

ANALISIS BUTIR SOAL EVALUASI TEMA 1 KELAS 4 SDN PLUMBUNGAN MENGUNAKAN PROGRAM ANATES

Mochammad Noor Akhmadi

Program Studi Teknologi Pendidikan Sekolah Pasca Sarjana Institution Universitas Adibuana Surabaya
Noorakhmadi49@gmail.com

Abstrak

Program anates digunakan untuk mengetahui ciri-ciri butir soal evaluasi tema 1 kelas 4. Pada penelitian ini pengumpulan data menggunakan Teknik dokumen dengan pendekatan dekriptif kuantitatif. Alat tes berupa soal evaluasi tema 1 kelas 4 berjumlah 20 soal berupa dokumen tes pilihan ganda. Soal evaluasi tema 1 kelas 4 direspon oleh 20 siswa dengan menggunakan aplikasi program anates versi 4.0.9 dengan jumlah 20 dokumen. Setelah dilakukan analisis, hasil yang didapat menunjukkan indeks daya pembeda dikategorikan tinggi karena jumlah butir soal yang bisa digunakan untuk membedakan kelompok atas dan bawah adalah 50%. Tingkat kesukaran butir soal kategori sedang dengan nilai prosentase sebesar 64%. Validitas tes sebesar 80% kategori sangat tinggi. Reliabilitas butir soal sangat tinggi dengan nilai indeks sebesar 0.80. Efektifitas distraktor pada soal evaluasi Penilaian Tengah Semester Tema 1 Kelas IV hanya 40 % yang berfungsi dengan baik.

Kata Kunci: Karakteristik Butir, Soal Evaluasi PTS, Program Anates

Abstract

The anates program was used to determine the characteristics of the items in the evaluation of the theme 1 class 4. In this study, data collection used document techniques with a quantitative descriptive approach. The test tool is in the form of evaluation questions on theme 1 in class 4 totaling 20 questions in the form of multiple choice test documents. The question of evaluating theme 1 in class 4 was responded to by 20 students using the Anates program application version 4.0.9 with a total of 20 documents. After the analysis was carried out, the results obtained showed that the index of discriminating power was categorized as high because the number of items that could be used to distinguish between upper and lower groups was 50%. The level of difficulty of the items in the medium category with a percentage value of 64%. The validity of the test is 80% very high category. The reliability of the items was very high with an index value of 0.80. The effectiveness of the distractor in the evaluation questions for the Middle Semester Theme 1 Class IV assessment was only 40% which functioned well.

Keywords: Item Characteristics, PTS Evaluation Questions, Anates Program

PENDAHULUAN

Evaluasi adalah kegiatan yang dilakukan untuk mengetahui, mengumpulkan, menentukan serta membuat keputusan untuk sejauh mana tercapainya tujuan pembelajaran. Informasi yang dikumpulkan akan digunakan untuk menghasilkan pilihan terbaik dalam menentukan keputusan (Arifin 2011). Dalam dunia Pendidikan sangat dibutuhkan dan sangat penting peran dari evaluasi. Dengan melakukan evaluasi dapat mengetahui sejauh mana perkembangan serta kemajuan Pendidikan serta hasil dari evaluasi dapat analisis menjadi data kualitatif atau kuantitatif dengan kriteria

yang sudah ditentukan.

Hasilnya dipakai untuk menentukan dalam membuat keputusan dalam bidang Pendidikan dan pengajaran. Pengukuran adalah membandingkan hasil tes dengan standar yang ditetapkan. Pengukuran bersifat kuantitatif. Evaluasi erat hubungannya dengan penilaian, dalam mengadakan penilaian seharusnya kita melaksanakan pengukuran terlebih dahulu. Sehingga pengukuran hasil belajar siswa dikhususkan pada ruang lingkup penilaian. Gronlund&Linn (1990:5) dalam Kusaeri & Suprananto berpendapat Penilaian adalah kegiatan yang terorganisir yang bertujuan untuk mengetahui sejauh mana seorang siswa atau kelompok belajar dalam

