

LAPORAN
PENELITIAN KOMPETISI FKIP UNMUL



PERBANDINGAN KUALITAS INSTRUMEN TES PILIHAN GANDA
MATEMATIKA SMP HASIL ANALISIS *BILOG-MG* DAN *QUEST*

Tim Penelitian:

Ketua

DR. SUGENG, M.PD.

NIP. 19581005 198503 1 003

NIDN. 0005105809

Dosen Pendidikan Matematika

Anggota

PRATIWI KURNIA SARI

NIM. 1405045014

Mahasiswa Pendidikan Matematika

KEMENTERIAN RISET, TEKNOLOGI, DAN PERGURUAN TINGGI
PROGRAM STUDI PENDIDIKAN MATEMATIKA
FAKULTAS KEGURUAN DAN ILMU PENDIDIKAN
UNIVERSITAS MULAWARMAN
OKTOBER 2019

**LEMBAR PENGESAHAN
PENELITIAN FKIP TAHUN 2019**

1. Judul penelitian: **Perbandingan Kualitas Instrumen Tes Pilihan Ganda
Matematika SMP Hasil Analisis BILOG-MG dan QUEST**

Rumpun Ilmu Penelitian: **Evaluasi Pendidikan Matematika**

2. Ketua Peneliti

- a. Nama : **Dr. Sugeng, M.Pd.**
b. NIP : 19581005 198503 1 003
c. NIDN : 0005105809
d. Pangkat/Gol. : Pembina Tk. I / IV-b
e. Jabatan Fungsional : Lektor Kepala
f. Fak/Jurusan/Prodi : FKIP Unmul/PMIPA/Pendidikan Matematika
g. Alamat Institusi : Jln.Muara Pahu, Kampus Gunung Kelua,
Samarinda 75119
h. Telpon/Fax/E-mail : 08156898316 / sugeng_ppg@yahoo.com

3. Anggota Mahasiswa,

- a. Nama : **Pratiwi Kurnia Sari**
b. NIM : 1405045014
c. Program Studi : Pendidikan Matematika FKIP Unmul

4. Biaya Penelitian : Penelitian ini dibiayai oleh DIPA FKIP UNMUL 2019
sebesar Rp.6.000.000,-



Prof. Dr. H. Han. Amir Masruhimi, M.Kes.
NIP. 196010271985031003

Samarinda, 24 Oktober 2019

Ketua Peneliti,

Dr. Sugeng, M.Pd.
NIP. 195810051985031003

PERBANDINGAN KUALITAS INSTRUMEN TES PILIHAN GANDA MATEMATIKA SMP HASIL ANALISIS BILOG-MG DAN QUEST

Oleh

Dr. Sugeng, M.Pd¹⁾

Pratiwi Kurnia Sari²⁾

¹⁾Dosen Pend. Matematika FKIP Universitas Mulawarman

²⁾Mahasiswa Pend. Matematika FKIP Universitas Mulawarman

RINGKASAN

Keakuratan seseorang dalam mengambil suatu keputusan sangat ditentukan oleh keakuratan data/informasi yang diperolehnya melalui pengukuran. Kondisi ini ditentukan oleh kualitas suatu instrumen pengukuran, terutama di bidang pembelajaran matematika, guna mendukung usaha peningkatan kualitas pendidikan.

Instrumen hasil belajar matematika ini berupa tes *Multiple Choise*. Kualitas instrumen tersebut dilakukan secara kuantitatif dan kualitatif. Pengujian kualitas instrumen secara kuantitatif dilakukan melalui bantuan program aplikasi *BILOG-MG* dan *QUEST*. Penggunaan keduanya bertujuan untuk membandingkan tingkat keakuratan hasil analisis butir dari masing-masing program tersebut sehingga diperoleh informasi atau data yang diharapkan. Hasil penelitian menunjukkan bahwa terdapat perbedaan hasil analisis butir terkait dengan parameter Tingkat Kesukaran. Program aplikasi *BILOG-MG* menghasilkan 80% butir soal memnuhi kriteria dan *QUEST* menghasilkan 100% memenuhi kriteria. Untuk parameter Daya pembeda keduanya menunjukkan hasil yang sama, yaitu 100% memenuhi kriteria. Hasil analisis *QUEST* juga memperlihatkan hasil Item Fit dan Case Fit yang memenuhi kriteria. Sebagai bentuk luaran dari hasil penelitian ini adalah terpublikasikannya dalam jurnal ilmiah atau prosiding jenjang nasional atau internasional.

KATA PENGANTAR

Puji syukur kami panjatkan ke hadirat Allah SWT atas rahmat-NYA karena penyusunan Laporan Pengabdian kepada Masyarakat yang berjudul “Perbandingan Kualitas Instrumen Tes Pilihan Ganda Matematika SMP Hasil Analisis BILOG-MG dan QUEST” ini dapat diselesaikan. Laporan penelitian ini mengungkap perbedaan hasil analisis butir soal dikotomus (Pilihan Ganda) menggunakan program aplikasi BILOG-MG dan QUEST, terutama secara kuantitatif.

