

**LAPORAN PENELITIAN
PENELITIAN DAN EVALUASI PENDIDIKAN
DANA DIPA PPS UNY
TAHUN ANGGARAN 2018**



**PARAREL TES DILIHAT DARI URUTAN BUTIR DAN
PENEMPATAN JAWABAN**

Diusulkan oleh:

Prof. Dr. Badrun Kartowagiran, M.Pd./NIP. 19530725 197811 1 001

Dr. Harun, M.Pd./NIP. 19560727 198503 1 024

Prof. Djemari Mardapi, Ph.D/ NIDK.11702470101624

Kriswantoro /NIM. 17701261018

Dian Normalitasari Purnama/NIM. 17701261020

**PROGRAM PASCASARJANA
UNIVERSITAS NEGERI YOGYAKARTA
TAHUN 2018**

HALAMAN PENGESAHAN

Judul : PARAREL TES DILIHAT DARI URUTAN DAN PENEMPATAN JAWABAN

Peneliti/Pelaksana
Nama lengkap : Prof. Dr. Badrun Kartowagiran, M.Pd.
Perguruan Tinggi : Universitas Negeri Yogyakarta
NIDN : 0025075303
Jabatan Fungsional : Guru Besar
Program Studi : Penelitian dan Evaluasi Pendidikan - S3
Nomor HP : +628122781548
Alamat surel (e-mail) : kartowagiran@uny.ac.id

Anggota (1)
Nama Lengkap : Dr. Harun, M.Pd.
NIDN : 0027075610
Perguruan Tinggi : Universitas Negeri Yogyakarta


Anggota (2)
Nama Lengkap : Prof. Djemari Mardapi, M.Pd.,Ph.D.
NIDN : 8881830017
Perguruan Tinggi : Universitas Negeri Yogyakarta

Institusi Mitra (jika ada)
Nama Institusi Mitra :
Alamat Institusi Mitra :
Penanggung Jawab :
Tahun Pelaksanaan :
Biaya Tahun Berjalan : Rp. 20.000.000,00

Mengetahui,
Kaprodi PEP,


Prof. Dr. Badrun Kartowagiran, M.Pd.
NIP 19530725 197811 1 001

Yogyakarta, 29 Oktober 2018
Ketua Pelaksana


Prof. Dr. Badrun Kartowagiran, M.Pd.
NIP 19530725 197811 1 001

Menyetujui,
Direktur PPs,


Prof. Dr. Marsigit, M.A.
NIP 19570749 198303 1 004



Parallel Tes Dilihat dari Urutan Butir dan Penempatan Jawaban

**Badrun Kartowagiran
Harun
Djemari Mardapi**

Abstrak

Selama ini untuk memperoleh tes yang paralel dilakukan dengan cara mengubah nomor urut dan atau mengubah urutan pilihan jawaban. Oleh karena itu penelitian ini dilakukan untuk membuktikan bahwa dengan cara mengubah urutan nomor butir soal dan atau urutan pilihan jawaban akan diperoleh tes yang paralel.

Penelitian ini menggunakan metode eksperimen yakni dengan dilakukan ujicoba sebelum dan sesudah perubahan nomor butir tes dan atau pilihan jawaban yang diujikan pada siswa yang sama. Populasi penelitian adalah siswa SMP di Kota Yogyakarta. Sampel penelitian ditentukan menggunakan teknik *stratified proporsional random sampling*. Sekolah yang ditetapkan menjadi sampel adalah SMP N 5 dan SMPN 8 Yogyakarta. Instrumen yang digunakan adalah perangkat tes matematika berbentuk tes objektif dengan 4 pilihan jawaban. Instrumen terdiri dari 5 paket soal yakni paket 1 (soal tanpa pengacakan), paket 2 (pengacakan nomor butir 1-40), paket 3 (pengacakan nomor butir 1-20 dan 21-40), paket 4 (pengacakan nomor 1-10, 11-20, 21-30 dan 31-40), dan paket 5 (pengacakan penempatan pilihan jawaban). Data yang telah didapatkan dianalisis menggunakan teori tes klasik (*quest*) dan teori respon butir (*quest*). Uji beda dilakukan dengan bantuan program SPSS.

Hasil penelitian menunjukkan bahwa tidak ada pengaruh pengacakan urutan nomor butir soal dan penempatan pilihan jawaban terhadap tingkat kesukaran butir. Hasil analisis menggunakan analisis varians menunjukkan bahwa tidak ada perbedaan yang signifikan tingkat kesukaran butir antara paket 1, paket 2, paket 3, dan paket 4. Hasil analisis juga menunjukkan bahwa tidak ada perbedaan yang signifikan tingkat kesukaran butir antara paket 1 dan paket 5.

Kata Kunci: *tes paralel, urutan butir, penempatan jawaban*

DAFTAR ISI

	Halaman
HALAMAN JUDUL.....	i
HALAMAN PENGESAHAN.....	ii
ABSTRAK.....	iii
KATA PENGANTAR.....	iv
DAFTAR ISI	v
DAFTAR TABEL.....	vi
BAB I. PENDAHULUAN.....	1
BAB II. TINJAUAN PUSTAKA.....	4
BAB III. METODE PENELITIAN	27
BAB IV. HASIL PENELITIAN DAN PEMBAHASAN.....	30
BAB V. KESIMPULAN DAN SARAN	48
DAFTAR PUSTAKA.....	49
LAMPIRAN.....	51

DAFTAR TABEL

Tabel 1.	Range Tingkat Kesukaran	18
Tabel 2.	Kategori Penerimaan Daya Beda Butir Soal	19
Tabel 3.	Paket Soal dan Jenis Pengacakan	28
Tabel 4.	Hasil Analisis Teori Tes Klasik Soal Paket 1	30
Tabel 5.	Hasil Analisis Teori Tes Klasik Soal Paket 2	31
Tabel 6.	Hasil Analisis Teori Tes Klasik Soal Paket 3	32
Tabel 7.	Hasil Analisis Teori Tes Klasik Soal Paket 4	33
Tabel 8.	Hasil Analisis Teori Tes Klasik Soal Paket 5	34
Tabel 9.	Tingkat Kesukaran 5 Paket Soal Berdasarkan Teori Tes Klasik	35
Tabel 10.	Hasil Analisis Deskriptif Berdasarkan SPSS	36
Tabel 11.	Uji Beda 5 Paket Soal Menggunakan Teori Tes Klasik	36
Tabel 12.	Hasil Analisis Uji Beda Tiap Paket Soal	37
Tabel 13.	Hasil Analisis Teori Respons Butir Paket 1	38
Tabel 14.	Hasil Analisis Teori Respons Butir Paket 2	39
Tabel 15.	Hasil Analisis Teori Respons Butir Paket 3	40
Tabel 16.	Hasil Analisis Teori Respons Butir Paket 4	41
Tabel 17.	Hasil Analisis Teori Respons Butir Paket 5	42
Tabel 18.	Tingkat Kesukaran 5 Paket Soal Berdasarkan Teori Respon Butir	43
Tabel 19.	Hasil Analisis Deskriptif Berdasarkan SPSS	44
Tabel 20.	Uji Beda 5 Paket Soal Menggunakan Teori Respon Butir	44
Tabel 21.	Hasil Analisis Uji Beda Tiap Paket Soal	45

BAB I

PENDAHULUAN

A. Latar Belakang Masalah

Pendidikan mempunyai peranan yang penting dalam meningkatkan kualitas hidup masyarakat. Dalam rangka mewujudkan pendidikan yang berkualitas sesuai dengan tuntutan masyarakat di era global saat ini diperlukan sumber daya manusia yang berkualitas. Sumber daya manusia yang berkualitas adalah insan yang berilmu pengetahuan, berketerampilan, berbudi pekerti luhur, berakhlak mulia, bertanggungjawab dan berupaya mencapai kesejahteraan diri serta memberikan sumbangan terhadap keharmonisan dan kemakmuran keluarga, masyarakat, dan negara (BSNP, 2006: 205). Sumber daya manusia yang berkualitas dapat diperoleh dari pendidikan yang berkualitas.

Ujian Nasional (UN) merupakan salah satu upaya pemerintah dalam rangka memacu peningkatan mutu atau kualitas pendidikan. Ujian Nasional selain berfungsi untuk mengukur dan menilai pencapaian kompetensi lulusan dalam mata pelajaran tertentu, serta pemetaan mutu pendidikan pada tingkat pendidikan dasar dan menengah, juga berfungsi sebagai motivator bagi pihak-pihak terkait untuk bekerja lebih baik guna mencapai hasil ujian yang baik (Puspendik, 2014: 1). Sistem penilaian dan kualitas pembelajaran adalah dua hal yang saling terkait. Sistem pembelajaran yang baik akan menghasilkan kualitas belajar yang baik (Djemari Mardapi, 2012: 12). Selanjutnya kualitas pembelajaran ini dapat dilihat dari hasil penilaian yang dilakukan oleh guru atau pendidik. Guru adalah pendidik profesional dengan tugas utama mendidik, mengajar, membimbing, mengarahkan, melatih, menilai, dan mengevaluasi siswa pada pendidikan anak usia dini jalur pendidikan formal, pendidikan dasar, dan pendidikan menengah (Undang-undang, 2005).