mencapai tujuan pembelajaran yang sudah ditetapkan dari pengetahuan, sikap serta keterampilan. Kegiatan tersebut meliputi pengumpulan, penganalisisan, dan penjelasan informasi melalui penilaian. Selama proses pembelajaran alat ukur yang digunakan untuk mengetahui hasil belajar siswa berupa tes. Tes sering diterapkan guna mengukur kemampuan siswa dalam tingkat perkembangan serta ketercapaian tujuan pembelajaran yang ditempuh selama pembelajaran dalam kurun waktu tertentu. Tes digunakan oleh guru dalam menilai untuk mendapatkan informasi kemajuan kompetensi siswa. Sebelum tes dilakukan hendaknya dilakukan analisis kualitas tes atau analisis butir soal untuk menghasilkan tes yang berkualitas dan memenuhi syarat tes yaitu validitas, reliabilitas, objektivitas kepraktisan dan ekonomis . Arikunto (2008). Teknis tes yang sering digunakan oleh guru adalah Penilaian tengah semester (PTS), penilaian akhir semester (PAS).

Pemerintah selalu mengupayakan peningkatan mutu Pendidikan dan pengendalian mutu Pendidikan sejak dulu. Peningkatan sumber daya manusia (SDM) sudah berada dalam program tersebut untuk itu diperlukan informasi serta alat ukur yang tepat apakah ada perubahan menuju kearah yang baik serta apakah program-program pendidikan yang sudah dilaksanakan selama ini sudah berjalan dengan baik sehingga hasil yang dicapai sesuai dengan tujuan. Agar evaluasi dapat mengenai sasaran yang tepat untuk itu penyusun evaluasi harus memperhatikan dengan sungguh-sungguh fungsi dari evaluasi. Evaluasi akan menghasilkan informasi tentang kemajuan hasil belajar siswa yang nantinya digunakan oleh guru dalam melakukan umpan balik dalam perbaikan kegiatan pembelajaran. Sudijono (2007) berpendapat perbaikan kualitas pendidikan dimulai dari melaksanakan proses evaluasi yang akan menghasilkan informasi dan umpan balik yang digunakan untuk pengukuran keefektifan sistem pembelajaran secara luas. Untuk itu dalam pelaksanaannya pengukuran memiliki peranan yang sangat penting dalam kegiatan evaluasi. Hasil dari pengukuran tergantung dari

baik buruknya kualitas evaluasi.

Untuk mengetahui kualitas dari soal dilakukan analisis butir soal terlebih dahulu agar sesuai dengan tujuan perbaikan dalam proses belajar mengajar . Uji mutu dilakukan dari perangkat soal dalam beberapa aspek dapat dilakukan secara kualitatif atau kuantitatif . Tujuan dari menganalisis butir soal adalah untuk memperoleh informasi ciri-ciri setiap butir soal yang digunakan untuk mengetahui kualitas instrument dan hasil belajar responden.

Membuat instrument tes yang sesuai dengan kreterian penyusunan intrumen evaluasi adalah tugas utama dari seorang guru selain tugas utama sebagai pengajar dan pendidik. Sementara temuan di lapangan masih banyak guru dalam melakukan evaluasi hanya pada aspek pengetahuan siswa dan melakukan analisis soal secara kualitatif. Disamping itu juga, analisis secara kauntitatif tidak dilakukan sehingga guru dalam memperoleh informasi dari pelaksanaan evaluasi masih belum sempurna akibat dariguru dalam melakukan pengukuran tidak sesuai dengan instrument evaluasi. Guru harus memiliki kompetensi yang baik dalam melaksanakan tugasnya sebagai pengajar dan pendidik secara menyeluruh mulai dari perancang proses pembelajaran, pelaksana pembelajaran sampai pada evaluasi pembelajaran untuk mengetahui ketercapaian tujuan pembelajaran dan umpan balik perbaikan pembelajaran. Instrumen tes yang sudah disusun oleh guru harus memiliki daya beda antara siswa yang pintar dan kurang pintar. Dalam menganalisis butir soal sudah banyak dikembangkan berbagai aplikasi yang memudahkan guru dalam melakukan analisis butir soal yaitu salah satunya aplikasi Anates versi 4.0.9.

