

ANALISIS KUALITAS DAN DETEKSI BIAS WILAYAH PADA BUTIR SOAL UAS BIOLOGI KELAS X DI KABUPATEN PATI

ITEM ANALYSIS AND BIAS OF REGION DETECTION IN BIOLOGY FINAL EXAM GRADE TEN AT PATI REGENCY

Oleh: Mahardika Himas Nugraeni¹⁾ (firdianrwahmawan@gmail.com)
Dr. Paidi, M.Si²⁾ (paidi@uny.ac.id), Budiwati, M.Si³⁾ (bwt_wati@yahoo.co.id)

Abstrak

Penelitian ini bertujuan untuk mengetahui: (1) kualitas butir soal UAS berdasarkan validitas, reliabilitas dan tingkat kesukaran butir soal, (2) ada tidaknya bias wilayah pada butir soal UAS. Penelitian ini termasuk penelitian deskriptif dengan metode sensus. Objek penelitian adalah butir soal UAS genap Biologi kelas X TA 2016/2017 yang digunakan oleh SMA negeri di Kabupaten Pati. Teknik pengumpulan data menggunakan wawancara dan dokumentasi. Teknik analisis data menggunakan pendekatan teori respon butir model satu parameter logistik dan metode *Lord's chi square*. Hasil penelitian menunjukkan: (1) 47 butir pilihan ganda dan 5 butir uraian cocok dengan model yang artinya butir soal valid, koefisien kappa butir pilihan ganda 0,45 yang artinya reliabel sedangkan koefisien kappa 5 butir uraian 0,22 artinya kurang reliabel, tingkat kesukaran butir soal terbagi menjadi 2 butir sangat sukar, 2 butir sangat mudah, 5 butir sukar, 5 butir mudah dan 34 butir sedang. (2) sejumlah 19 butir soal terindikasi bias dimana 7 butir menguntungkan SMA perkotaan sedangkan 12 butir menguntungkan SMA perdesan.

Kata Kunci: analisis kualitas butir soal, bias wilayah, ulangan akhir semester

Abstract

The aim of the research was to know: (1) the quality of the items based on validity, reliability, and item difficulty, (2) any bias of region detected in the items. The research was descriptive with census method. The object was the items of biology final exam grade X used by public senior high school in Pati. The techniques to collect data were interview and documentation. The techniques to analyze data were item response theory one parameter logistic model and Lord's chi square method. The results were (1) 47 multiple choice items and 5 essay items were fit with the model which meant final exam items were valid. Kappa coefficient of multiple choice items was 0.45 meant reliable meanwhile Kappa coefficient of essay items was 0.22 meant less reliable. The item difficulties classified into 2 items were really difficult, 2 items were really easy, 5 items were difficult, 5 items were easy and 34 items were medium. (2) there was 19 items detected as bias where 7 items advantageous to urban schools while 12 items advantageous to rural schools.

Keywords: bias of region, final exam, items analysis

PENDAHULUAN

Dalam kegiatan pembelajaran tidak lepas dari kegiatan pengukuran dan penilaian. Kegiatan pengukuran dapat dilaksanakan menggunakan teknik tes ataupun non tes. Teknik pengukuran yang sering digunakan guru

di sekolah adalah teknik penilaian jenis tes. Salah satu contohnya adalah Ulangan Akhir Semester.

Ulangan Akhir Semester (UAS) adalah salah satu bentuk pengukuran dalam rangka penilaian sumatif dengan teknik tes yang

dilakukan oleh lembaga sekolah. Tes sumatif menurut Arikunto (2008:33) adalah tes yang diberikan kepada peserta didik diakhir sekelompok program atau sebuah program pembelajaran besar seperti program satu semester. Manfaat tes sumatif antara lain untuk menentukan nilai peserta didik, untuk menentukan peserta didik dapat mengikuti program selanjutnya atau tidak, dan untuk mengisi catatan kemajuan peserta didik yang akan dilaporkan kepada orang tua peserta didik.