Berdasarkan UU No. 14 Tahun 2005 tentang Guru dan Dosen di atas, dapat diketahui bahwa peran pendidik tidak hanya merancang dan melaksanakan pembelajaran saja, melainkan juga melakukan asesmen atau penilaian. Penilaian hasil belajar siswa yang dilakukan oleh pendidik hendaknya dilakukan secara berkesinambungan. Penilaian tidak hanya dimaksudkan untuk mengetahui tercapai tidaknya tujuan-tujuan yang telah ditetapkan, tetapi juga untuk mengetahui apakah

tujuan-tujuan tersebut penting bagi siswa dan bagaimana siswa mencapainya. Hasil kajian menunjukkan bahwa sebanyak 87 persen guru masih kesulitan dalam memahami cara melakukan penilaian (Rusilowati, 2013). Tidak tersosialisasikannya langkah-langkah dalam melakukan penilaian menjadi salah satu kendalanya. Hal ini mengindikasikan bahwa kemampuan guru untuk melakukan penilaian secara baik masih perlu ditingkatkan.

Penilaian yang baik dapat dilakukan dengan cara mengumpulkan bukti akurat terkait pencapaian hasil belajar siswa dan menjadikan proses penilaian kelas dan hasilnya bermanfaat bagi siswa, yaitu mampu meningkatkan motivasi dan prestasi belajarnya (Stiggins & Chappuis, 2012: 3). Dengan demikian, penilaian hasil belajar harus dapat menilai kemajuan belajar siswa. Informasi kemajuan belajar siswa ini dapat digunakan sebagai dasar untuk mengambil suatu keputusan tentang status siswa dalam kelompoknya serta menentukan langkah-langkah selanjutnya. Oleh karena itu dalam melakukan penilaian hasil belajar, perlu dirancang langkah-langkahnya secara rinci agar dapat memotret kompetensi siswa secara tepat.

Salah satu memotret kompetensi siswa yakni dengan melakukan tes. Hasil tes merupakan informasi tentang karakteristik individu atau sebuah kelompok (Harun Rasyid & Mansur, 2008: 11). Salah satu syarat tes yang baik adalah reliabel atau dengan kata lain instrumen tes yang baik adalah instrumen yang dapat menghasilkan skor/jawaban yang konsisten meskipun digunakan pada orang yang berbeda dan pada waktu yang berbeda tetapi kondisinya sama. Terdapat beberapa cara dalam mengetahui reliabilitas suatu tes, salah satu caranya yakni dengan menggunakan tes paralel.

Tes paralel adalah dua buah tes atau lebih yang mempunyai kesamaan tujuan, tingkat kesukaran, dan susunan, tetapi butir-butir soalnya berbeda. Selama ini untuk memperoleh tes yang paralel dilakukan dengan cara mengubah nomor urut dan atau mengubah urutan pilihan jawaban dari satu paket soal menjadi beberapa paket soal. Hal ini masih menjadi anggapan saja untuk sebagian besar pendidik, sedangkan kenyataannya belum ada penelitian yang menunjukkan bahwa strategi tersebut memang benar, yakni dengan cara mengubah nomor urut dan atau urutan pilihan jawaban akan diperoleh tes yang paralel. Oleh karena itu penelitian ini dilakukan

untuk mengetahui bagaimana pengaruh pengacakan urutan nomor butir soal dan atau urutan pilihan jawaban terhadap tingkat kesulitan.

B. Tujuan Penelitian

Tujuan penelitian ini adalah untuk mengetahui: pengaruh pengacakan urutan nomor soal dan pilihan jawaban terhadap tingkat kesukaran butir.

C. Manfaat Penelitian

Manfaat penelitian ini adalah mengetahui apaitu tes parallel dan dampak dari perubahan peletakan urutan nomor dan penempatan pilihan jawaban. Hal ini akan memudahkan guru dalam mengetahui kemampuan siswa setelah mengerjakan soal yang telah dirubah peletakan urutan nomor dan penempatan jawaban, adakah perbedaannya atau tidak.

BAB II

TINJAUAN PUSTAKA

A. Kajian Teori

1. Tes

Pengukuran dilakukan dengan bantuan alat yaitu perangkat tes (Cohen & Swerdlik, 2009: 5). Allen & Yen (1979: 1) menyebutkan bahwa tes sebagai instrumen atau prosedur sistematis untuk mengukur perilaku sampel. Tes adalah suatu prosedur yang sistematis untuk mengamati dan menggambarkan satu atau lebih karakteristik seseorang dengan bantuan skala *numeric* atau sistem kategorik. Sejalan dengan pengertian di atas, Miller, Linn & Grondlund (2009: 28) mendefinisikan tes sebagai tipe tertentu dari penilaian yang diberikan selama periode waktu tertentu dalam kondisi cukup sebanding untuk semua siswa. Tes biasanya terdiri dari satu set pertanyaan. Tujuan tes adalah untuk menjawab pertanyaan seberapa baik individu melakukan sesuatu, baik dibandingkan dengan orang lain atau dibandingkan dengan domain tugas kinerja.

Berdasarkan uraian di atas dapat dikatakan bahwa tes adalah alat ukur yang berisi berbagai pertanyaan, pernyataan atau serangkaian tugas yang harus dikerjakan peserta didik untuk mengukur pencapaian hasil belajar seseorang. Tes dilakukan dengan prosedur yang sistematis. Tes juga menghasilkan informasi tentang karakteristik seseorang atau sekelompok orang. Jika ditinjau dari tujuannya, ada 4 macam tes yang banyak digunakan dalam dunia pendidikan, yaitu tes penempatan, tes diagnosik, tes formatif dan tes sumatif (Djemari Mardapi, 2008: 68). Tes penempatan dilaksanakan pada awal pelajaran. Tes diagnosik berguna untuk mengetahui kesulitan belajar yang dihadapi peserta didik. Tes formatif bertujuan untuk memperoleh masukan tentang tingkat keberhasilan pelaksanaan proses pembelajaran, sedangkan tes sumatif diberikan di akhir pembelajaran atau akhir semester.

Selain itu, Zainal Arifin (2009: 117-124) menjelaskan penggolongan tes dalam beberapa kategori. Berdasarkan jumlah peserta didik, tes dibagi menjadi dua jenis yaitu tes kelompok dan tes perorangan. Berdasarkan cara penyusunannya, tes dibagi menjadi tes buatan guru dan tes standar. Berdasarkan kajian psikologi, tes dibagi menjadi tes intelegensia umum, tes kemampuan khusus, tes prestasi belajar dan tes kepribadian. Selain itu, tes berdasarkan bentuk jawaban peserta didik, tes dibagi menjadi

tes tertulis, tes lisan dan tes tindakan. Dari tes tertulis tersebut dapat dibagi menjadi dua yaitu tes uraian yang terdiri dari uraian bebas dan uraian terbatas dan tes objektif yang terdiri dari tes benar salah, pilihan ganda, menjodohkan dan melengkapi.

Dalam penelitian ini, jenis tes yang digunakan adalah soal pilihan ganda. Soal pilihan ganda menghendaki siswa memilih respon yang tepat dari pilihan yang disediakan. Pada beberapa item pilihan siswa memilih jawaban terbaik dari daftar alternatif. Pemilihan jenis tes ini karena beberapa pertimbangan mengenai kelebihan dan kekurangan jenis tes pilihan ganda (Reynolds, Livingstone, & Wilson, 2009: 184-186; dan Cohen & Swerdlik, 2009: 261). Keunggulan soal pilihan ganda adalah mencakup sejumlah relatif besar item dengan pilihan jawaban dalam pengujian, dapat menilai secara efisien, objektif, dan cara yang dapat diandalkan. Selain itu, tes pilihan ganda sangat baik untuk mengukur tujuan tingkat yang lebih rendah (misalnya pengetahuan, pemahaman, dan aplikasi) serta dapat mengurangi pengaruh konstruk tertentu faktor yang tidak relevan.

Kelemahan soal pilihan ganda adalah relatif sulit untuk dibuat karena membutuhkan waktu lebih banyak untuk membuat soal pilihan ganda, tidak mampu untuk menilai semua tujuan pendidikan (misalnya kemampuan menulis), meskipun item dengan pilihan jawaban sangat cocok untuk menilai tujuan kognitif tingkat yang lebih rendah, mereka tidak cocok untuk menilai tujuan tingkat yang lebih tinggi (yaitu analisis, sintesis, dan evaluasi) dan dapat terjadi tebakan acak.

Kelebihan dan kekurangan tes pilihan ganda tersebut menjadikan pertimbangan dalam pemilihan jenis tes. Tes yang akan digunakan dalam penelitian ini cocok menggunakan soal pilihan ganda. Hal itu karena jumlah siswa yang akan diberikan tes relatif banyak dan membutuhkan soal yang dapat mengukur kemampuan secara akurat.

Pengembangan tes pada penelitian ini dilakukan dengan beberapa langkah. Djemari Mardapi (2012: 110) dan Heri Retnawati (2015) menyebutkan bahwa instrumen tes dapat disusun dengan langkah-langkah tertentu. Pembuatan tes diawali dengan menentukan tujuan tes dilanjutkan dengan mencari cakupan materi, menyusun indikator butir soal, menyusun butir instrumen, melakukan validasi, merevisi soal berdasarkan hasil validasi, melakukan uji coba, melakukan analisis hasil uji coba, dan merakit instrumen. Langkah-langkah tersebut dilakukan dengan harapan instrumen yang baik dapat diperoleh.