Dalam penyusunan intrumen untuk mengukur kemampuan siswa, guru harus mampu menyusun intrumen yang baik dan tepat . Dalam penelitian ini akan menguji butir soal evaluasi tema 1 kelas 4 yang sudah di susun oleh guru berdasarkan instrumen penyusunan

evaluasi pada siswa kelas 4 SDN Plumbungan menggunakan aplikasi Anates untuk mendapatkan informasi validitas, reliabilitas, tingkat kesukaran, daya beda dan efektifitas pengecoh. Berdasarkan latar belakang diatas untuk rumusan masalah dalam penelitian ini yaitu : Bagaimana validitas, reliabilitas, tingkat kesukaran, daya beda dan efektifitas pengecoh butir-butir soal penilaian tengah semester tema 1 kelas 4 SDN Plumbungan dengan menggunakan Program aplikasi Anates ? Adapun tujuan dari penelitian ini adalah untuk mengetahui validitas, reliabilitas, tingkat kesukaran, daya beda dan efektifitas pengecoh butir-butir soal penilaian tengah semester tema 1 kelas 4 SDN Plumbungan dengan menggunakan Program aplikasi Anates. Pada penelitian ini dikategorikan penelitian deskriptif kuantitatif.

METODE

Dalam penelitian ini menggunakan metode dengan pendekatan kajian lapangan (field research) atau kancah. Pengambilan data dalam peneliti ini dari siswa kelas 4 SDN Plumbungan kecamatan Sukodono kabupaten Sidoarjo. Penelitian ini dilakukan pada siswa kelas 4 SDN Plumbungan kecamatan Sukodono kabupaten Sidoarjo. Tempat atau lokasi ini dipilih karena memiliki ciri-ciri yang bisa menggambarkan kondisi dan situasi yang sesuai dengan harapan penelitian.

Subjek dalam penelitian ini adalah sumber data seluruh dokumen soal dan seluruh lembar jawaban penilaian tengah semester tema 1 kelas 4 dengan jumlah soal 20 butir dengan tes yang dibuat dalam bentuk pilihan ganda dengan pilihan jawaban atau opsi berjumlah 4 . Objek dalam penelitian ini adalah semua masalah yang nantinya akan dikaji, dibahas dan di teliti berdasarkan hasil analisis butir soal penilaian tengah semester tema 1 kelas 4 SDN Plumbungan yang berjumlah 20 butir soal dengan jumlah 20 siswa. Pengumpulan data pada penelitian ini menggunakan Teknik wawancara dan

dokumentasi. Metode dokumentasi membantu dan mempermudah peneliti dalam mengumpulkan informasi kisi-kisi soal, soal, lembar jawaban siswa penilaian tengah semester tema 1 kelas 4 SDN Plumbungan. Penelitian ini menggunakan program Anates versi 4.0.9. Butir soal dianalisis secara kuantitatif untuk informasi tentang mengetahui daya pembeda, validitas, reliabilitas, tingkat kesukaran, dan efektifitas distraktor pada butir soal Evaluasi Penilaian Tengah Semester Tema 1 kelas 4.

PEMBAHASAN

Kegiatan analisis butir-butir soal dengan harapan menghasilkan informasi yang akan digunakan untuk memperbaiki soal atau tes disebut analisis butir soal. Hasil dari analisis akan digunakan untuk melakukan umpan balik sebagai proses perbaikan dan penyempurnaan butir-butir soal yang akan digunakan untuk kegiatan tes. Butir butir soal yang dirancang oleh guru pada saat digunakan dalam kegiatan pengukuran harus benar-benar bisa menjalankan fungsinya sebagai alat ukur hasil belajar yang mempunyai ketepatan yang baik. Berfungsi dan tidaknya sebuah soal dalam menjalankan fungsinya dilakukan dengan analisis soal. Dalam sebuah analisis butir soal dapat dilakukan dengan analisis kualitatif dan analisis kuantitatif. Dalam penelitian ini digunakan analisis kuantitatif dengan mengumpulkan dokumen-dokumen hasil dari pengukuran sehingga bisa dibuktikan secara empiric. Penelitian ini menggunakan analisis kuantitatif yang didasarkan pada bukti empirik. Dalam Analisa yang berdasarkan dengan bukti empirik dapat menunjukkan apakah analisis pada setiap butir soal dapat membedakan antara responden yang memiliki kemampuan tinggi dengan rendah. Program aplikasi anates versi 4.0.9 adalah salah satu program aplikasi yang bisa digunakan untuk menganalisis butir soal dengan mudah, cepat dan akurat yang dikembangkan khusus untuk menganalisis butir soal pilihan ganda dan tes uraian.

