

BIAS GENDER UJIAN AKHIR SEMESTER GENAP FISIKA KELAS X SMA NEGERI KABUPATEN PURWOREJO TAHUN PELAJARAN 2012/2013

Nur Triyatno, Sriyono, Nur Ngazizah

Program Studi Pendidikan Fisika

Universitas Muhammadiyah Purworejo

Jalan KHA. Dahlan 3 Purworejo, Jawa Tengah

Email: ntriyatno8@gmail.com



Intisari- Telah dilakukan penelitian dengan tujuan : 1) mengetahui indeks reliabilitas; 2) mengetahui keberadaan fungsi pengecoh; 3) mengetahui besar kelsalahan pengukuran; 4) mengetahui tipe butir soal yang terkena bias gender dalam ujian akhir semester genap mata pelajaran Fisika kelas X SMA Negeri di Kabupaten Purworejo tahun pelajaran 2012/2013. Metode pengumpulan data dengan menggunakan metode dokumentasi berupa lembar jawab siswa pada soal UAS genap kelas X SMA di Kabupaten Purworejo tahun pelajaran 2012/2013. Sampel dalam penelitian ini sebanyak 640 siswa. Pengelolaan data menggunakan metode klasik dan Mantel-Haenzsel. Hasil analisis besarnya indeks reliabilitas perangkat adalah 0,807 dengan fungsi pengecoh yang berfungsi adalah 30 atau 85,71% dan estimasi kesalahan pengukuran menggunakan pendekatan klasik sebesar 2,576. Tipe soal yang terkena bias gender pada butir soal menggunakan Mantel-Haenzsel adalah tipe ingatan pada nomor 33 dan 35, tipe aplikasi pada nomor 3, 4 dan 30, serta tipe aplikasi pada nomor 10, 15, 16, 22, 24, 25, 29 dan 34.

Kata kunci: Bias Gender, Ujian Akhir Semester Genap, Mantel-haenzsel

I. PENDAHULUAN

Pembelajaran IPA menekankan pada pemberian pengalaman langsung untuk mengembangkan kompetensi siswa. Pembelajaran IPA di SMA yang meliputi Fisika, Kimia dan Biologi menekankan pada proses pemahaman IPA yang berkaitan dengan cara mencari tahu tentang alam secara sistematis. Hasil pembelajaran IPA bagi siswa bukan hanya kumpulan pengetahuan berupa fakta, konsep atau prinsip saja tetapi juga proses penemuan sehingga siswa mampu menghadapi perubahan yang terjadi sebagai akibat perkembangan ilmu pengetahuan dan teknologi

Dalam proses pembelajaran, evaluasi merupakan salah komponen penting dan tahap yang harus ditempuh oleh guru untuk mengetahui keefektifan pembelajaran. Hasil yang diperoleh dapat dijadikan *feed-back* bagi guru dalam memperbaiki dan menyempurnakan program dan kegiatan pembelajaran. Di sekolah sering guru memberikan tes antara lain ulangan harian, UAS, ujian blok, tagihan, tes lisan, tes tertulis [1].

Ulangan Akhir Semester (UAS) merupakan salah satu ujian yang wajib dilaksanakan siswa pada setiap akhir semester gasal maupun genap (biasa disebut ulangan kenaikan kelas) pada tingkat pendidikan dari mulai sekolah dasar sampai dengan perguruan tinggi. Ulangan akhir semester sendiri bertujuan untuk mengukur kemampuan siswa yang sebenarnya. Sehingga perbedaan nilai atau skor yang diperoleh siswa hanyalah semata-mata mengenai

perbedaan kemampuan mereka dan bukan disebabkan oleh adanya faktor lain seperti bias pada butir tes.

Kesenjangan pada bidang pendidikan telah menjadi faktor utama yang sangat berpengaruh terhadap bidang lain di Indonesia. Hampir semua sektor, seperti lapangan pekerjaan, jabatan, peran di masyarakat, sampai pada masalah menyuarakan pendapat antara laki-laki dan perempuan. Meskipun secara formal dalam UUD 1945 hak laki-laki dan perempuan tidak dibedakan tetapi dalam kenyataannya sangat berbeda. Faktor penyebab karena kesenjangan pendidikan yang belum setara selain masalah-masalah klasik yang cenderung menjustifikasi ketidakadilan gender seperti interpretasi teks-teks keagamaan yang tekstual dan kendala sosial budaya lainnya.