2. Penulisan Butir Soal

Butir butir soal yang dituliskan harus memiliki peringkat kognitif yang tinggi. Menurut Moore, B dan Stanley T (2010), taksonomi Bloom yang mencakup: mengingat, memahami, menerapkan, menganalisis, mengevaluasi, dan melakukan kreasi merupakan urutan, dari yang paling rendah (peringkat 1) ke yang paling tinggi (peringkat 6). Selanjutnya, Moore, B dan Stanley T (2010), menambahkan bahwa urutan nomor 1 – 3 dikategorikan *the lower level of thinking* dan 4 -6 *the higher level of thinking (HOT)*. Hal ini senada dengan pendapat Thomas, A. dan Thorne, G. (2007) yang mengatakan *HOT is thinking on a higher level than memorizing facts or telling something back to someone exactly the way the it was told to you. When a person memorizes and gives back the information without having to think about it, we call it rote memory. That's because it's much like a robot; it does what it's programmed to do, but it doesn't think for itself.*

1. Langkah-langkah Penulisan Butir Soal

Seperti yang dijelaskan di atas bahwa langkah-langkah pengembangan suatu tes prestasi belajar adalah : (1) penentuan tujuan tes, (2) penyusunan kisi-kisi, (3) penulisan soal, (4) penelaahan soal (review dan revisi soal), (5) uji coba soal, termasuk analisis dan perbaikan, dan (6) perakitan soal menjadi perangkat tes.

a. Penentuan tujuan/penyusunan *blueprint*

Dalam melakukan pengetesan pasti ada tujuan yang ingin dicapai. Tujuan ini dapat berupa tujuan khusus, misal untuk mengetahui penguasaan materi, tes diagnostik, atau tes seleksi; dan tujuan umum, misal untuk mengetahui pengetahuan umum dari sekelompok responden atau sekelompok orang. Dalam kesempatan ini, tujuan pemberian tes adalah untuk mengetahui penguasaan peserta didik pada kompetensi/sub kompetensi tertentu setelah diajarkan. Penguasaan ini dapat diartikan, sejauh mana peserta didik memahami atau mungkin menganalisis materi tertentu yang telah dibahas di ruang kelas. Dengan kata lain, pada tingkat kognitif mana mereka menguasai materi yang telah diberikan, ditugaskan, atau dibahas, yang biasanya direncanakan dalam bentuk *blue print*. Tujuan tes harus jelas agar arah dan ruang lingkup pengembangan tes selanjutnya juga jelas.

b. Penyusunan Kisi-kisi

Kisi-kisi adalah panduan atau acuan dalam menyiapkan bahan ajar, menyelenggarakan pembelajaran, dan mengembangkan butir-butir soal uji. Kisi-kisi soal

tes yang merupakan bagian dari silabus ini biasanya berisi standar kompetensi, kompetensi dasar, materi pokok, kegiatan pembelajaran, penilaian, waktu, dan sumber belajar. Hal yang harus diperhatikan dalam menyusun kisi-kisi adalah indikator jabaran dari kompetensi dasar (KD), kompetensi dasar jabaran dari standar kompetensi (SK), standar kompetensi jabaran dari standar kompetensi lulusan mata pelajaran (SKL-MP), dan standar kompetensi lulusan mata pelajaran jabaran dari standar kompetensi lulusan satuan pendidikan (SKL-P), dan standar kompetensi lulusan satuan pendidikan jabaran dari Tujuan Pendidikan Nasional.

Kompetensi lulusan dijabarkan ke dalam subkompetensi, selanjutnya subkompetensi dijabarkan menjadi indikator esensial dan deskriptor. Sama halnya pada kompetensi dan subkompetensi, kata utama dalam indikator esensial dan deskriptor juga kata kerja, hanya saja skopanya sama atau lebih sempit dan peringkat kognitifnya sama atau lebih rendah. Contoh format Kisi-kisi dapat dilihat pada Lampiran 1.

c. Penulisan butir-butir soal/tes

Penulisan butir-butir soal merupakan langkah penting dalam upaya pengembangan alat ukur kemampuan atau tes yang baik. Penulisan soal adalah penjabaran indikator jenis dan tingkat perilaku yang hendak diukur menjadi pertanyaan-pertanyaan yang karakteristiknya sesuai dengan perinciannya dalam kisi-kisi. Butir soal merupakan jabaran atau dapat juga ujud dari indikator, Dengan demikian setiap pernyataan atau butir soal perlu dibuat sedemikian rupa sehingga jelas apa yang ditanyakan dan jelas pula jawaban yang diminta. Mutu setiap butir soal akan menentukan mutu soal tes secara keseluruhan. Butir-butir soal harus memiliki tingkat penalaran tinggi atau memiliki *Higher Order Thinking (HOT)*.

d. Telaah Soal atau Analisis Kualitatif Soal

Telaah soal atau analisis kualitatif soal adalah mengkaji secara teoritik soal tes yang telah tersusun. Telaah ini dilakukan dengan memperhatikan tiga aspek, yaitu aspek materi, aspek konstruksi, dan aspek bahasa. Tabel telaah butir dapat dilihat pada Lampiran 1.

e. Ujicoba Soal

Ujicoba soal pada dasarnya adalah upaya untuk mengetahui kualitas soal tes berdasarkan pada empirik atau respon dari peserta tes. Hal ini dapat terwujud manakala

dilakukan analisis empirik atau analisis kuantitatif, baik menggunakan teori klasik maupun teori modern.

f. Analisis Empirik

Untuk mengetahui kualitas butir soal, maka hasil uji coba harus dianalisis secara empirik. Ada dua pendekatan yang digunakan untuk melakukan analisis empirik ini, yaitu: teori klasik dan teori respon. Masing-masing pendekatan ada kelebihan dan kekurangannya. Untuk responden yang kecil (kurang dari 100) lebih cocok menggunakan teori klasik, sebaliknya untuk responden yang besar (lebih besar dari 200) lebih cocok menggunakan teori respon butir.

Pada analisis butir soal dengan pendekatan teori tes klasik, parameter butir yang diestimasi adalah tingkat kesulitan butir, daya beda, dan keberfungsian distraktor. Selain parameter butir, pada pendekatan teori tes klasik juga mengestimasi reliabilitas tes. Sementara itu, pada pendekatan teori tes modern (Item Response Theory /IRT) 3 parameter butir (3 PL), yang diestimasi adalah tingkat kesulitan, daya beda, dan *factor guessing*. Bila 2 PL, yang diestimasi adalah tingkat kesulitan dan daya beda, sedangkan untuk 1 PL yang diestimasi hanya tingkat kesulitan butir saja. Sebagai pengganti reliabilitas, pada pendekatan dengan IRT dihitung Fungsi Informasi; baik fungsi informasi butir maupun fungsi informasi tes.

g. Perakitan Soal Tes

Agar skor tes yang diperoleh tepat dan dapat dipercaya maka soal tes harus valid dan reliabel. Butir-butir soal perlu dirakit menjadi alat ukur yang terpadu. Hal-hal yang dapat mempengaruhi validitas skor tes adalah urutan nomor soal, pengelompokan bentuk-bentuk soal, tata letak soal, dan sebagainya. Untuk itu, ada baiknya soal tes disajikan mulai dari butir mudah ke yang susah, pengelompokan rapi, tata letak bagus dan tidak terpotong-potong kalimatnya, dan kemasannya menarik.

3. Analisis Perangkat Tes

a. Validitas Tes

Suatu tes harus mempunyai validitas tes yang baik. Tes yang baik adalah tes yang dapat mengukur apa yang diukur (Allen & Yen, 1979: 95) dan relevan dengan apa yang diukur (Carroll, 2005: 14). Sawyer (2004: 95) dan Koretz (2008: 215) menyebutkan

bahwa validitas merupakan kriteria terpenting untuk mengevaluasi tes prestasi belajar. Miller, Linn, & Grondlund (2009: 71) menjelaskan lebih lanjut mengenai hubungan validitas dan reliabilitas tes yaitu reliabilitas itu dibutuhkan tetapi tidak selalu menjadi kondisi penentu bagi validitas. Suatu tes dapat mempunyai reliabilitas yang baik namun tidak mengukur apa yang diukur sehingga dapat dikatakan bahwa hasil tes yang valid memerlukan bukti reliabilitas, namun tes tidak cukup hanya reliabel tanpa mempunyai bukti validitas. Validitas dapat dikategorikan menjadi 3 yaitu validitas isi, validitas kriteria dan validitas konstruk (Cohen & Swerdlik, 2009: 173; dan Djemari Mardapi, 2012: 39-45).