Anates merupakan software pada komputer yang bertujuan untuk

menganalisis setiap butir soal secara otomatis dengan cara memeriksa jawaban yang benar dan salah secara praktis dan cepat. Program ini khususnya berguna bagi para guru umumnya untuk pemerhati evaluasi Pendidikan. Kelebihan Anates adalah menggunakan Bahasa Indonesia pada setiap petunjuk menjalankan program sehingga para pengguna khususnya para pendidik Indonesia tidak mengalami kesulitan. Anates digunakan untuk untuk mengetahui analisis setiap butir soal yang meliputi: daya pembeda, reliabilitas, kelompok unggul dan asor, tingkat kesukaran, korelasi skor butir dengan skor total dan kualitas pengecoh.

Hasil penelitian dari analisis butir soal evaluasi Penilaian Tengah Semester Tema 1 Kelas 4 SDN Plumbungan dengan menggunakan program Anates versi 4.0.9 diperoleh sebagai berikut:

Daya Pembeda

Untuk mengetahui kemampuan butir soal membedakan hasil tes dengan hasil jawaban yang benar dan hasil tes dengan hasil jawaban yang salah adalah fungsi dari daya pembeda. Pengertian yang lain daya beda adalah kemampuan soal dalam membedakan siswa yang berkemampuan tinggi dan siswa yang berkemampuan rendah. Pernyataan ini didukung dengan pendapat Arikunto bahwa daya pembeda soal adalah kemampuan suatu butir soal untuk membedakan antara responden yang memiliki kemampuan tinggi dengan rendah.

Sudijono (2007) mengatakan bahwa mencari daya beda bertujuan untuk membedakan apakah butir soal tersebut dapat membedakan kelompok siswa yang diukur sesuai dengan kemampuan yang berbeda dalam kelompok tersebut. Nilai koefisien korelasi biserial digunakan pada penelitian ini karena menunjukkan hubungan dua skor, skor butir soal dan skor keseluruhan dari peserta tes yang sama. Koefisien daya beda berkisar antara $-1,00$ sampai dengan $+1,00$. Daya beda $+1,00$ artinya kelompok atas semuanya menjawab soal dengan benar, sedangkan kelompok bawah dalam menjawab soal salah satu terhadap butir soal itu. Sebaliknya daya beda $-1,00$

artinya semua kelompok atas dalam menjawab butir soal itu salah, dan kelompok bawah dalam menjawab butir soal benar salah satu terhadap butir soal itu.

Nilai koefisien korelasi biserial digunakan pada penelitian ini karena menunjukkan keterkaitan dua skor yaitu skor butir soal dan skor keseluruhan dari peserta tes yang sama. Nilai koefisien daya beda berkisar antara $-1,00$ sampai $+1,00$. Daya beda $+1,00$ artinya kelompok atas semuanya menjawab soal dengan benar, sedangkan kelompok bawah menjawab soal salah satu terhadap butir soal itu. Sebaliknya daya beda $-1,00$ artinya semua kelompok atas dalam menjawab butir soal itu salah, dan kelompok bawah dalam menjawab butir soal benar salah satu terhadap butir soal tersebut. Kriteria indeks daya pembeda dapat dijelaskan pada tabel berikut:

Tabel.1 Kriteria Indeks Daya Pembeda Soal

Indeks Daya Pembeda	Kriteria Daya Pembeda
Negatif – 9%	Sangat buruk
10% – 19%	Buruk
20% – 29%	Cukup
30% – 49%	Baik
50% ke atas	Sangat Baik

Rentang koefisien daya pembeda berkisar antara $-1,00$ sampai $+1,00$, nilai positif saja yang mempunyai arti semakin besar indeks daya beda, maka semakin mendekati 1 maka butir tersebut mampu membedakan kelompok atas dan kelompok bawah. Butir yang dianggap mempunyai daya beda yang baik apabila butir tersebut $> 0,5$.

Berdasarkan tabel indeks daya pembeda di atas dan analisis butir soal PTS Tema 1 Kelas 4 dengan menggunakan anates, diperoleh kesimpulan bahwa daya pembeda dapat dikategorikan sangat baik. Terdapat 8 butir soal dengan kriteria sangat baik yaitu nomor 4, 6, 14, 15, 16,17,18,19 dan 20, 3 soal dengan kriteria baik yaitu nomor 2,3 dan 12. 3 soal kriteria agak baik atau cukup yaitu nomor butir soal 7, 11,13. 3 butir soal dengan kriteria sangat

buruk yaitu nomor 1, 9 dan 10, itu berarti butir soal ini harus diperbaiki. Hal ini diklasifikasi berdasarkan tabel di bawah ini :