Soal UAS diharapkan memiliki kualitas yang baik dikarenakan soal akan digunakan untuk menentukan pencapaian belajar peserta didik selama satu semester. Kualitas butir soal dapat diketahui antara lain dengan melakukan analisis kualitas butir soal. Analisis kualitas butir soal menurut Arikunto (2008: 205) adalah prosedur sistematis yang akan memberikan informasi-informasi yang sangat khusus terhadap butir tes yang disusun. Grondlund (1981: 255) menyatakan manfaat analisis butir soal antara lain meningkatkan kualitas butir soal, data hasil analisis dapat menjadi dasar untuk diskusi yang efisien tentang hasil tes, hasil analisis dapat menjadi dasar untuk membuat kerja remedial, hasil analisis dapat digunakan untuk meningkatkan pembelajaran secara umum dan melakukan analisis akan meningkatkan kemampuan penyusunan soal.

Permendikbud no 23 Tahun 2016 tentang Standar Penilaian Pendidikan menyatakan bahwa Instrumen penilaian yang digunakan oleh satuan pendidikan dalam bentuk penilaian akhir dan/atau ujian

sekolah/madrasah memenuhi persyaratan substansi, konstruksi, dan bahasa, serta memiliki bukti validitas empirik. Salah satu tanda valid dan reliabelnya suatu tes yaitu tidak adanya bias. Bias menurut Osterlind dalam Retnawati (2014:125) memiliki makna yang tidak baik, rasial, menekan atau terlalu fanatik dengan objek yang akan diteliti. Jika suatu butir terindikasi bias maka butir tidak akan mampu mengukur kemampuan peserta didik yang sebenarnya sehingga hasil pengukuran menjadi tidak valid dan reliabel.

Soal UAS mata pelajaran Biologi kelas X tahun ajaran 2016/2017 di Kabupaten Pati disusun oleh Musyawarah Guru Mata Pelajaran (MGMP) Biologi Kabupaten Pati. Soal tersebut digunakan oleh hampir seluruh SMA yang ada di Kabupaten Pati, baik yang daerah perkotaan maupun perdesaan. Penggunaan satu jenis soal untuk seluruh wilayah ini memungkinkan adanya bias. Soal UAS yang disusun mungkin akan menguntungkan atau merugikan peserta didik dari salah satu wilayah. Misalnya dalam soal UAS tersebut memunculkan objek yang lebih dikenal atau familiar oleh salah satu kelompok peserta didik, atau menggunakan konteks soal yang hanya dipahami oleh sebagian kelompok peserta didik. Contohnya soal menggunakan objek seperti mall tertentu, hotel tertentu, atau jenis makanan yang hanya dijumpai di wilayah perkotaan. Contoh lainnya soal menggunakan konteks yang lebih dipahami oleh peserta didik dari wilayah perdesaan, seperti suasana atau kondisi danau alami, sawah alami, atau sungai alami. Selain hal diatas informasi mengenai kualitas butir soal UAS mata pelajaran Biologi kelas X tahun

ajaran 2016/2017 di Kabupaten Pati juga belum tersedia.

Tujuan penelitian ini adalah untuk mengetahui (1) kualitas butir soal berdasarkan validitas, reliabilitas dan tingkat kesukaran butir soal, (2) ada tidaknya bias wilayah pada butir soal UAS mata pelajaran Biologi kelas X tahun ajaran 2016/2017 di Kabupaten Pati.

METODE PENELITIAN

Jenis Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan jenis penelitian deskriptif.

Waktu dan Tempat Penelitian

Penelitian dilakukan mulai dari Februari sampai November 2017. Penelitian dilaksanakan di beberapa SMA negeri di Kabupaten Pati.

Target/Objek Penelitian

Objek penelitian ini adalah butir soal Ulangan Akhir Semester genap mata pelajaran Biologi kelas X tahun ajaran 2016/2017 buatan MGMP Biologi yang digunakan oleh SMA negeri di Kabupaten Pati.