Validitas isi tes dapat diperoleh dengan menganalisis hubungan antara isi tes dan konstruk yang ingin diukur oleh pakar di bidang yang diukur. Pada tes hasil belajar, beberapa pakar akan menilai apakah isi tes sesuai dengan materi yang ingin diukur (Crocker & Algina, 2008: 218). Validitas isi sering dijelaskan melalui validitas tampak dengan memeriksa item-item tes apakah tes mengukur aspek yang relevan serta validitas logik yang dapat ditunjukkan dengan spesifikasi tes yang ditulis sebagai pedoman menyusun tes (Allen & Yen, 1979: 96; dan Djemari Mardapi, 2012: 40-41). Pada tes *achievement*, validitas isi sangat penting (Koretz, 2008: 31) dan hasil analisis item adalah hal penting kedua setelah validitas isi (Fernandes, 1984: 29). Lebih lanjut Heri Retnawati (2014) menjelaskan bahwa proses validasi instrumen minimal melibatkan 3 *judgement*. Validitas isi dapat dihitung dengan rumus I (Aiken, 1980: 956).

$$V = \frac{\sum s}{n(c-1)} \dots\dots\dots(1)$$

Keterangan:

V: indeks kesepakatan rater mengenai validitas butir.

s : skor yang ditetapkan setiap rater dikurangi skor terendah dalam kategori yang dipakai.

n : banyaknya rater

c : banyaknya kategori yang dapat dipilih rater.

Validitas kriteria dicari dengan bukti berdasarkan hubungan dengan variabel lain. Validitas ini dicari dengan menganalisis variabel eksternal yang diharapkan memprediksi hal yang diukur (Djemari Mardapi, 2012: 44; dan Allen & Yen, 1979: 97). Validitas ini

berdasarkan waktu pengambilan datanya dibagi menjadi dua yaitu validitas konkuren dan validitas prediktif.

Bukti berdasarkan respon merupakan analisis yang dilakukan terhadap respon siswa yang akan digunakan untuk mencari validitas konstruk dan validitas faktorial. Validitas konstruk menguji sejauh mana tes dapat mengukur konsep dari suatu teori yang menjadi dasar dalam menyusun tes. Validitas faktorial dilakukan untuk mengetahui berapa banyak faktor atau dimensi yang diukur tes.

Analisis soal juga dilakukan dengan analisis kualitatif. Analisis dilakukan dengan pedoman analisis butir soal secara kualitatif menggunakan format penelaahan soal yang dikeluarkan Departemen Pendidikan Nasional (2008: 6-7). Format penelaahan soal digunakan sebagai dasar untuk menganalisis setiap butir soal. Aspek yang diperhatikan dalam analisis butir soal secara kualitatif mencakup aspek materi, konstruksi, bahasa, dan kunci jawaban/pedoman penskorannya. Penelaah perlu mempersiapkan bahan-bahan penunjang seperti kisi-kisi tes, kurikulum yang digunakan, buku sumber, dan kamus bahasa Indonesia dalam melakukan penelaahan butir soal.

b. Reliabilitas Tes

Reliabilitas merupakan penerjemahan dari kata *reliability*. Suatu pengukuran yang mampu menghasilkan data yang memiliki tingkat reliabilitas tinggi disebut sebagai pengukuran yang reliabel (*reliable*). Menurut Saifuddin Azwar (2013), walaupun istilah reliabilitas mempunyai berbagai nama lain seperti konsistensi, keterandalan, keterpercayaan, kestabilan, keajegan, dan sebagainya, namun gagasan pokok yang terkandung dalam konsep reliabilitas adalah sejauhmana hasil suatu proses pengukuran dapat dipercaya.

Suatu tes dikatakan reliabel jika skor amatan mempunyai korelasi yang tinggi dengan skor sebenarnya (Allen & Yen, 1979: 72). Dengan demikian pengertian yang dapat diperoleh dari pernyataan tersebut adalah suatu tes reliabel jika hasil pengukuran mendekati keadaan peserta tes yang sebenarnya. Koefisien reliabilitas yang baik adalah diatas 0,70 (Linn, 1989: 106; Djemari Mardapi, 2014: 3) meskipun secara teoritik besarnya koefisien reliabilitas berkisar dari 0,00 sampai 1,00 (Saifuddin Azwar, 2013: 13). Apabila suatu tes berisi butir-butir yang diberi skor dikotomi, sedangkan jumlah butirnya tidak begitu banyak, maka estimasi reliabilitas dilakukan

melalui formula Alpha yang disesuaikan dengan data dikotomi yang dikenal dengan formula *Kuder-Richardson-20* (KR-20) (Saifuddin Azwar, 2013: 73).

Dalam pendidikan, pengukuran tidak dapat langsung dilakukan pada ciri atau karakter yang akan diukur. Ciri atau karakter ini bersifat abstrak. Hal ini menyebabkan sulitnya memperoleh alat ukur yang stabil untuk mengukur karakteristik seseorang (Mehrens & Lehmann, 1973: 103). Allen & Yen (1979: 62) menyatakan bahwa 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.

Kaitannya dengan tes, reliabilitas skor hasil tes merupakan informasi yang diperlukan dalam pengembangan tes. Mehrens & Lehmann (1973: 102) menyatakan bahwa reliabilitas merupakan derajat keajegan (consistency) di antara dua buah hasil pengukuran pada objek yang sama. Definisi ini dapat diilustrasikan dengan seseorang yang diukur tinggi badannya akan diperoleh hasil yang tidak berubah walaupun menggunakan alat pengukur yang berbeda dan skala yang berbeda. Dalam kaitannya dengan dunia pendidikan, prestasi atau kemampuan seorang siswa dikatakan reliabel jika dilakukan pengukuran, hasil pengukuran akan sama informasinya, walaupun pengujinya berbeda, korektornya berbeda atau butir soal yang berbeda tetapi memiliki karakteristik yang sama.

Pembuatan alat ukur dalam dunia pendidikan harus dilakukan secara cermat mungkin dan disesuaikan dengan kaidah-kaidah yang telah ditentukan oleh ahli-ahli pengukuran di bidang pendidikan (Retnawati). Untuk melihat reliabilitas suatu alat ukur, yang berupa suatu indeks reliabilitas, dapat dilakukan penelaahan secara statistik. Nilai ini biasa dinamakan dengan koefisien reliabilitas (reliability coefficient).

Koefisien reliabilitas dapat diartikan sebagai koefisien keajegan atau kestabilan hasil pengukuran. Alat ukur yang reliabel adalah alat ukur yang mampu membuahkan hasil pengukuran yang stabil (Lawrence, 1994) dan konsisten (Mehrens & Lehmann, 1973: 102). Artinya suatu alat ukur dikatakan memiliki koefisien

reliabilitas tinggi manakala digunakan untuk mengukur hal yang sama pada waktu berbeda hasilnya sama atau mendekati sama. Dalam hal ini, reliabilitas merupakan sifat dari sekumpulan dari skor (Frisbie, 2005). Dalam kaitannya dengan dunia pendidikan, dengan alat ukur yang reliabel, hasil pengukuran akan sama informasinya walaupun pengujian berbeda, koreksinya berbeda atau butir soal yang berbeda tetapi mengukur hal yang sama dan memiliki karakteristik butir yang sama.

Penghitungan reliabilitas disebut dengan estimasi. Estimasi reliabilitas tes yang dapat dilakukan dengan dua cara, baik konsistensi eksternal dan maupun konsistensi internalnya.

1) Pengukuran Konsistensi Eksternal

Reliabilitas eksternal diperoleh dengan cara mengolah hasil pengujian yang berbeda, baik dari instrument yang berbeda maupun yang sama. Ada dua cara untuk estimasi reliabilitas eksternal suatu instrument yaitu dengan teknik ulang dan teknik paralel.

a) Metode Test Ulang (Test-Retest-Method)

Untuk mengetahui sampai dimana suatu pengukuran dapat diandalkan, pengukuran ini dapat dilakukan dua kali, pengukuran pertama dan ulangnya. Kedua pengukuran ini dapat dilakukan oleh orang yang sama atau berbeda. Dalam hal ini perlu diatur bahwa proses pengukuran kedua, keadaan yang diukur itu harus benar-benar sama. Selanjutnya hasil pengukuran yang pertama dan yang kedua dikorelasikan dan hasilnya menunjukkan reliabilitas dari tes ini. Memang teknik ulangan ini akan dapat memenuhi sasaran bila keadaan subjek yang diukur (dites) tetap bertahan dan tidak mengalami perubahan pada saat pengukuran yang pertama maupun pada pengukuran yang kedua. Karena keadaan pribadi anak itu selalu dalam keadaan berkembang, tidak statis, maka sebenarnya teknik ini kurang tepat digunakan. Disamping itu pada pengukuran yang kedua akan dijumpai adanya "testing effect" anak telah mendapat tambahan pengetahuan karena sudah mengalami tes yang pertama. Reliabilitas tes retes ini penting ketika kita menafsirkan koefisien tes-retes untuk mengetahui : a) jangka waktu antara kedua pengambilan penilaian, b) stabilitas yang diharapkan dari kinerja yang diukur. Secara umum, semakin lama antara interval pelaksanaan tes yang berulang, semakin rendah tingkat reliabilitasnya. Estimasi reliabilitas dengan pendekatan tes-retes akan menghasilkan koefisien stabilitas. Untuk memperoleh

koefisien reliabilitas melalui pendekatan tes-retes dapat dilakukan dengan menghitung koefisien korelasi linier antara distribusi skor subjek pada pemberian tes pertama dengan skor subjek pada pemberian skor kedua.

b) Metode Bentuk Paralel (*Equivalent*)

Tes paralel atau tes *equivalent* adalah dua buah tes yang mempunyai kesamaan tujuan, tingkat kesukaran dan susunan tetapi butir-butir soalnya berbeda, dalam istilah bahasa Inggris disebut *alternate-forms method* (parallel forms). Dengan metode bentuk paralel ini, dua buah tes yang paralel, misalnya tes Matematika seri A yang akan dicari reliabilitasnya dan tes Seri B diteskan kepada sekelompok siswa yang sama, kemudian dikorelasikan. Koefisien korelasi dari kedua hasil tes inilah yang menunjukkan koefisien reliabilitas tes seri A.