Dari hasil analisis di atas mempunyai makna jika hasil tes siswa memperoleh skor yang tinggi berarti sudah menjawab soal dengan benar dan jika hasil tes siswa memperoleh skor yang rendah berarti dalam menjawab soal tersebut salah. Dari butir soal tersebut dapat menggambarkan perbedaan siswa yang berkemampuan tinggi adalah siswa yang sudah memahami materi dan siswa yang berkemampuan rendah adalah siswa yang kurang memahami materi. Butir soal yang menunjukkan kategori sangat memuaskan dan sangat tidak memuaskan adalah butir soal yang tidak bisa membedakan antara siswa yang berkemampuan tinggi dan siswa yang berkemampuan rendah. Butir soal yang bertanda negatif sebaiknya tidak digunakan karena siswa kelompok atas lebih banyak menjawab salah karena kualitasnya sangat jelek.

Tingkat Kesukaran

Tingkat kesukaran soal adalah peluang untuk menjawab benar suatu soal pada tingkat kemampuan tertentu yang biasanya dinyatakan dalam bentuk indeks. Indeks tingkat kesukaran ini pada umumnya dinyatakan dalam bentuk proporsi yang besarnya berkisar 0,00 - 1,00 (Bungin 2008). Semakin besar indeks tingkat kesukaran yang diperoleh dari hasil hitungan, berarti semakin mudah soal itu. Suatu soal memiliki $TK = 0,00$ artinya bahwa tidak ada siswa yang menjawab benar dan bila memiliki $TK = 1,00$ artinya bahwa siswa menjawab benar. Perhitungan indeks tingkat kesukaran ini dilakukan untuk setiap nomor soal. Pada prinsipnya, skor rata-rata yang diperoleh peserta didik pada butir soal yang bersangkutan dinamakan tingkat kesukaran butir soal itu. Soal yang terlalu mudah tidak merangsang siswa untuk mempertinggi usaha memecahkannya. Sebaliknya soal yang terlalu sukar akan menyebabkan siswa menjadi putus asa dan tidak mempunyai semangat untuk mencoba lagi karena di luar jangkauannya (Arikunto 2008).

Jumlah responden yang mampu mengerjakan soal dengan benar pada suatu instrument berhubungan dengan taraf kesukaran butir soal tersebut dengan kriteria sedang, sukar dan mudah. dan dapat dilihat pada tabel di bawah ini :

Tabel 2. Kriteria Tingkat Kesukaran Butir Soal

Kriteria	Kategori
$P \geq 0,76$	Mudah
$0,25 \leq p \leq 0,75$	Sedang
$P \leq 0,24$	Sukar

Dari hasil analisis butir soal penilaian tengah semester tema 1 kelas 4 SDN Plumbungan adalah ada 2 butir soal yang memiliki kriteria sangat mudah, 4 butir soal yang memiliki kriteria mudah, 10 butir soal yang memiliki kriteria sedang, 2 sukar dan 1 butir soal yang memiliki kriteria sangat sukar. Jika dilihat dari hasil analisis butir soal yang memiliki kriteria sedang adalah butir soal yang tidak terlalu mudah dan tidak terlalu sulit sehingga siswa akan terangsang untuk bisa memecahkan butir soal tersebut, sebaliknya jika butir soal itu terlalu sulit ini akan menyebabkan siswa akan putus asa dalam memecahkan butir soal karena siswa merasa tidak mampu untuk memecahkan soal tersebut.

Validitas

Untuk mengetahui kualitas butir soal pada penelitian ini dilakukan analisis secara kuantitatif berdasarkan data empirik khususnya soal pilihan ganda penilaian tengah semester tema 1 kelas 4 SDN Plumbungan. Pada penelitian ini menggunakan program aplikasi anates versi 4.0.9 untuk mempermudah analisis dan menghitung karakteristik aspek-aspek yang akan dihitung, Analisis kuantitatif pada soal penilaian tengah semester tema 1 kelas 4 SDN Plumbungan meliputi aspek validitas butir soal, aspek reliabilitas butir soal, aspek tingkat kesukaran, aspek daya beda butir soal, dan aspek fungsi distractor atau pengecoh. Suatu instrumen dikatakan baik apabila memiliki tingkat kesahihan atau

kevalidan yang tinggi (Arikunto 2008). Dalam pengumpulan data harus memperhatikan prinsip keandalan instrumen sehingga pengukuran dan pengamatan menghasilkan data yang valid. Menurut (Arikunto 2008) hasil pengalamannya dapat mengetahui validitas dari sebuah tes. Ada dua kelompok validitas tes yaitu validitas logis dan validitas empiris. Validitas logis meliputi validitas isi dan validitas konstruk, sedangkan validitas empiris adalah meliputi validitas kongruen dan validitas perkiraan.