Berdasarkan uraian diatas, penelitian ini akan menggunakan metode klasik untuk mengetahui indeks reliabilitas, mengetahui keberadaan fungsi pengecoh, mengetahui besar kesalahan pengukuran, mengetahui tipe butir soal yang terkena bias gender dalam ujian akhir semester genap mata pelajaran Fisika kelas X SMA Negeri di Kabupaten Purworejo tahun pelajaran 2012/2013.

II. LANDASAN TEORI

A. Hakekat Pembelajaran Fisika

Pembelajaran Fisika seharusnya dapat memberikan pengalaman langsung pada siswa sehingga menambah kemampuan dalam

mengkonstruksi, memahami, dan menerapkan konsep yang telah dipelajari. Dengan demikian, siswa akan terlatih menemukan sendiri berbagai konsep secara holistik, bermakna, otentik serta aplikatif untuk kepentingan pemecahan masalah.

B. Penilaian Pembelajaran Fisika

Penilaian merupakan sebuah kegiatan yang kompleks, komperhensif, dan melibatkan berbagai komponen. Berbagai definisi penilaian yang dikemukakan, pada hakekatnya memiliki persamaan dan tidak berlawanan/bertolak belakang. Informasi kualitatif dan kuantitatif dapat diperoleh lewat berbagai macam teknik, yang harus sesuai dengan kompetensi yang dibutuhkan. Jika data yang dimiliki akurat, pembuatan keputusan dalam penilaian lebih mudah dibuat dan sekaligus dapat dipertanggung jawabkan.

C. Tes Hasil Belajar

Tes hasil belajar atau *achievement test* ialah tes yang dipergunakan untuk menilai hasil-hasil pelajaran yang telah diberikan oleh guru kepada siswa dalam jangka waktu tertentu [2]. Untuk melaksanakan evaluasi hasil belajar guru dapat menggunakan dua macam tes, yakni tes yang telah distandarkan dan tes buatan pendidik itu sendiri. Yang dimaksud tes yang telah distandarkan itu ialah tes yang telah mengalami proses standarisasi, yakni proses validasi dan reliabilitas sehingga tes benar-benar valid.

D. Pendekatan Klasik

Teori tes klasik dikembangkan sejak tahun 1940, dengan menggunakan pengukuran yang sangat sederhana. Teori tes klasik atau teori skor-murni klasik didasarkan pada suatu model aditif yakni skor amatan yang merupakan penjumlahan dari skor sebenarnya dan skor kesalahan pengukuran [3].

Tes dikatakan reliabel jika skor amatan mempunyai korelasi yang tinggi dengan skor yang sebenarnya. Selanjutnya dinyatakan bahwa reliabilitas merupakan koefisien korelasi antara dua skor amatan yang diperoleh dari hasil pengukuran menggunakan tes yang paralel. Dengan demikian, pengertian yang dapat diperoleh dari pernyataan tersebut adalah suatu tes itu reliabel jika hasil pengukuran mendekati keadaan peserta tes yang sebenarnya [3].

Pengertian validitas dinyatakan sejauhmana skor tampak X dapat mendekati besarnya skor murni T . Skor tampak tidak akan sama dengan skor murni kecuali apabila alat ukur yang bersangkutan memiliki validitas yang sempurna. Validitas ada tiga tipe, yaitu validitas isi, validitas konstruk, dan validitas kriteria [4].

Indeks kesukaran suatu butir sebaiknya terletak pada interval 0,3-0,7 karena pada interval ini informasi tentang kemampuan siswa akan diperoleh

secara maksimal. Sedangkan daya pembeda suatu butir soal merupakan suatu parameter yang memberikan informasi tentang seberapa besar daya soal itu untuk dapat membedakan peserta tes yang berkemampuan tinggi dengan peserta tes yang berkemampuan rendah [3].