Jika koefisiennya tinggi maka tes tersebut sudah reliabel dan dapat digunakan sebagai alat dan pengetes yang terandalkan. Dalam menggunakan metode tes paralel pengetes harus menyiapkan dua buah tes, dan masing-masing dicobakan kepada sekelompok siswa yang sama. Penggunaan metode ini baik karena siswa dihadapkan kepada dua macam tes sehingga tidak ada factor “ masih ingat-ingat soalnya” yang dalam evaluasi disebut adanya *practice effect*- dan *carry-over-effect*. Artinya ada factor yang dibawa oleh pengikut tes karena sudah mengerjakan soal tersebut. Kelemahan dari metode ini adalah bahwa pengetes pekerjaannya berat karena harus menyusun dua seri tes. Lagipula harus tersedia waktu yang lama untuk mencobakan dua kali tes. Mengenai pertanyaan bagaimana proses melaksanakan tes reliabilitas secara ekivalen? Berikut ini akan ditunjukkan beberapa langkah-langkah tersebut diantaranya adalah sebagai berikut:

- 1) Menentukan subjek sasaran yang hendak di tes
- 2) Melakukan tes yang dimaksud kepada sasaran subjek yang dimaksud
- 3) Diadministrasi dengan baik
- 4) Dalam waktu yang tidak begitu lama melakukan tes yang kedua pada
- 5) kelompok tersebut.
- 6) Mengkorelasikan antara kedua skor tes tersebut.

Jika hasil koefisien ekivalen tinggi, berarti tes memiliki reliabilitas ekivalen baik. Sebaliknya, jika ternyata koefisien rendah maka reliabilitas ekivalen tes adalah rendah.

Reliabilitas ekivalen merupakan salah satu bentuk yang diterima dan umum dipakai penelitian terutama penelitian pendidikan.

Pada metode ini suatu tes dibagi dua, bisa gasal dan genap atau bagian awal dan akhir. persyaratan yang harus dipenuhi pada metode ini adalah rerata bagian pertama dan kedua sama, demikian pula varians kedua belahan sama, serta materi yang diukur juga sama (Mardapi, 2016: 55).

Pada belahan pertama:

$$X_1 = T_1 + E_1 \longrightarrow \mu_{x_1} = \mu_{T_1} + 0$$

Pada belahan kedua:

$$X_2 = T_2 + E_2 \longrightarrow \mu_{x_2} = \mu_{T_2} + 0$$

Besarnya rerata skor bagian satu dan bagian dua adalah sebagai berikut:

$$\mu_{x_1} = \mu_{x_2} \text{ maka } \sigma^2_{T_1} = \sigma^2_{T_2} \text{ dan } \sigma^2_{x_1} = \sigma^2_{x_2}$$

$$\sigma^2_{x_1} = \sigma^2_{T_1} + \sigma^2_{E_1}$$

$$\sigma^2_{x_2} = \sigma^2_{T_2} + \sigma^2_{E_2}$$

$$\text{Karena } \sigma^2_{E_1} = \sigma^2_{E_2} \text{ maka } \sigma^2_{x_1} = \sigma^2_{x_2}$$

Jadi model paralel klasik adalah $\mu_1 = \mu_2$ dan $\sigma^2_{x_1} = \sigma^2_{x_2}$

Estimasi reliabilitas dengan metode paralel klasik yang juga dikenal Spearman-Brown, dilakukan dengan menghitung korelasi skor tampak antara dua tes yang paralel yang disajikan pada kelompok subjek yang sama. Dikarenakan hampir tidak mungkin untuk memperoleh dua tes yang paralel, maka biasanya digunakan bentuk alternatif. Bentuk alternatif tes mengukur satu trait yang sama dan disusun sedemikian rupa sehingga menjadi separalel mungkin sehingga keduanya dapat mempunyai distribusi skor tampak dengan rerata dan korelasi dengan tes lain yang sama besar.

Salah satu prosedur dalam metode konsistensi internal sangat populer adalah yang menghasilkan estimasi reliabilitas *split-half* (Mardapi, 2016: 57). Tes yang akan diestimasi reliabilitasnya, dibelah menjadi dua bagian yang diusahakan paralel. pembelahan ini dilakukan setelah keseluruhan tes sebagai kesatuan dikenakan pada para subjek. kemudian diperoleh distribusi skor subjek untuk belahan pertama dan belahan kedua. Pendekatan konsistensi internal dalam estimasi reliabilitas dimaksudkan, antara lain, untuk menghindari permasalahan yang biasanya ditimbulkan oleh pendekatan

tes-ulang dan pendekatan bentuk-paralel. Dalam pendekatan konsistensi internal data skor diperoleh melalui prosedur satu kali pengenaaan satu tes kepada sekelompok individu sebagai subjek (*single-trial administration*), sehingga metode ini mempunyai nilai praktis dan efisiensi yang tinggi dibanding prosedur tes-ulang dan bentuk-paralel.

Dengan suatu kali pengenaaan tes akan diperoleh hanya satu distribusi skor tes saja dari kelompok sampel yang bersangkutan sehingga analisis tidak dilakukan terhadap skor tes melainkan terhadap distribusi skor aitem-aitem dalam tes. Makna konsistensi internal adalah konsistensi di anatar aitem-aitem dalam tes sebagai indikasi bahwa tes yang bersangkutan memiliki fungsi pengukuran yang reliable. Dengan kata lain, prosedur estimasi reliabilitasnya harus dilakukan melalui analisis terhadap distribusi skor aitem atau distribusi skor kelompok-kelompok aitem, tidak dilakukan melalui analisis terhadap skor tes. Bila estimasi reliabilitas pengukuran dilakukan dengan cara melihat konsistensi di antara kelompok-kelompok aitem maka perlu dibuat beberapa kelompok aitem yang disebut sebagai bagian atau belahan tes. Setiap bagian atau belahan tes dapat berisi beberapa aitem, bahkan dapat berisi hanya satu aitem saja. Banyaknya belahan adalah dua, tiga, atau sebanyak jumlah aitemnya.

Pembelahan tes harus dilakukan sedemikian rupa sehingga sedapat mungkin setiap belahan berisi aitem dalam jumlah yang sama banyak dan berkarakteristik yang setara. Apabila banyaknya aitem yang ada dalam tes adalah genap hendaknya tes itu dibelah menjadi hanya dua belahan agar masing-masing belahan berisi cukup banyak aitem dalam jumlah yang sama. Apabila banyaknya aitem dalam tes tidak berjumlah genap tapi masih dapat dipisahkan menjadi tiga bagian yang sama panjang, hendaknya tes dibelah menjadi tiga kelompok yang masing-masing berisi aitem dalam jumlah yang sama banyak dan berkarakteristik setara. Bila aitem dalam tes berjumlah hanya sedikit, misalnya kurang dari 30 aitem, sekalipun dapat dipisahmenjadi dua atau tiga bagian namun sebaiknya tes dibelah menjadi sebanyak jumlah aitemnya sehingga setiap bagian hanya berisi satu aitem saja.

Akan ditemukan kasus tes dengan jumlah aitem yang tidak dapat dibelah menjadi dua atau menjadi tiga belahan yang sama panjang, misalnya tes yang berisi 51 aitem, 67 aitem, 73 aitem, dan semacamnya. Dalam hal seperti itu, tes dapat dibelah menjadi dua belahan yang tidak sama panjang namun tetap diusahakan agar karakteristik dan isi aitem dalam kedua belahan setara satu sama lain. Tersedia formula khusus yang dapat

digunakan untuk estimasi reliabilitas dalam kasus tersebut. Pilihan cara pembelahan tes banyak tergantung pada kisi-kisi tes (pertimbangan aspek atau komponen), banyaknya aitem, karakteristik aitem (isinya homogeny atau tidak), sifat dan fungsi tes (*power test* atau *speed test*), dan lain-lain. Cara pembelahan itu, pada gilirannya, akan ikut menentukan pula formula mana yang harus digunakan dalam menghitung koefisien reliabilitasnya. Tes yang skornya sedikit-banyak ikut dipengaruhi oleh kecepatan kerja (*speed-test*), misalnya, menghendaki cara pembelahan yang tidak sama dengan cara pembelahan yang dilakukan terhadap tes yang mengukur kemampuan maksimum (*power-test*). Suatu tes yang berisi aitem-aitem yang mempunyai taraf kesukaran homogen akan lebih terbuka terhadap berbagai cara pembelahan dibandingkan dengan tes yang berisi aitem-aitem dengan tingkat kesukaran yang sangat bervariasi. Tentu tidak setiap jenis dan karakteristik tes menghendaki cara pembelahan khusus, akan tetapi setiap cara pembelahan tes yang digunakan hendaknya ditekankan pada usaha maksimal untuk memperoleh bagian-bagian atau belahan-belahan yang parameter dan isi aitem-aitemnya relatif setara. Metode ini hanya memerlukan satu kali penyajian tes saja, dikenal dengan nama *single-trial administration*, dan karena itu masalah-masalah yang timbul akibat penyajian yang berulang dapat dihindari. Pendekatan dalam melakukan estimasi reliabilitas dengan konsistensi internal ada tiga yaitu klasik paralel, tau-ekuivalen, dan konginerik. Masing-masing pendekatan menggunakan persyaratan yang berbeda, yaitu sesuai dengan karakteristik data.