Menurut Sukardi (2008) tes dikatakan valid jika koefisien sebesar 0,5 dan dapat diterima apabila tes dilakukan satu kali, apabila ada tes dengan dugaan lain yang sama dan memiliki nilai koefisien yang lebih tinggi maka koefisien 0,5 tidak dapat diterima. Sehingga jika tes dikatakan valid apabila koefisiennya paling sedikit 0,5. Menurut Azwar (2008) apabila butir soal dianggap valid nilai koefisiennya yang diperoleh antara 0,30 – 0,50. Tes yang sah atau valid artinya tes dalam menjalankan fungsi pengukuran sudah memenuhi ketepatan dan kecermatan atau sudah tepat sarannya. Semakin tepat dan cermat suatu tes maka nilai koefisiennya semakin tinggi. Berdasarkan hasil analisis data pada penelitian ini dengan menggunakan antes koefisien yang diperoleh menunjukkan validitas tes sebesar 0,80 artinya tes yang sudah dilakukan sudah menjalankan fungsi pengukurannya dengan tepat dan cermat sehingga memiliki kalsifikasi yang tinggi. Dari 20 butir soal penilaian tengah semester tema 1 kelas 4 SDN Plumbungan yang dianalisis menggunakan program aplikasi anates dihasilkan butir soal dengan kreteria sangat signifikan berjumlah 6 butir yaitu butir 14, 15, 16, 17, 18 dan 19. Butir soal dengan kreteria signifikan berjumlah 3 butir yaitu butir 4, 6 dan 20. Dan butir soal dengan kreteria yang tidak signifikan berjumlah 10 yaitu 1, 2, 3, 5, 7, 8, 9, 10, 11, 12 dan 13. Apabila suatu alat tes memiliki validitas dengan kategori tinggi maka alat tes tersebut sudah sesuai dengan fungsinya sebagai alat ukur, atau memperlihatkan informasi bahwa hasil pengukuran sudah

sesuai dengan tujuan awal tes tersebut dijalankan. Melainkan, jika alat tes yang dihasilkan memiliki koefisien validitas rendah maka tujuan pengukuran tidak sesuai dengan apa yang diharapkan.

Reliabilitas

Menurut Rufi'i (2020) Mengukur serta mengamati objek yang diukur yang dilakukan dengan konsisten pada sebuah tes disebut reliabilitas. Jika jawaban siswa terhadap sebuah pertanyaan dapat konsisten dan stabil dari waktu ke waktu dapat dikatakan bahwa tes tersebut handal atau reliabel. Akurasi, prediksi, konsistensi dan stabilitas merupakan hasil dari reliabilitas suatu tes. Jika pengukuran dikatakan reliabel maka pengukuran itu memiliki atau menghasilkan reliabilitas yang signifikan atau tinggi. Informasi yang dihasilkan dari suatu tes dapat dipercaya reliabel, apabila memiliki konsistensi mengukur apa yang seharusnya diukur. Apabila hasil pengukuran yang dilakukan menunjukkan nilai yang hampir sama meskipun dilakukan berulang kali maka dapat dikatakan soal tersebut memiliki indeks reliabilitas tinggi. Dan sebaliknya, jika menunjukkan perbedaan yang relatif jauh maka dapat dikatakan reliabilitas soal tersebut adalah rendah. Reliabilitas menunjukkan seberapa dipercaya hasil dari pengukuran tes yang sudah dilakukan. Realibilitas bertujuan untuk menentukan seberapa besar keanekaragaman yang terjadi karena adanya kesalahan pengukuran dan seberapa besar keanekaragaman tes yang sebenarnya (Supranata 2004).