Kami menghaturkan terima kasih sebesar-besarnya kepada yang terhormat:

1. Dekan Fakultas Keguruan dan Ilmu Pendidikan beserta Staf, atas pemberian kesempatan dan kebijakannya sehingga penulisan Laporan Penelitian kepada Masyarakat ini dapat penulis selesaikan.
2. Kepala Diknas Tingkat II Kota Samarinda atas kerjasamanya dan telah mengizinkan guru-guru di wilayah Samarinda untuk mengikuti kegiatan Penelitian.
3. Kepala SD Muhammadiyah 2 Samarinda Ulu yang telah memberikan tempat pelaksanaan kegiatan Pengabdian kepada Masyarakat ini.
4. Pihak lain yang telah memberikan bantuan (materiil maupun moril) kepada penulis sehingga penulisan Laporan Penelitian ini dapat diselesaikan.

Semoga Allah swt melimpahkan rahmat dan balasan berlipat ganda kepada pihak-pihak yang telah membantu dalam penulisan Laporan Penelitian ini. Selain itu, semoga hasil Penelitian ini bermanfaat bagi pengembangan pengukuran bagi guru.

Samarinda, 25 September 2019

Sugeng

Pratiwi Kurnia Sari

DAFTAR ISI

HALAMAN SAMPUL.....	i
HALAMAN PENGESAHAN.....	ii
RINGKASAN	iii
KATA PENGANTAR.....	iv
DAFTAR ISI.....	v
BAB I. PENDAHULUAN	
A. Latar Belakang Masalah.....	1
B. Perumusan Masalah Penelitian.....	3
C. Tujuan Penelitian.....	3
D. Manfaat Penelitian.....	3
E. Luaran	4
BAB II. TINJAUAN PUSTAKA	
A. Kajian Teoretis	5
1. Pengukuran	5
2. Teori Respons Butir	7
3. Program Aplikasi Pengukuran	9
a. BILOG-MG	9
b. QUEST	10
B. Estimasi	11
1. Estimasi Item	11
2. Estimasi Testi (Person)	11
3. Pengujian Validitas	11
4. Estimasi Reliabilitas	12
C. Instruman Tes Matematika	13
D. Kualitas Instrumen Tes Matematika	14
1. Secara Kuantitatif	14
2. Secara Kualitatif	15
BAB III. METODE PENELITIAN	
A. Waktu dan Tempat	17

B. Sampel dan Populasi.....	17
C. Teknik Pengumpulan Data	17
1. Instrumen Tes Matematika.....	17
2. Kisi-kisi Instrumen.....	18
D. Teknik Analisis Data	19
1. Analisis Data Kuantitatif.....	19
2. Analisis Data Kualitatif.....	19
BAB IV. METODE PENELITIAN	
A. Hasil Penelitian.....	21
1. Hasil Analisis Butir dengan BILOG-MG	21
2. Hasil Analisis Butir dengan BILOG-MG Untuk Teori Tes Klasik.....	27
3. Hasil Analisis Butir dengan QUEST.....	30
a. Estimasi untuk Item (Item Fit) Klasik.....	31
B. Hasil Penelitian.....	42.
BAB V. KESIMPULAN DAN SARAN	
1. Kesimpulan	51
2. Saran.....	52
DAFTAR PUSTAKA	53
LAMPIRAN.....	57

BAB I

PENDAHULUAN

A. Latar Belakang Masalah

Pengukuran merupakan sarana terpenting dalam pelaksanaan kegiatan evaluasi program pembelajaran. Dalam praktik pengukuran, senantiasa menggunakan instrumen atau alat ukur yang bersesuaian dengan keperluan evaluasi. Penggunaan instrumen dalam kegiatan pengukuran bertujuan untuk memperoleh informasi sebagaimana diharapkan. Informasi yang diperoleh dari pengukuran tersebut merupakan data yang sangat bermanfaat dalam kegiatan penilaian ataupun pengambilan keputusan tentang program yang dievaluasi. Berdasarkan data yang diperoleh tersebut memberikan arah kepada akurat atau tidaknya keputusan yang dihasilkannya. Oleh karena itu, dalam kegiatan pengukuran, salah satu tujuan utamanya adalah memperoleh informasi atau data yang berkualitas atau akurat.