Teori tes klasik memiliki beberapa kelemahan mendasar. Kebanyakan statistik yang digunakan dalam model tes klasik seperti tingkat kesukaran dan daya pembeda soal sangat tergantung pada sampel yang dipergunakan dalam analisis. Rerata tingkat kemampuan, rentang, dan sebaran kemampuan siswa yang dijadikan sampel dalam analisis sangat mempengaruhi nilai statistik yang diperoleh.

E. Bias Instrument Tes

Butir tes yang baik harus bebas dari bias. Tes yang baik tidak memihak pada kelompok tertentu atau golongan tertentu dari peserta tes. Tes yang baik akan memberikan hasil pengukuran yang sama terhadap peserta tes yang memiliki kemampuan sama meskipun berasal dari kelompok atau golongan yang berbeda. Bila tes memberikan hasil berbeda maka tes tersebut bias, yang berarti perangkat tersebut tidak valid secara konstruktif. Sebuah tes yang validitasnya rendah berarti tes tersebut tidak akurat mengukur apa yang seharusnya diukur.

F. Bias Gender

Dunia pendidikan sering kita tanpa sadari terjadinya bias gender. Pengertian gender itu sendiri adalah dimensi sosiokultural dalam psikologis pria dan wanita. Istilah gender dibedakan dari jenis kelamin (sex). Sex berhubungan dengan dimensi dari pria dan wanita. peran gender adalah ekspektasi sosial yang merumuskan bagaimana pria dan wanita seharusnya berpikir, merasa dan berbuat. Faktor-faktor penentu ketimpangan gender meliputi : 1) masalah lama; 2) nilai gender yang dianut oleh masyarakat; 3) nilai dan peran gender yang terdapat dalam buku ajar; 4) nilai gender yang ditanamkan oleh guru; dan 5) kebijakan yang bias gender.

G. Kajian Pustaka

Penelitian yang dilakukan Zulkarnaen (2008) tentang Keberfungsian Butir Diferensial Perangkat Tes Ujian Akhir Sekolah Berstandar Nasional (UASBN) Mata Pelajaran Matematika di Kota Kendari. Hasil penelitian ini menunjukkan bahwa analisis deskriptif berdasarkan teori tes klasik menunjukkan bahwa dari 40 butir soal yang dianalisis 28 butir yang berkualitas baik, indeks reliabilitas 0,810 dan estimasi kesalahan pengukuran sebesar 2,779. Berdasarkan teori respons butir 3 parameter, 16 butir berkualitas baik dengan nilai informasi 7,728, estimasi kesalahan pengukuran sebesar 0,359.

Kesalahan pengukuran lebih besar jika menggunakan teori tes klasik. Analisis pendeteksian DIF dengan metode Mantel-Haenszel tidak ditemukan adanya butir yang terdeteksi mengandung DIF berdasarkan perbedaan jenis kelamin (*gender*).

Penelitian yang dilakukan Wahyu Widhiarso (2012) tentang Aplikasi Uji Keberfungsian Butir Deferensial (DIF) melalui Teknik Mantel Haenszel pada Program SPSS. Penelitian ini menunjukkan bahwa pada butir yang terkait dengan kondisi diri (misalnya merasa cukup berharga, banyak hal-hal yang baik dalam diri). Perempuan cenderung mendapatkan skor tinggi. Sebaliknya pada butir terkait dengan kemampuan, laki-laki cenderung mendapatkan skor rendah. Butir – butir yang terjerangkit DIF ini perlu dievaluasi lebih lanjut agar hasil pengukuran benar-benar memberikan hasil yang akurat.

Penelitian yang dilakukan oleh Abdurahman Jalal (2005) tentang Pendeteksian *Differential Item Function* Soal Ujian Nasional SMP/MTs Provinsi Maluku Utara Tahun 2005. Hasil penelitian ini menunjukkan bahwa analisis pendeteksian DIF dengan metode delta plot menghasilkan semua butir soal pada mata pelajaran tidak terdeteksi memuat DIF. Pada metode Mantel Haenszel terdeteksi 9 butir (15%) soal bahasa Indonesia, 3 butir (10%) soal Matematika, dan 8 butir (13%) soal Bahasa Inggris memuat DIF. Hasil analisis metode uji perbedaan parameter b didapatkan 10 butir (17%) soal bahasa Indonesia, 2 butir (7%) soal matematika, dan 8 butir (13%) soal Bahasa Inggris memuat DIF.