Prosedur

Penelitian dimulai dengan observasi ke sekolah untuk mengetahui soal UAS yang dipakai apakah buatan guru atau MGMP. Kemudian pengambilan data dilakukan setelah UAS selesai di SMA yang sudah ditentukan.

Data, Instrumen, dan Teknik Pengumpulan Data

Data yang diperlukan untuk analisis pada penelitian ini adalah jawaban UAS peserta didik. Pemilihan SMA yang dijadikan sumber data berdasarkan proporsi SMA perkotaan:SMA perdesaan di Kabupaten Pati, yaitu 3:1. Alasan selanjutnya yaitu proporsi tersebut sudah memenuhi minimal jumlah sampel yang diperlukan yaitu antara 250-500(Le, 2006:4). Berdasarkan alasan tersebut SMA yang dijadikan sumber data adalah 3 SMA perkotaan yaitu SMA N 2 Pati, SMA N 3 Pati, SMA N 1 Juwana, dan 1 SMA perdesaan yaitu SMA N 1 Jakenan. Data diperoleh dengan teknik dokumentasi..

Teknik Analisis Data

Kegiatan analisis data dalam penelitian ini meliputi dua hal yaitu analisis kualitas soal dan pendeteksian bias wilayah atau *DIF*. Analisis kualitas soal menggunakan pendekatan teori respon butir, yaitu analisis dengan model satu parameter logistik atau *Rasch*, sedangkan pendeteksian bias wilayah atau *DIF* menggunakan metode *Lord's chi square* model satu parameter logistik.

HASIL PENELITIAN DAN PEMBAHASAN

Hasil analisis kualitas butir soal UAS adalah sebagai berikut:

a. Validitas butir soal

Validitas dalam penelitian ini dianalisis menggunakan program QUEST dengan model satu parameter logistik atau

Rasch. Pada program QUEST suatu butir atau testi dikatakan fit terhadap model jika nilai INFIT MNSQ berkisar dari 0,77 – 1,30 dengan alfa 5% (Subali & Suyata, 2011: 10). Hasil analisis dari 48 butir soal pilihan ganda adalah sebagai berikut:

Tabel 1. Kecocokan Butir Soal Pilihan Ganda Terhadap Model Satu Parameter Logistik (Rasch)

No	Nilai Infit MNS Q	Kecocokan terhadap model	Butir soal	Jml
1	0,77 – 1,30	Cocok	1,2,3,4,5,6,7,8, 9,10,11,12,13 14,15,15,17,18 ,19,20,21,22 24,25,27,28,29 ,30,31,32,33 34,35,36,37,38 ,39,40,41,43 44,45,46,47,48 ,49,50	47
2	<0,77 atau >1,30	Tidak cocok	42	1
Jumlah				48

Berdasarkan tabel di atas sejumlah 47 butir soal memiliki INFIT MNSQ berkisar 0,77-1,30 sedangkan terdapat 1 butir soal tidak cocok terhadap model, yaitu butir 42. Secara garis besar butir soal pilihan ganda UAS genap mata pelajaran Biologi kelas X tahun ajaran 2016/2017 Kabupaten Pati adalah valid.

Hasil analisis untuk kelima butir uraian adalah sebagai berikut:

Tabel 2. Kecocokan Butir Soal Uraian Terhadap Model Satu Parameter Logistik (Rasch)

No	Nilai Infit MNSQ	Kecocokan terhadap model	Butir soal	Jml
1	0,77 – 1,30	Cocok	1,2,3,4, 5	5
2	<0,77 atau >1,30	Tidak cocok	-	0
Jumlah				5

Berdasarkan tabel di atas terlihat semua butir soal uraian cocok dianalisis menggunakan model Rasch. Jadi butir soal uraian tergolong valid.

b. Relibilitas butir soal

Reliabilitas pada penelitian ini ditentukan menggunakan nilai taksiran koefisien kappa yang diperoleh dari konversi nilai reliabilitas menurut Kuder-Richardson-20 atau Alpha Cranbach. Penentuan koefisien Kappa memerlukan nilai z dan r . Nilai r adalah reliabilitas yang dihitung secara klasik yaitu dapat berupa reliabilitas Kuder-Richardson-20 atau Alpha Cranbach. Nilai z dihitung menggunakan rumus:

$$z = \frac{(c - 0,5 - M)}{S}$$

Nilai z dan r kemudian disesuaikan dengan ke tabel taksiran koefisien kappa (Subali dan Suyata, 2011: 11,41).