4. Teori Tes Klasik

Penggunaan teori tes klasik dalam konstruksi dan interpretasi skor telah berjalan beberapa dekade lamanya. Menurut Djemari Mardapi (2008: 32), teori tes klasik menggunakan model pengukuran yang sangat sederhana, yakni skor yang tampak merupakan penjumlahan dari skor murni dan kesalahan pengukuran. Hubungan antara skor tampak atau *observed score* (X), skor murni atau *true score* dan kesalahan pengukuran atau *error* (E) dapat ditulis dalam persamaan $X = T + E$.

Teori klasik merupakan sebuah teori yang mudah dalam penerapannya serta model yang cukup berguna dalam mendeskripsikan bagaimana kesalahan dalam

pengukuran dapat mempengaruhi skor tampak. Kesalahan pengukuran digolongkan menjadi dua, yaitu acak dan sistemik (Djemari Mardapi, 2008: 67). Kesalahan acak disebabkan karena kesalahan dalam menentukan sampel isi tes, variasi emosi seseorang, termasuk variasi emosi pemeriksa lembar jawaban jika lembar jawaban siswa diperiksa secara manual; sedangkan kesalahan sistemik disebabkan karena soal terlalu mudah atau sukar dan kecenderungan guru untuk memberikan skor yang mahal atau murah.

a. Tingkat kesukaran

Bermutu atau tidaknya butir tes dapat diketahui dari tingkat kesukaran yang dimiliki oleh masing-masing butir tersebut. Butir tes dikatakan baik jika butir tersebut tidak terlalu sukar dan tidak pula terlalu mudah, atau dengan kata lain tingkat kesukaran butir adalah sedang atau cukup. Tingkat kesukaran butir biasanya dikaitkan dengan tujuan tes (Mehrens & Lehmann, 1973: 195; Heri Retnawati, 2012). Menurut Haryani (2011), untuk keperluan ujian semester, digunakan butir dengan indeks kesukaran sedang; untuk keperluan seleksi digunakan butir dengan indeks kesukaran rendah (sukar); dan untuk kepentingan diagnostik digunakan butir yang memiliki indeks kesukaran tinggi (mudah).

Tingkat kesukaran suatu butir dinyatakan oleh suatu indeks yang dinamakan indeks kesukaran butir dan disimbolkan oleh huruf p , yang merupakan rasio antara penjawab butir dengan benar dan banyaknya penjawab butir (Saifuddin Azwar, 2002: 134). Secara teoritik dapat dikatakan bahwa p sebenarnya merupakan probabilitas empirik untuk lulus butir tertentu bagi kelompok siswa tertentu. Dengan kata lain, indeks kesukaran butir tidak selalu sama antara satu kelompok dengan kelompok siswa yang lain. Selain itu, harga p yang dihitung merupakan indeks kesukaran butir bagi seluruh kelompok, bukan indeks kesukaran item bagi masing-masing individu.

Hal senada juga diutarakan oleh Anastasi & Urbina (1997: 173) bahwa pada pengukuran butir tes, tingkat kesukaran butir berhubungan dengan presentase orang-orang yang dapat menjawab soal dengan benar. Tingkat kesukaran ini merupakan proporsi jumlah peserta tes yang menjawab benar dari suatu butir soal, yaitu perbandingan antara jumlah peserta tes seluruhnya. Besarnya tingkat kesukaran item terletak antara 0 sampai 1. Semakin besar tingkat kesukaran, maka semakin mudah butir tes tersebut. Sebaliknya semakin kecil tingkat kesukaran, maka butir tes akan

semakin sukar (Saifuddin Azwar, 2002: 137). Harga tingkat kesukaran butir pada suatu tes, menurut Djemari Mardapi (2008: 141) dapat ditentukan dengan rumus:

$$P = \frac{B}{T}$$

Keterangan :

P : Tingkat kesukaran (*proportion correct*)

B : Jumlah peserta tes yang menjawab butir dengan benar

T : Jumlah seluruh peserta tes yang mengerjakan butir

Indeks kesukaran butir yang dapat diterima adalah 0,30-0,70 karena pada interval ini informasi kemampuan siswa diperoleh dengan maksimal (Allen & Yen, 1979: 121). Rahma Zulaiha (2008: 17) menambahkan bahwa butir soal yang memiliki indeks kesukaran di atas 0,7 termasuk butir mudah sedangkan butir soal yang memiliki indeks kesukaran di bawah 0,3 termasuk butir sukar. Akan tetapi, Nitko (1983) secara lebih tegas mengklasifikasikan tingkat kesukaran menjadi:

Tabel 1. Range Tingkat Kesukaran

Range tingkat kesukaran	Kategori	Keputusan
0,30 s.d 0,70	Sedang	Diterima
0,10 s.d 0,29 atau 0,70 s.d 0,90	Sukar atau mudah	Direvisi
< 0,10 atau > 0,90	Sangat sukar atau sangat mudah	Ditolak

b. Daya Pembeda

Daya pembeda atau daya beda (indeks diskriminasi) suatu butir adalah kemampuan suatu butir dalam membedakan siswa yang mempunyai kemampuan tinggi dan siswa yang mempunyai kemampuan rendah (Saifuddin Azwar, 2002; Reynolds, Livingston & Wilson, 2009). Suatu butir yang dikatakan mempunyai daya beda tinggi haruslah dijawab dengan benar oleh semua atau sebagian besar subjek kelompok tinggi dan tidak dapat dijawab dengan benar oleh semua atau sebagian besar subjek kelompok rendah. Semakin besar perbedaan antara proporsi penjawab benar kelompok tinggi dan kelompok rendah, semakin besarlah daya beda suatu butir. Jika proporsi penjawab benar dari kedua kelompok tersebut sama, maka butir tersebut tidak dapat membedakan antara mereka yang berkemampuan tinggi dan berkemampuan rendah.

Apalagi bila suatu butir ternyata dijawab dengan benar oleh sebagian besar subjek kelompok bawah sedangkan subjek kelompok atas tidak banyak yang menjawab benar, maka jelas butir yang seperti ini akan menyesatkan karena daya bedanya terbalik. Berdasarkan uraian di atas, dapat disimpulkan bahwa indeks daya beda ditetapkan dari selisih proporsi yang menjawab dari masing-masing kelompok. Indeks ini dapat dihitung dengan korelasi point biserial dan korelasi biserial (Allen & Yen, 1989: 122). Korelasi *point biserial* maupun *biserial* merupakan korelasi product momen yang diterapkan pada data. Korelasi point biserial digunakan jika variabel skor butir soal bersifat dikotomi dan variabel skor atau sub-skor bersifat kontinum. Variabel skor tes bersifat dikotomi karena skor-skor yang terdapat pada tes hanya ada 1 dan 0, sedangkan variabel skor total bersifat kontinum karena diperoleh dari jumlah jawaban benar. Korelasi *point biserial* menurut Crocker & Algina (2008: 317) ditentukan dengan menggunakan rumus:

Indeks daya beda secara klasik terletak antara -1,00 sampai +1,00 (Saifuddin Azwar, 2002: 139). Baik Reynolds, Livingston & Wilson (2010: 152) maupun Badrun Kartowagiran (2012) menyarankan bahwa butir soal yang diterima adalah butir soal yang memiliki indeks daya pembeda lebih dari 0,3. Namun, Crocker & Algina (2008: 315) mengklasifikasikan kriteria besarnya koefisien daya beda sebagai berikut.

Tabel 2. Kategori Penerimaan Daya Beda Butir Soal

Indeks diskriminasi	Keputusan
0,40 atau lebih	Bagus sekali
0,30-0,39	Cukup bagus
0,20-0,29	Belum memuaskan
Kurang dari 0,20	Jelek dan harus dibuang

c. Efektivitas Distraktor

Efektivitas distraktor (pengecoh) yang ada pada suatu butir dianalisis dari distribusi jawaban terhadap butir yang bersangkutan pada setiap alternatif yang disediakan (Saifuddin Azwar, 2002). Sumarna Surapranata (2005b: 43) juga menambahkan bahwa melalui distribusi jawaban dapat diketahui: (1) banyaknya peserta tes yang menjawab benar, (2) pengecoh yang tidak dipilih peserta tes berkemampuan rendah, (3) pengecoh yang menyesatkan, dan (4) pengecoh yang menjadi daya tarik bagi peserta tes berkemampuan rendah. Dalam memeriksa pola

distribusi jawaban untuk melihat efektivitas distraktor, seringkali lebih baik menggunakan proporsi atau persentase jawaban bagi setiap alternatif sebagai ganti frekuensi jawaban dikarenakan proporsi dapat dibandingkan langsung antara satu item dengan item lain tanpa tergantung pada besarnya N. Menurut Saifuddin Azwar (2002), proporsi bagi setiap alternatif diperoleh dari membagi frekuensi jawaban pada setiap pilihan dengan jumlah penjawab atau N.