III. METODE PENELITIAN

Penelitian ini dilaksanakan mulai bulan april 2013 sampai bulan maret 2014. Subjek penelitian ini berupa lembar jawab pilihan ganda dari 640 siswa yang diambil dengan teknik pertimbangan peneliti. Data diperoleh dengan metode dokumentasi. Pengolahan data menggunakan metode klasik dan Mantel-Haenszel. Respon siswa disunting dalam komputer, analisis dilakukan dengan dua tahapan. Tahap pertama melihat kualitas butir soal menggunakan bantuan program komputer *MicroCat ITEMAN* versi 3.00 kemudian pada tahap kedua untuk melihat adanya kandungan DIF menggunakan metode Mantel-Haenszel.

IV. HASIL DAN PEMBAHASAN

A. Hasil Analisis dengan ITEMAN

Hasil analisis ITEMAN dibedakan dalam 2 bagian. Bagian pertama disajikan beberapa hasil analisis yang mencakup analisis statistik butir soal.

Yang tercakup dalam bagian tersebut antara lain, tingkat kesukaran, daya beda, dan pengecoh.

1. Tingkat Kesukaran

Tabel 1. Distribusi tingkat kesukaran

Prop. Correct	kategori	Nomor	jumlah
$P > 0,70$	Mudah	5, 8, 13	3
$0,30 \leq P \leq 0,70$	Sedang	1, 2, 3, 4, 6, 7, 9, 10, 11, 12, 14, 15, 16, 17, 18, 20, 22, 23, 25, 26, 27, 28, 31, 32, 33, 34, 35	27
$P < 0,30$	Sukar	19, 21, 24, 29, 30	5
jumlah			35

Berdasarkan tabel 1 dapat dilihat bahwa 3 atau 8,57% butir soal tergolong mudah, 27 atau 77,14% butir soal tergolong sedang, dan 5 atau 14,29% butir soal tergolong sukar.

2. Daya Beda Soal

Tabel 2. Distribusi Indeks Daya Beda

Daya beda soal	kategori	Nomor	jumlah
0,41 – 1,00	Baik	2, 3, 4, 6, 7, 9, 10, 11, 13, 15, 16, 17, 18, 19, 20, 22, 23, 24, 25, 26, 28, 29, 30, 31, 32, 34, 35	27
0,21 – 0,40	Sedang	8, 12, 14, 27, 33	5
< 0,2	Jelek	1, 5, 21	3
jumlah			35

Berdasarkan tabel 2 dapat dilihat bahwa 27 atau 77,14% butir soal tergolong baik atau dapat diterima, 5 atau 14,29% soal tergolong sedang dan dapat diterima akan tetapi perlu direvisi dan 3 atau 8,57% butir soal tergolong jelek dan dibuang.

3. Pengecoh

Tabel 3. Statistik Keberadaan Pengecoh pada tiap Butir Tes

Statistik	Nomor	Jumlah
Berfungsi	1, 2, 3, 4, 5, 6, 7, 8, 9, 12, 13, 15, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 33, 35	30
Tidak berfungsi	10, 11, 14, 16, 34	5
Jumlah		35

Berdasarkan tabel 3, terdapat 30 butir soal atau 85,71% pengecohnya dapat berfungsi, dan 5 butir soal atau 14,29% pengecohnya kurang tidak berfungsi dengan baik.

4. Skala Statistik

Berdasarkan 35 soal Fisika yang telah diujikan pada sampel 640 siswa SMA kelas X didapatkan besar harga rata-rata skor siswa sebesar 17,384 dengan standar deviasi 5,886 atau varian 34,405. Nilai tertinggi 33,00 dan terendah 3,00 dengan nilai tengah 17,00. Indeks daya beda soal rata-rata 0,474 dengan keiringan (*skew*) sebesar 0,437. Indeks reliabilitas perangkat tes/ α $r_{xx'} = 0,807$ dengan kesalahan pengukuran $SEM = 2,576$.