Pada soal pilihan ganda nilai $c=29$, $M=22,71$ dan $S=6,14$ sehingga nilai z yang dihitung berdasarkan rumus adalah 0,94. Nilai r soal pilihan ganda adalah 0,74. Nilai

$z=0,94$ dan $r=0,74$ sehingga nilai taksiran koefisien kapa soal pilihan ganda berdasarkan tabel adalah 0,45.

Pada soal uraian nilai $c=15$, $M=17,26$ dan $S=3,24$ sehingga nilai z adalah 0,85. Nilai r soal uraian adalah 0,36. Nilai $z=0,85$ dan $r=0,36$ sehingga nilai taksiran koefisien kapa dipilih yang mendekati yaitu 0,22.

Menurut Subali dan Suyata (2011:48) suatu tes yang diterapkan dalam suatu kelas pada periode penuh (misalkan 1 semester) dapat menggunakan nilai koefisien kapa sebesar 0,35 sampai 0,50. Koefisien kapa butir soal pilihan ganda sebesar 0,45 yang artinya butir soal memiliki keandalan atau reliabel. Sedangkan butir soal uraian memiliki reliabilitas yang kurang dikarenakan nilai koefisien kapa hanya berkisar di 0,22.

c. Tingkat kesukaran butir soal

Tingkat kesukaran pada penelitian ini dihitung berdasarkan model parameter 1 logistik menggunakan bantuan program QUEST. Pada program QUEST tingkat kesukaran pada skala dikotomus sama besarnya dengan nilai threshold. Hasil analisis tingkat kesukaran butir soal pilihan ganda adalah sebagai berikut:

No	Kategori	Butir soal	Jml
1	Sangat sukar ($b>2$)	48,27	2
2	Sukar ($1<b\leq 2$)	16,20,21,25,45	5
3	Sedang ($-1\leq b\leq 1$)	15,35,33,39,40,22,24,32 7,29,14,18,28,13,34,50,31 11,12,2,10,17,49,4,37,43,38 1,5,6,44,8,30,9	34
4	Mudah ($-1>b\geq -2$)	3,41,47,42,46	5
5	Sangat mudah ($b<-2$)	36,19	2
Jumlah Total			48

Pada hasil analisis diatas terlihat butir soal nomer 48 dan 27 memiliki nilai threshold mendekati 2 sehingga dikatakan sebagai *outlayer* yang artinya soal sangat sukar sampai jarang ada peserta didik yang dapat menjawab dengan benar. Butir soal nomer 36 dan 19 dengan nilai threshold mendekati -2 juga termasuk *outlayer*, soal sangat mudah sehingga tidak dapat membedakan kemampuan peserta didik. Soal-soal *outlayer* untuk selanjutnya perlu direvisi jika akan dipakai kembali. Jadi pada butir soal pilihan ganda terdapat 2 butir soal tergolong sangat sukar, 2 butir soal tergolong sangat mudah, 5 butir soal tergolong sukar, 34 tergolong sedang dan 5 butir soal tergolong mudah.

