjau dari kategori daya pembeda didapat 19 butir soal (47,5%) berkategori jelek, 13 butir soal (32,5%) berkategori cukup, 8 butir soal (20%) berkategori baik. Ditinjau dari segi efektivitas pengecoh/distractor terdapat 5 butir soal (12,5%) soal dalam kategori sangat baik, 11 butir soal (27,5%) soal dalam kategori baik, 15 butir soal (37,5%) dalam kategori kurang baik, dan 9 butir soal (22,5%) dalam kategori tidak baik.

## Saran

Berdasarkan dari hasil kesimpulan yang telah diperoleh di atas maka ada beberapa saran yang bisa disampaikan yaitu: a) Berdasarkan hasil analisis terhadap soal pilihan ganda ujian tengah semester Mata Pelajaran Pendidikan Jasmani Olahraga dan Kesehatan sudah diketahui soal yang berkualitas dan yang tidak. Maka diharapkan untuk soal yang sudah baik bisa dipertahankan untuk dijadikan alat ukur ujian, sedangkan untuk soal yang belum baik untuk dilakukan revisi sebelum digunakan kembali. Untuk soal yang perlu diganti yaitu untuk soal yang masuk dalam kategori tidak baik dan sangat tidak baik. Sedangkan untuk soal yang masuk dalam kategori baik dan cukup baik perlu diperbaiki lagi untuk aspek yang belum berfungsi; b) Dalam merancang atau membuat ukur yang akan dijadikan evaluasi sebaiknya guru menjalankan semua proses yang sudah menjadi prosedur karena jika soal yang akan dijadikan bahan ujian tersebut masih belum baik maka proses evaluasi yang akan dilakukan belum bisa dijadikan alat ukur kualitas peserta didik. Analisis butir soal penting dilakukan agar evaluasi berjalan maksimal; c) Proses evaluasi kinerja setiap guru untuk lebih ditingkatkan lagi mengingat evaluasi merupakan bagian penting dari pembelajaran sebagai upaya untuk meningkatkan kualitas peserta didik.

## DAFTAR PUSTAKA

- Aiken, Lewis R. (1994). *Psychological testing and assessment* (Eight ed.). Boston: Allyn and Bacon.
- Arikunto, S. (2013). *Dasar-dasar evaluasi pendidikan* Edisi 2. Jakarta: PT Bumi Aksara.
- Daryanto. (2008). *Evaluasi pendidikan*. Jakarta: Rineka Cipta.

- Gronlund, N. E. (1982). *Measurement and evaluation in teaching* (4th ed.). Cliffs, NY: Macmillan.
- Kusaeri & Suprananto. (2012). *Pengukuran dan penilaian pendidikan*. Yogyakarta: Graha Ilmu.
- Kusumawati, M. & Hadi, S. (2018). An analysis of multiple choice questions (MCQs): Item and test statistics from mathematics assessments in senior high school. *REiD (Research and Evaluation in Education)* 4(1) from doi: <https://doi.org/10.21831/reid.v4i1.20202>
- Mardapi, D. (2012). *Pengukuran, penilaian, dan evaluasi pendidikan*. Yogyakarta: Nuha Medika.
- Purwanto. (2011). *Evaluasi hasil belajar*. Yogyakarta: Pustaka Pelajar.
- Sudaryono. (2012). *Dasar-dasar evaluasi pembelajaran*. Yogyakarta: Graha Ilmu.
- Surapranata, S. (2009). *Analisis, validitas, reliabilitas dan interpretasi hasil tes*. Bandung: PT Remaja Rosdakarya.
- Sriundy M., I Made. (2014). *Evaluasi pengajaran* (Edisi Revisi). Surabaya: Unesa University Press.
- Sudijono, A. (2009). *Pengantar evaluasi pendidikan*. Jakarta: PT. Raja Grafindo Persada.
- Sudijono, A. (2011). *Pengantar evaluasi pendidikan*. Jakarta: PT. Raja Grafindo Persada.
- Sugiyono. (2010). *Metode Penelitian pendidikan pendekatan kuantitatif, kualitatif, dan R&D*. Bandung: Alfabeta.
- Supardi. (2016). *Penilaian autentik pembelajaran afektif, kognitif dan psikomotor* (Konsep dan Aplikasi). Jakarta: Rajawali Pres.
- Tshabalala, T., & Ncube, A. C. (2014). The effectiveness of measurement and evaluation in Zimbabwean primary schools: Teachers and heads' perceptions. *International Journal of Innovation and Applied Studies*, 8(1), 141-148.
- Zulaiha, R. (2008). *Analisis soal secara manual*. Jakarta: Puspendik Balitbang Depdiknas.