B. Hasil Analisis dengan Mantel-Haenzel

Berdasarkan analisis perangkat soal ujian akhir semester genap mata pelajaran Fisika Kelas X SMA Negeri di Kabupaten Purworejo tahun pelajaran 2012/2013 berdasarkan gender dengan menggunakan metode Mantel-Haenzel diketahui terdapat 14 atau 37% butir soal yang terkena DIF yaitu soal nomor 3, 4, 10, 15, 16, 22, 24, 25, 29, 30, 33, 34 dan 35. Sedangkan 22 atau 63% butir soal tidak terkena DIF yaitu soal nomor 1, 4, 5, 6, 7, 8, 9, 11, 12, 13, 14, 17, 18, 19, 20, 21, 23, 26, 27, 28, 31 dan 32 tidak terkena DIF.

Jika $\alpha_{MH} > 1$, maka butir yang diselidiki terkena DIF yang menguntungkan kelompok acuan (kelompok laki-laki). Jika $\alpha_{MH} < 1$, maka butir yang diselidiki terkena DIF yang menguntungkan kelompok fokus (kelompok perempuan). Berdasarkan data yang diperoleh dengan melihat besar angka *common odds ratio* dan disignifikasikan dengan *chi-square* berhasil didapat, pada butir soal nomor 4, 15, 16, 22, 24, 25, 33, 34, dan 35 menguntungkan

kelompok acuan dan pada butir soal nomor 3, 10, 29 dan 30 menguntungkan kelompok fokus.

Hasil yang didapat, ke-14 butir soal tersebut dikaji kembali dengan melihat butir soal perangkat tes ujian akhir semester genap mata pelajaran Fisika kelas X SMA Negeri di Kabupaten Purworejo tahun pelajaran 2012/2013. Seteelah dilihat kembali butir soal tersebut didapat perbandingan tipe butir soal ingatan, analisis dan aplikasi yaitu sebanyak 2:8:3. Dengan tipe ingatan pada butir soal nomor 33 dan 35, tipe analisis pada butir nomor 10, 16, 22, 24, 25, 29, dan 34, serta tipe aplikasi pada butir nomor 3, 4, dan 30.

V. KESIMPULAN

Berdasarkan hasil analisis soal ujian akhir semester genap mata pelajaran Fisika Kelas X SMA Negeri di Kabupaten Purworejo tahun pelajaran 2012/2013 yang telah dilakukan dapat ditarik kesimpulan besarnya indeks reliabilitas perangkat adalah 0,807 dengan fungsi pengecoh yang berfungsi adalah 30 atau 85,71% dan estimasi kesalahan pengukuran menggunakan pendekatan klasik sebesar 2,576. Tipe soal yang terkena bias gender pada butir soal menggunakan Mantel-Haenzel adalah tipe ingatan pada nomor 33 dan 35, tipe aplikasi pada nomor 3, 4 dan 30, serta tipe aplikasi pada nomor 10, 15, 16, 22, 24, 25, 29 dan 34.

UCAPAN TERIMAKASIH

Penulis mengucapkan terimakasih kepada Drs. H. Ashari sebagai *reviewer* jurnal ini dan Sriyono, M.Pd. sebagai pembimbing.

PUSTAKA

Buku:

- [1] Arifin, Zainal. 2011. *Penelitian Pendidikan*. Bandung: PT Remaja Rosdakarya.
- [2] Purwanto, Ngalm. 2009. *Prinsip-Prinsip dan Teknik Evaluasi Pengajaran*. Bandung: PT. Remaja Rosdakarya.
- [3] Allen, M. J., & Yen, W. M. (1979). *Introduction to measurement theory*. Belmont, CA: Wadsworth, Inc.
- [4] Azwar, Saifuddin. 2010. *Metode Penelitian*. Yogyakarta: Pustaka Pelajar.

Skripsi:

- [5] Triyatno, Nur. 2014. *Bias Gender Ujian Akhir Semester Genap Fisika Kelas X SMA Negeri Kabupaten Purworejo Tahun Pelajaran 2012/2013*. Universitas Muhammadiyah Purworejo. Purworejo.