Hasil analisis butir soal uraian yang menggunakan bantuan program QUEST tingkat kesukaran butir soal ditunjukkan oleh nilai *DIFFICULTY*. Berdasarkan nilai *difficulty* kelima butir soal uraian termasuk

Tabel 3. Tingkat Kesukaran Butir Soal Pilihan Ganda

memiliki tingkat kesukaran sedang. Hasil tersebut disajikan dalam tabel berikut:

Tabel 4. Tingkat Kesukaran Butir Soal Uraian

No	Kategori	Butir soal	Jumlah
1	Sangat sukar ($b > 2$)	-	0
2	Sukar ($1 < b \leq 2$)	-	0
3	Sedang ($-1 \leq b \leq 1$)	1,2,3,4,5	5
4	Mudah ($-1 > b \geq -2$)	-	0
5	Sangat mudah ($b < -2$)	-	0
Jumlah Total			5

d. Pendeteksian bias wilayah

Pendeteksian bias butir atau *DIF* pada penelitian ini dilakukan dengan bantuan program R paket *difR* buatan David Magis dan kawan-kawan. Pada analisis ini derajat kebebasan yang digunakan adalah 1 karena menggunakan model parameter 1 logistik. Nilai *chi square* tabel pada alfa 5% dan $db=1$ adalah 3.8415. Hasil pendeteksian bias butir menggunakan metode Lord's *chi square* satu parameter logistik adalah 19 butir terindikasi bias wilayah atau *DIF*.

Wilayah atau *DIF*

No	Indikasi bias wilayah atau <i>DIF</i>	Butir	Jml
1	Terindikasi bias	2,13,15,19,20,21,22,25,30,32,33,34,37,41,42,44,47,49,50	19
2	Tidak terindikasi bias	1,3,4,5,6,7,8,9,10,11,12,14,16,17,18,24,27,28,29,31,35,36,38,39,40,43,45,46,48	29
Jumlah Total			48

Perhitungan untuk menentukan butir soal lebih menguntungkan atau merugikan kelompok mana dihitung berdasarkan presentase menjawab benar pada masing-masing butir yang bias pada tiap kelompok SMA. Hasil perhitungan adalah sebagai berikut:

Tabel 6. Distribusi Butir Soal yang Terindikasi Bias Wilayah

No	Kelompok wilayah yang diuntungkan	Butir soal	Jumlah
1	Perkotaan	13, 15, 20, 21, 25, 32, 41	7
2	Perdesaan	2, 19, 22, 30, 33, 34, 37, 42, 44, 47, 49, 50	12
Jumlah total			19

Berdasarkan tabel diatas terlihat bahwa lebih banyak butir soal yang menguntungkan peserta didik di SMA wilayah perdesaan. Jadi, secara umum butir soal UAS semester genap mata pelajaran Biologi kelas X tahun ajaran 2016/2017 di Kabupaten Pati sedikit lebih

Tabel 5. Data Butir Soal Terindikasi Bias

menguntungkan peserta didik di SMA wilayah perdesaan.

Terjadinya bias wilayah pada butir soal UAS dapat terjadi karena berbagai faktor. Kemungkinan yang menjadi faktor penyebab adanya bias wilayah antara SMA perkotaan dengan SMA perdesaan di Kabupaten Pati antara lain perbedaan sumber belajar yang tersedia dan perbedaan fasilitas dalam membantu proses belajar mengajar.

Menurut peraturan Kepala Badan Pusat Statistik nomor 37 tahun 2010 perbedaan antara wilayah perkotaan dan perdesaan antara lain dipengaruhi oleh jumlah pasar, pertokoan, bioskop, dan rumah sakit. Berbagai tempat tersebut dapat digunakan sebagai sumber belajar bagi peserta didik tergantung materi atau konsep apa yang akan dipelajari. Misalnya untuk peserta didik di SMA wilayah perkotaan dapat lebih mudah belajar secara faktual ke tempat seperti pasar, pertokoan atau rumah sakit. Sedangkan untuk peserta didik di SMA perdesaan akan lebih mudah menjangkau sumber belajar yang faktual seperti sawah, kebun atau sungai. Perbedaan ini mungkin ikut mempengaruhi adanya bias pada butir soal UAS yang digunakan di Kabupaten Pati. Masih diperlukan studi lebih lanjut mengenai sumber belajar faktual yang dipakai oleh SMA wilayah perkotaan dan perdesaan untuk mengetahui seberapa besar

pengaruhnya terhadap adanya bias pada soal UAS ini.