Instrumen yang digunakan dalam kegiatan pengukuran untuk memperoleh sekumpulan data atau informasi dapat berupa tes dan non-tes (Djemari Mardapi, 2008; Sumadi Suryabrata, 2000). Informasi yang diperoleh dengan instrumen tes berbeda dengan informasi yang diperoleh menggunakan non-tes. Menurut Sumadi Suryabrata (2000), penggunaan sarana instrumen tes cenderung untuk mengukur hasil kegiatan kognitif. Kondisi mengisyaratkan bahwa penggunaan instrumen non-tes lebih mengarah kepada pengumpulan data selain kognitif. Pada sisi lain, Mehrens & Lehmann (1978) mengungkapkan bahwa antara *test*, *measurement*, dan *evaluation* merupakan tiga aspek yang saling berkaitan dalam praktik evaluasi. Kondisi ini menunjukkan bahwa kegiatan pengukuran merupakan salah satu kegiatan utama dalam melakukan evaluasi hasil pembelajaran siswa. Selain hasil pembelajaran, informasi lain yang diperoleh, antara lain ketersediaan dan kelayakan

BAB V.

KESIMPULAN DAN SARAN

1. Kesimpulan

Berdasarkan hasil analisis butir soal menggunakan program BILOG-MG, dan dengan program aplikasi QUEST, terutama untuk parameter Tingkat Kesukaran dan Daya Pembeda butir soal jenis Pilihan Ganda, disimpulkan bahwa:

- a. Hasil analisis butir soal matematika dengan menggunakan program aplikasi BILOG-MG diketahui bahwa untuk (a) parameter Tingkat Kesukaran butir terdapat 4 butir soal matematika (atau 20%) berada di luar kategori yang sebagai butir soal yang diijinkan; dan selebihnya yaitu 15 butir soal (atau 75%) dengan kategori Sedang, terdapat 1 butir soal (atau 5%) dengan kategori Sukar.
- b. Hasil analisis dengan program aplikasi BILOG-MG untuk parameter Daya pembeda, diperoleh bahwa rentangan nilai antara 0,248 sampai dengan 0,888, atau $0,248 \leq a_i \leq 0,888$. Berarti, berarti semua butir soal memiliki indeks diskriminan yang memenuhi persyaratan sebagai butir yang layak untuk digunakan lebih lanjut.
- c. Hasil pengujian terhadap Item Fit dengan program aplikasi QUEST, menunjukkan bahwa untuk semua butir soal (nomor 1 sampai dengan nomor 20) memiliki nilai *INFIT MNSQ* dalam rentang 0.83 sampai dengan 1.06, atau $0.83 \leq \text{INFIT MNSQ} \leq 1.06$. Artinya, seluruh butir soal berada dalam kategori fit terhadap model pengukuran.
- d. Hasil pengujian *Case Fit* menunjukkan bahwa koefisien estimasi person berada interval $-1.93 \leq \text{Case Fit} \leq 0.72$. Harga estimasi tersebut masih berada di dalam rentangan antara -2.0 sampai dengan +2.0; atau $-2,0 \leq \text{case fit} \leq +2,0$. Oleh

DAFTAR PUSTAKA

- Adams, R. J., & Khoo, S. -T. (1996). *QUEST: The interactive test analysis system*. Camberwell, VA: ACER Press.
- Adams R. J., & Khoo S-T. (1998). QUEST - Version 2.1. The interactive test analysis system. *Rasch Measurement Transactions* 11:4 p. 598.
- Bambang Subali & Pujiyati Suyata. (2011). *Panduan Analisis Data Pengukuran Pendidikan Untuk Memperoleh Bukti Empirik Kesahihan Menggunakan Program QUEST*. Yogyakarta: Lembaga Penelitian dan Pengabdian pada Masyarakat, Universitas Negeri Yogyakarta.
- Bond, T.G. & Fox, Ch.M. (2007). *Applying the rasch model: Fundamental measurement in the human sciences*. 2-nd ed. Mahwah, New Jersey: Lawrence Erlbaum Associates, Publishers.
- Caroll, J. B. (1979). Measurement of abilities constructs. Dalam *Construct Validity in Psychological Measurement* (pp. 23-41). Proceedings of a Colloquium on Theory and Application in Education and Employment. Princeton, NJ: ETS.
- Center for Educational Assessment Research Report No. 652. Amherst, MA: University of Massachusetts, Center for Educational Assessment.
- DeMars, C. (2010). *Item response theory*. New York: OXFORD University Press.
- Djemari Mardapi. (1994). Analisis butir dengan teori tes klasik dan teori tes respons butir. *Laporan Penelitian*. Yogyakarta: IKIP Yogyakarta.
- Djemari Mardapi. (1999). *Pengukuran, penilaian, dan evaluasi*. Makalah disajikan pada Penataran Evaluasi Pembelajaran Matematika SLTP untuk Guru Inti Matematika di MGMP SLTP Tanggal 8 Nopember sampai 23 Nopember 1999, di Pusat Pendidikan dan Penataran Guru Matematika Yogyakarta.
- Djemari Mardapi. (2008). *Teknik penyusunan instrumen tes dan nontes*. Yogyakarta: Mitra Cendikia Press.
- Embretson, S. E. & Reise, S. P. (2000). *Item response theory for psychologist*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Garver, M. S. & Mentzer, J. T. (1999). Logistics research methods: Employing structural equation modeling to test for construct validity. *Journal of Business*, 20(1), 33-57.