Efektivitas distraktor pada intinya diperiksa untuk melihat apakah semua distraktor atau semua pilihan jawaban yang bukan kunci telah berfungsi sebagaimana mestinya, yaitu apakah distraktor-distraktor tersebut telah dipilih oleh lebih banyak (atau semua) siswa kelompok rendah, sedangkan siswa dari kelompok tinggi hanya sedikit (atau tidak ada) yang memilihnya. Distraktor dikatakan bekerja efektif manakala setidaknya dipilih oleh 5% siswa (Djemari Mardapi, 2008: 143) atau 2% siswa (Brown dalam Fernandes, 1984: 29). Saifuddin Azwar (2002: 142) menyatakan bahwa efektifitas distraktor dapat dilihat dari dua kriteria: (a) distraktor dipilih oleh siswa dari kelompok rendah, dan (b) pemilih distraktor tersebut relatif proporsional pada masing-masing distraktor. Sebuah pengecoh dikatakan baik, apabila memiliki nilai korelasi point biserial negatif. Nilai negatif menunjukkan bahwa siswa yang berkemampuan rendah cenderung memilih pengecoh tersebut sebagai jawaban, sebaliknya siswa yang berkemampuan tinggi akan memilih kunci jawaban sebagai jawabannya.

5. Teori Respons Butir

a). Konsep Dasar dan Asumsi

Berdasarkan kelemahan-kelemahan yang dimiliki teori tes klasik maka dikembangkanlah teori tes baru, yaitu Teori tes modern atau *Item Response Theory* (IRT) pertama kali dimunculkan oleh Lord dalam disertasinya tahun 1952. Analisis butir soal secara modern yaitu penelaahan butir soal dengan menggunakan IRT. Tujuan utama IRT adalah memberikan kesamaan antara statistik soal dan estimasi kemampuan. Selain itu, Devi (2015) menjelaskan bahwa dalam teori respon butir melepaskan keterpisahan antara butir uji tes dengan peserta tes. Xitao Fan (1998) menjelaskan bahwa pada pendekatan teori respon butir lebih menekankan pada level informasi butir tes, yang berbeda dengan pendekatan klasik yang lebih menekankan pada level informasi perangkat tes.

Menurut Heri Retnawati (2014: 1), dalam teori respon butir, probabilitas subjek untuk menjawab butir dengan benar tergantung pada kemampuan subjek dan karakteristik butir. Ini berarti peserta tes dengan kemampuan tinggi akan mempunyai probabilitas menjawab benar lebih besar jika dibandingkan dengan peserta yang mempunyai kemampuan rendah. 2) Dengan menggunakan teori respons butir, ciri butir akan tetap sama. Maksudnya adalah parameter butir tidak tergantung dengan kemampuan siswa. Selain itu, Edwards (2009) menyatakan bahwa *"IRT can provide item-level weights which more accurate."*

Menurut Hambleton, Swaminathan & Rogers (1991: 7), teori respon butir dikembangkan berdasarkan dua postulat, yaitu: 1) prestasi peserta uji pada suatu tes dapat diprediksikan dengan seperangkat faktor yang disebut kemampuan laten (latent traits), trait adalah dimensi kemampuan seseorang seperti kemampuan verbal, kemampuan psikometer, kemampuan kognitif, dan sebagainya, dan 2) hubungan antara prestasi uji pada suatu butir tes dan perangkat kemampuan yang mendasarinya sesuai dengan grafik fungsi naik monoton tertentu yang disebut kurva karakteristik butir (item characteristic curve). Kurva ini menggambarkan bahwa semakin tinggi kemampuan peserta uji semakin meningkat pula peluang menjawab benar suatu butir tes.

Beberapa asumsi yang melandasi teori respons butir (Hambleton, Swaminathan, & Rogers, 1991: 9-12), yaitu sebagai berikut.

a) Satu Dimensi (*Unidimensi*)

Menurut Heri Retnawati (2014: 1), unidimensi artinya setiap butir tes hanya mengukur satu kemampuan. Oleh karena itu, asumsi unidimensi dapat ditunjukkan hanya jika tes mengandung satu komponen dominan yang mengukur prestasi subjek. Asumsi ini sangat sulit untuk dipenuhi karena banyaknya faktor-faktor yang mempengaruhi seperti kognitif, kepribadian, dan bahasa. Akan tetapi dapat diasumsikan jika dalam hasil analisis terdapat satu faktor yang lebih dominan dari faktor yang lain, maka dapat dikatakan bahwa tes tersebut unidimensi (Dali, 1992: 166).

Menurut De Mars (2010: 39), ada 3 cara yang digunakan untuk menguji asumsi unidimensi, yakni analisis nilai eigen dari matriks korelasi interbutir, uji-Stout pada uji asumsi unidimensi, dan indeks berdasarkan residual pada penyelesaian unidimensi.

b) Kebebasan Lokal (*local Independence*)

Uji asumsi independensi lokal ini bertujuan untuk mengetahui apakah kemampuan peserta tes dalam mengerjakan soal independen terhadap butir/*item* soal. Maksudnya adalah respon peserta tes terhadap suatu butir tidak berhubungan dengan butir lainnya dalam tes tersebut. Menurut Hambleton & Swaminathan (1991: 10), independensi lokal secara matematis dapat dinyatakan sebagai berikut.

$$P(U_1, U_2, \dots, U_n | \theta) = P(U_1 | \theta), P(U_2 | \theta) \dots, P(U_n | \theta) \\ = \prod_{i=1}^n P(U_i | \theta)$$

Keterangan:

i = 1, 2, 3, ... n

n = banyaknya butir tes

$P(u_i | \theta)$ = probabilitas peserta tes yang memiliki kemampuan dapat menjawab butir ke- i dengan benar.

$P(U_1, U_2, \dots, U_n | \theta)$ = probabilitas peserta tes yang memiliki kemampuan dapat menjawab butir ke-1 sampai ke- n dengan benar

c) Invariansi Parameter

Menurut Retnawati (2014: 3), invariansi parameter artinya karakteristik butir soal tidak tergantung pada distribusi parameter kemampuan peserta tes dan parameter yang menjadi ciri peserta tes tidak bergantung pada ciri butir soal. Kemampuan seseorang tidak akan berubah hanya karena mengerjakan tes yang berbeda tingkat kesukarannya dan parameter butir tes tidak akan berubah hanya karena diujikan pada kelompok peserta tes yang berbeda tingkat kemampuannya.

Invariansi parameter kemampuan dapat diselidiki dengan mengajukan dua perangkat tes atau lebih yang memiliki tingkat kesukaran yang berbeda pada sekelompok peserta tes. Invariansi parameter kemampuan akan terbukti jika hasil estimasi kemampuan peserta tes tidak berbeda walaupun tes yang dikerjakan berbeda tingkat kesukarannya. Invariansi parameter butir dapat diselidiki dengan mengujikan tes pada kelompok peserta yang berbeda. Invariansi parameter butir terbukti jika hasil estimasi parameter butir tidak berbeda walaupun diujikan pada kelompok peserta yang berbeda tingkat kemampuannya.

b). Model-Model Teori Respons Butir

Pada umumnya dalam teori respons butir digunakan model distribusi logistik. Hambleton & Swaminathan (1985: 49), menyebutkan beberapa model logistik dalam teori respon butir, diantaranya: 1) model logistik satu parameter, ditentukan oleh satu karakteristik butir yaitu tingkat kesukaran, 2) model logistik dua parameter ditentukan oleh dua karakteristik butir yaitu tingkat kesukaran dan daya pembeda, dan 3) model logistik tiga parameter ditentukan oleh tiga karakteristik butir yakni tingkat kesukaran, daya pembeda, dan faktor tebakan.

Mehtap Erguven (2014) menyatakan bahwa teori respon butir memungkinkan penggunaannya untuk menentukan fungsi matematika pada model hubungan antara sifat laten θ , dan probabilitas bahwa sebuah ujian dengan diberikan θ akan menjawab benar soal tes. Siswa yang memiliki θ tertentu akan memiliki probabilitas menjawab benar pada butir soal tertentu. Hal ini berbeda dengan pendekatan teori tes klasik.

a) Model Logistik Satu Parameter

Model logistik satu parameter ini diajukan oleh George Rasch pada tahun 1966, kemudian dikembangkan di Amerika oleh Benyamin Wright di Universitas Chicago. Hambleton, Swaminathan & Rogers (1991: 14) menyatakan bahwa model logistik satu parameter sering dinamakan model Rasch karena secara matematika ekuivalen dengan Model Rasch. Tingkat kesukaran butir digunakan sebagai parameter pada model ini, sedangkan parameter lainnya seperti daya pembeda dianggap sama, dan *guessing* atau dugaan sama dengan nol. Dengan kata lain semua kurva karakteristik butir soal adalah sejajar atau mendekati sejajar, karena besarnya slope kurva sama dan memotong sumbu y di titik 0 atau mendekati 0.