Fasilitas untuk membantu proses belajar mengajar tentu berbeda antara satu sekolah dengan sekolah lain. Perbedaan fasilitas belajar ini antara lain seperti kelengkapan buku pelajaran yang digunakan, kelengkapan alat peraga yang dimiliki, kelengkapan koleksi perpustakaan, kelengkapan alat dan bahan laboratorium, tersedianya kebun biologi ataupun *greenhouse*, tersedianya koneksi internet yang baik, kemudian tidak lupa kondisi ruang kelas ataupun gedung sekolah secara umum. Perbedaan fasilitas ini kemungkinan mempengaruhi adanya bias pada butir soal UAS Biologi kelas X di Kabupaten Pati. Untuk mendapatkan informasi mengenai perbedaan fasilitas yang dimiliki SMA di wilayah perdesaan atau SMA perkotaan di Kabupaten Pati diperlukan penelitian yang lebih lanjut mengenai hal tersebut.

SIMPULAN DAN SARAN

Simpulan

1. Kualitas butir soal adalah sebagai berikut:
 - a. Sejumlah 47 dari total 48 butir soal pilihan ganda fit dengan model Rasch. Seluruh butir uraian sejumlah 5 juga fit dengan model. Dengan demikian butir soal UAS tergolong valid.
 - b. Butir soal pilihan ganda memiliki koefisien kappa 0,45. Sedangkan butir

soal uraian dengan koefisien kappa 0,22. Dengan demikian reliabilitas butir pilihan ganda baik sedangkan butir soal uraian kurang baik.

- c. Sejumlah 48 butir soal pilihan ganda terbagi menjadi 2 butir soal termasuk kategori sangat sukar dan 2 butir soal termasuk kategori sangat mudah, 5 butir soal termasuk kategori sukar, 5 butir soal termasuk kategori mudah dan 34 butir soal termasuk kategori sedang. Tingkat kesukaran butir soal uraian yaitu kelima butir termasuk kategori sedang.
2. Pendeteksian bias wilayah menggunakan metode Lord's chi square pada soal pilihan ganda menunjukkan 19 butir soal terindikasi bias atau *DIF*. Soal yang terindikasi bias adalah butir nomor 2, 13, 15, 19, 20, 21, 25, 30, 32, 33, 37, 41, 42, 44, 47, 49, dan 50. Sejumlah 7 butir soal menguntungkan peserta didik SMA Perkotaan sedangkan 12 butir soal menguntungkan peserta didik SMA Perdesaan. Jadi butir soal UAS sedikit lebih menguntungkan peserta didik di SMA Negeri wilayah perdesaan.

Saran

Dalam penyusunan butir soal UAS berikutnya jika akan digunakan dalam skala kabupaten maka perlu lebih teliti dan berhati-hati. Butir soal yang kurang baik seperti butir soal yang outlayer, butir dengan reliabilitas kurang atau butir yang

terdeteksi bias perlu dilakukan perbaikan sebelum digunakan kembali atau dijadikan bank soal. Penyusunan butir soal UAS untuk menghindari adanya bias dapat dilakukan dengan beberapa cara diantaranya memahami berbagai variabel yang mungkin mempengaruhi jawaban peserta tes, menambah guru tim penyusun butir soal yang mewakili berbagai daerah sekolah sehingga lebih adil, lebih banyak saran dan koreksi yang masuk dikarenakan berdasarkan hasil wawancara diketahui bahwa kolaborasi dan kerjasama antar guru dalam tim penyusun butir soal UAS masih kurang.

Analisis kualitas butir soal pada suatu tes alangkah baiknya jika dianalisis secara kesatuan antara butir pilihan ganda dengan butir uraian sehingga hasil analisis akan lebih menggambarkan kualitas butir soal tes secara kesatuan.

Hasil analisis kualitas butir soal menggunakan model satu parameter logistik dirasa belum mampu menggambarkan semua faktor yang terlibat dalam menentukan kualitas suatu butir soal sehingga analisis kualitas butir soal menggunakan 2 atau 3 parameter logistik disarankan untuk penelitian selanjutnya.