Cara menentukan reliabilitas suatu tes adalah : (1) Metode dua tes, yaitu dengan melakukan satu tes yang diberikan kepada siswa yang mengikuti tes. Korelasinya dicari dari hasil kedua tes tersebut; (2) Metode satu tes, yaitu memberikan 2 kali tes kepada siswa namun dalam kurun waktu yang berbeda, maka korelasinya dicari dari hasil kedua tes tersebut; (3) Metode "split-half" (satu tes), yaitu dengan melakukan tes yang di bagi menjadi dua dengan memberikan kesamaan pada tingkat kesukaran, isi dan bentuk. Korelasinya dicari

dengan melihat skor masing-masing bagian setengah tes. (Purwanto 2010).

Dari 20 butir soal Penilaian Tengah Semester Tema 1 kelas 4 SDN Plumbungan yang dianalisis menggunakan program aplikasi anates dihasilkan indeks reliabilitas sebesar 0,80 dengan menggunakan teknik belah yaitu ganjil dan genap. Klasifikasi hasil koefisien reliabilitas dapat dilihat pada tabel dibawah ini:

Tabel 3. Klasifikasi Tingkat Reliabilitas

Tes Kategori Reliabilitas	Koefisien Korelasi
Sangat Tinggi	0,800 – 1,000
Tinggi	0,600 – 0,799
Cukup	0,400 – 0,599
Rendah	0,200 – 0,399
Sangat Rendah	0,000- 0,199

Hasil analisis yang diperoleh dari program aplikasi anates versi 4.0.9 soal evaluasi penilaian tengah semester tema 1 kelas 4 SDN Plumbungan yaitu sebesar 0,80, dengan melihat pada table diatas maka masuk dalam kelompok kategori sangat tinggi dengan rentang 0,800-1,000. Dari hasil analisis tes penilaian tengah semester tema 1 kelas 4 SDN Plumbungan menunjukkan tingkat reliabilitas dengan kriteria sangat baik dalam mengukur kemampuan siswa. Alat tes dapat dipercaya dan dapat konsisten jika reliabel. Alat tes perlu di uji coba untu membuktikan apakah alat tes tersebut reliabel pada saat digunakan. Menurut (Azwar 2008) Jika alat tes digunakan dalam beberapa kali dalam kegiatan pengukuran terhadap kelompok dengan subjek yang sama dengan hasil yang hampisama maka dapat dikatakan alat ukur tersebut dapat dipercaya.

Efektifitas Pengecoh

Kemungkinan-kemungkinan jawaban yang dipasangkan pada setiap butir jawaban pada soal pilihan ganda yang digunakan untuk menggambarkan pola disebut distraktor. Jika pilihan jawaban pada setiap item sekurang-kurangnya dipilih oleh siswa sebesar 5% maka distraktor dinyatakan sudah menjalankan fungsinya dengan benar. Pedoman yang digunakan dalam

menentukan efektifitas distrktor dapat dilakukan dengan cara perkalian antara 5% X Jumlah siswa. Menurut (Supranata 2004) Untuk mengecoh kemungkinan - kemungkinan jawaban pada siswa yang berkemampuan tinggi adalah distraktor.

Di dalam soal pilihan ganda terdapat alternatif jawaban yang berfungsi untuk mengecoh jawaban yang benar, Fungsi dari pengecoh ini sendiri adalah untu mengecoh dari jawaban yang benar. Pilihan alternatif jawaban berkisar anatara 3 samapai dengan 5 buah, yang terpasang pada setiap butir jawaban soal.

Ada beberapa factor yang mempengaruhi distraktor tidak dipilih adalah susah membedakan butir soal dan cenderung sama. Penyebab lain bisa juga dikarenakan materi belum sepenuhnya di pahami oleh siswa dan pilihan jawaban yang tersedia tidak menarik. Purwanti (2014) berpendapat materi yang kurang dipahami siswa akan menyebabkan pengecoh tidak memiliki daya Tarik kepada siswa. Pengecoh yang tidak berfungsi dengan baik untuk selanjutnya akan diganti oleh guru dengan pengecoh yang lainnya dan pengecoh yang berfungsi dengan baik akan dipergunakan oleh guru dalam melakukan kegiatan evaluasi dan akan di dokumenkan sebagai bank soal dengan perangkat tes yang lainnya dalam kurun waktu yang ditentukan yaitu 5 tahun.