Penelitian ini belum mampu menjelaskan penyebab terjadinya bias wilayah pada soal UAS Biologi buatan MGMP Kabupaten Pati oleh sebab itu disarankan penelitian lebih lanjut untuk

mengetahui penyebab terjadinya bias wilayah pada soal UAS semester genap mata pelajaran Biologi kelas X tahun ajaran 2016/2017 di Kabupaten Pati.

DAFTAR PUSTAKA

- An, X. & Yung, Y-F. (2014). Item Response Theory: What it is and How You Can Use the IRT Procedure to Apply it. *Artikel Institusi*. Diambil pada 10 Juli 2017, dari <http://support.sas.com/SAS364-2014>
- Arifin, Z. (2011). *Evaluasi Pembelajaran. Edisi Revisi*. Jakarta: Direktorat Jenderal Pendidikan Islam Kementerian Agama RI.
- Arikunto, S. (2008). *Dasar-dasar Evaluasi Pendidikan. Edisi revisi*. Jakarta: Bumi Aksara.
- Badan Pusat Statistik. (2010). *Peraturan Kepala Badan Pusat Statistik Nomor 37, Tahun 2010, tentang Klasifikasi Perkotaan dan Perdesaan di Indonesia*.
- Gronlund, N.E. (1981). *Measurement and Evaluation in Teaching. 4th*. New York: Collier Macmillan International.
- Kusaeri & Suprananto. (2012). *Pengukuran dan Penelitian Pendidikan*. Yogyakarta: Graha Ilmu.
- Le, Luc. (April 2006). *Analysis of Differential Item Functioning*. Makalah disajikan dalam Annual Meetings of the American Educational Research Association di San Francisco.
- Magis, D. (2010). A General Framework And an R Package For The Detection of Dichotomous DIF. *Behavior Research Methods*, 42(3), 847-862.
- Magno, C. (2009). Demonstrating the Difference between Classical Test Theory and Item Response Theory Using Derived Test Data. *The International Journal of Educational and Psychological Assessment*, Vol.1 Issue 1, pp.1-11.
- Retnawati, H. (2014). *Teori Respon Butir Dan Penerapannya: Untuk peneliti, Praktisi Pengukuran dan Pengujian, Mahasiswa Pascasarjana*. Yogyakarta: Nuha Medika.
- Retnawati, H & Hidayati, K. (2006). Pendeteksian Bias Tes Dan Butir Perangkat Soal Matematika Ujian Nasional SLTP Berdasarkan Teori Respon Butir. *Artikel Institusi*. Diambil pada 30 Januari 2017, dari <http://staffnew.uny.ac.id/upload/132296141>
- Subali, B. (2012). *Prinsip Asesmen & Evaluasi Pembelajaran*. Yogyakarta: UNY Press.
- Subali, B. & Suyata, P. (2011). *Panduan Analisis Data Pengukuran Pendidikan Untuk Memperoleh Bukti Empirik Kesahihan Menggunakan Program Quest*. Yogyakarta: Lembaga Penelitian dan Pengabdian kepada Masyarakat UNY.
- Sudaryono. (2012). Kajian Metode Deteksi Differential Item Functioning (DIF) Butir Soal Ujian Nasional Dengan Teori Tes Klasik. *Jurnal Pendidikan dan Kebudayaan*, Vol 18 No 2. Diambil pada 30 Januari 2017, dari <http://jurnaldikbud.kemdikbud.go.id>
- Sudijono, A. (2010). *Pengantar Evaluasi Pendidikan*. Jakarta: PT. Raja Grafindo Persada.
- Venables, W.N., Smith, D. M., & the R Core Team. (2017). An Introduction to R. Diambil pada 3 Mei 2017, dari <http://cran.r-project.org/doc/R-intro>
- Wiberg, M. (2007). Measuring And Detecting DIF In Criterion-Referenced Licensing Test: A theoretic Comparison of Methods. *Artikel Institusi*. Diambil pada 11 Juli 2017, dari http://edusci.umu.se/59534_em-no-60