Berdasarkan hasil analisis soal evaluasi Penilaian Tengah Semester Tema 1 kelas 4 SDN Plumbungan disimpulkan bahwa soal evaluasi memiliki efektifitas pengecoh dengan kriteria sangat baik, baik, buruk dan sangat buruk. Butir soal dikategorikan memiliki pengecoh yang baik apabila butir soal tersebut dipilih oleh peserta tes sebanyak 5% dari jumlah responden (Supranata 2004). Pengecoh yang dipilih secara merata oleh para responden, maka pengecoh tersebut dapat dikatakan sangat baik. Membuat jawaban pengecoh pada soal evaluasi tidaklah gampang, karena akan berdampak pada daya pembeda dan tingkat kesukaran butir soal itu sendiri. Artinya, jika efektifitas pengecoh buruk maka indeks daya pembeda akan rendah dan tingkat kesukaran soal tersebut juga rendah. Begitu

sebaliknya, jika efektifitas pengecoh suatu jawaban pada butir soal itu baik maka indeks daya pembeda akan baik dan tingkat kesukaran soal menjadi tinggi. Butir soal yang mempunyai pengecoh yang baik pada soal evaluasi Penilaian Tengah Semester Tema 1 Kelas 4 adalah pada butir soal dengan nomor 2,5,11,13,16,17,18,19 dan 20.

SIMPULAN

Berdasarkan hasil analisis soal evaluasi Penilaian Tengah Semester tema 1 kelas 4 SDN Plumbungan yang diperoleh secara keseluruhan, maka dapat ditarik kesimpulan bahwa : Daya pembeda soal dengan nomor 4, 6, 14, 15, 16,17,18,19 dan 20 dikategorikan sangat baik. 3 soal dengan kriteria baik adalah nomor 2, 3 dan 12. 3 soal dengan kriteria agak baik atau cukup adalah nomor 7, 11,13. 3 soal dengan kriteria sangat buruk adalah nomor 1, 9 dan 10,

dan harus diperbaiki. Indeks tingkat kesukaran soal sangat mudah berjumlah 2 butir, kriteria mudah berjumlah 4 butir, kriteria sedang berjumlah 10 butir, kriteria sukar berjumlah 2 butir, dan kriteria sangat sukar berjumlah 1 butir. Uji Validitas yang diperoleh dengan menggunakan aplikasi Anates versi 4.0.9 adalah 0,80 dengan klasifikasi sangat tinggi. Dari 20 soal yang dianalisis diperoleh hanya 6 butir soal yang sangat signifikan yaitu butir 14, 15, 16, 17, 18 dan 19. Butir yang signifikan berjumlah 3 butir yaitu butir 4, 6 dan 20. Dan sedangkan yang tidak signifikan berjumlah 10 yaitu 1, 2, 3, 5, 7, 8, 9, 10, 11, 12, 13. Uji Reliabilitas tes pada penelitian ini diperoleh sebesar 0,80 dengan menggunakan teknik belah yaitu genap dan ganjil. Distraktor atau pengecoh yang baik pada soal evaluasi Penilaian Tengah Semester Tema 1 kelas 4 adalah pada butir soal 2,5,11,13,16,17,18,19 dan 20.

DAFTAR PUSTAKA

- Arifin, Zainal. 2011. *Evaluasi Pembelajaran*. Bandung: Remaja Rosdakarya.
- Arikunto, Suharsimi. 2008. *Dasar-Dasar Evaluasi Pendidikan*. Jakarta: Bumi Aksara.
- Azwar, Saifuddin. 2008. *Reliabilitas Dan Validitas*. Yogyakarta: Pustaka Pelajar.
- Bungin, Burhan. 2008. *TiMetodologi Penelitian Kuantitatif*. Jakarta: Kencana.
- Purwanti, Mustika. 2014. "Analisis Butir Soal Ujian Akhir Mata Pelajaran Akuntansi Keuangan Menggunakan Microsoft Excel 2010." *Jurnal Pendidikan Ajuntansi Indonesia, XII (1)*. 2014.
- Purwanto. 2010. *Prinsip-Prinsip Dan Teknik Evaluasi Pengajaran*. Bandung: Rosdakarya.
- Rufi'i, noor achmad. 2020. *Evaluasi Dan Pengukuran*. Surabaya: Adi Buana University Press.
- Sudijono, Anas. 2007. *Pengantar Evaluasi Pendidikan*. Jakarta: Raja Grafindo.
- Sukardi. 2008. *Evaluasi Pendidikan Prinsip & Operasionalnya*. Jakarta: Bumi Aksara.
- Supranata, Sumarna. 2004. *Analisis, Validitas, Reliabilitas Dan Interpretasi Hasil Tes*. Bandung: Remaja Rosdakarya.