Hambleton, Swaminathan & Rogers (1991: 12) merumuskan model logistik satu parameter secara matematis sebagai berikut.

$$P_i(\theta) = \frac{e^{(\theta-b_i)}}{1+e^{(\theta-b_i)}} \quad (3)$$

Keterangan:

$P_i(\theta)$ = probabilitas menjawab benar butir i oleh peserta dengan kemampuan θ
 b_i = tingkat kesukaran butir i

- θ = kemampuan peserta tes
- e = bilangan transenden yang besarnya mendekati 2,718

Parameter b_i merupakan suatu titik pada skala kemampuan yang berkaitan dengan peluang menjawab benar sebesar 50%. Misalnya, apabila suatu butir soal mempunyai nilai parameter $b_i = 0,4$ maka diperlukan kemampuan minimal pada skala 0,4 untuk menjawab benar 50%. Semakin besar nilai parameter b_i maka semakin besar kemampuan yang dituntut untuk memperoleh 50% peluang menjawab dengan benar, artinya semakin besar nilai parameter b_i maka butir soal semakin sulit dan sebaliknya.

Model Rasch adalah suatu formulasi matematika yang menghubungkan kemungkinan hasil (benar atau salah untuk butir soal yang dikotomi) jika seseorang menjawab butir soal sesuai dengan karakteristik peserta tes dan butir soal tersebut. Apabila kemampuan peserta tes lebih tinggi dari tingkat kesukaran soal, maka nilai kemungkinan peserta tes tersebut menjawab soal dengan benar lebih besar dari 0,5. Sebaliknya apabila kemampuan peserta tes lebih rendah dari tingkat kesukaran butir, maka kemungkinan orang tersebut menjawab soal dengan benar kurang dari 0,5. Jika kemampuan orang sama persis dengan tingkat kesukaran butir soal, maka nilai kemungkinan orang tersebut menjawab soal dengan benar adalah 0,5 (Bahrul Hayat, 2000: 5).

b) Model Logistik Dua Parameter

Parameter yang digunakan dalam model ini ada dua yaitu tingkat kesukaran butir soal dan daya pembeda. Dugaan atau *guessing* butir dianggap sama dengan nol. Hambleton, Swaminathan, & Rogers (1991: 15) merumuskan model logistik dua parameter secara matematis sebagai berikut.

$$P_i(\theta) = \frac{e^{Da_i(\theta-b_i)}}{1+e^{Da_i(\theta-b_i)}} \quad (4)$$

Keterangan:

- $P_i(\theta)$ = probabilitas menjawab benar butir i oleh peserta dengan kemampuan θ
- θ = kemampuan peserta tes
- a = daya pembeda butir i
- b_i = tingkat kesukaran butir i
- D = suatu konstanta faktor penskalaan bernilai 1,7.

c) Model Logistik Tiga Parameter

Model logistik tiga parameter menggunakan parameter tingkat kesukaran butir, daya pembeda dan *guessing* (tebakan semu). Hambleton, Swaminathan, & Rogers (1991: 17) model ini secara matematis dirumuskan sebagai berikut.

$$P_i(\theta) = c_i + (1 - c_i) \frac{e^{Da_i(\theta - b_i)}}{1 + e^{Da_i(\theta - b_i)}} \quad (5)$$

Keterangan:

- $P_i(\theta)$ = probabilitas menjawab benar butir i oleh peserta dengan kemampuan θ
 θ = kemampuan peserta tes
 a_i = daya pembeda butir i
 b_i = tingkat kesukaran butir i
 c_i = parameter kebetulan menjawab benar (besarnya tebakan semu).
 D = suatu konstanta faktor penskalaan bernilai 1,7.

B. Kerangka Berpikir

Belum ada penelitian yang menunjukkan bahwa kegiatan strategi itu memang benar yakni dengan cara mengubah nomor urut dan atau urutan pilihan jawaban akan diperoleh tes yang paralel. Oleh karena itu penelitian ini dilakukan untuk membuktikan bahwa dengan cara mengubah urutan nomor butir soal dan atau urutan pilihan jawaban akan diperoleh tes yang paralel. Perubahan urutan nomor butir dan peletakan pilihan jawaban untuk mata pelajaran Matematika tingkat SMP. Selanjutnya dari hasil analisis perubahan urutan nomor butir dan peletakan pilihan jawaban akan diketahui bagaimana tingkat ksukasar pada masing-masing pengacakan tersebut.

C. Pertanyaan Penelitian

Berdasarkan tujuan penelitian, kajian teori yang relevan, dan kerangka berpikir, pertanyaan penelitian ini adalah sebagai berikut.

1. Apakah ada pengaruh tingkat kesukaran butir soal sebelum dilakukan pengacakan urutannomor butir soal dengan setelah dilakukan pengacakan urutannomor butir soal 1-40 berdasarkan teori tes klasik?

2. Apakah ada pengaruh tingkat kesukaran butir soal sebelum dilakukan pengacakan urutan nomor butir soal dengan setelah dilakukan pengacakan urutan nomor butir 1-20 dan 21-40 berdasarkan teori tes klasik?
3. Apakah ada pengaruh tingkat kesukaran butir soal sebelum dilakukan pengacakan urutan nomor butir soal dengan setelah dilakukan pengacakan urutan nomor butir 1-10, 11-20, dan 21-30, dan 31-40 berdasarkan teori tes klasik?
4. Apakah ada pengaruh perbedaan tingkat kesukaran butir soal sebelum dilakukan pengacakan urutan pilihan jawab (A, B, C, D) dengan setelah dilakukan pengacakan urutan pilihan jawaban (D, C, B, A) berdasarkan teori tes klasik?
5. Apakah ada pengaruh tingkat kesukaran butir dan parameter kemampuan sebelum dilakukan pengacakan urutan nomor butir soal dengan setelah dilakukan pengacakan urutan nomor butir soal 1-40 berdasarkan teori respon butir?
6. Apakah ada pengaruh tingkat kesukaran butir dan parameter kemampuan sebelum dilakukan pengacakan urutan nomor butir soal dengan setelah dilakukan pengacakan urutan nomor butir 1-20 dan 21-40 berdasarkan teori respon butir?
7. Apakah ada pengaruh tingkat kesukaran butir dan parameter kemampuan sebelum dilakukan pengacakan urutan nomor butir soal dengan setelah dilakukan pengacakan urutan nomor butir 1-10, 11-20, dan 21-30, dan 31-40 berdasarkan teori respon butir?
8. Apakah ada pengaruh perbedaan tingkat kesukaran butir dan parameter kemampuan sebelum dilakukan pengacakan urutan pilihan jawab (A, B, C, D) dengan setelah dilakukan pengacakan urutan pilihan jawaban (D, C, B, A) berdasarkan teori respon butir?
9. Teknik pengacakan soal manakah yang paralel diantara pengacakan urutan yang sudah dilakukan?

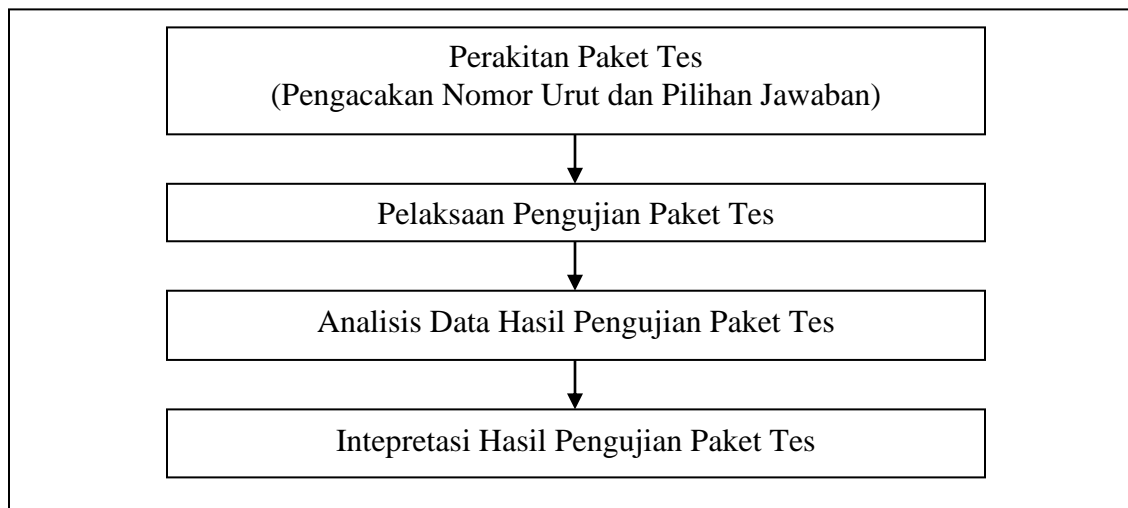
BAB III METODE PENELITIAN

A. Jenis Penelitian

Penelitian ini termasuk dalam jenis penelitian eksperimen dengan pendekatan kuantitatif. Penelitian ini menggunakan *one group pre test post test design* untuk membandingkan kemampuan siswa dan melihat tingkat keparalelan soal dengan menggunakan beberapa paket soal yang telah diacak nomor urut dan pilihan jawabannya.

B. Prosedur Penelitian

Pada penelitian ini, soal Matematika yang akan diujikan dilakukan pengacakan pada urutan nomor soal dan penempatan jawaban untuk membuktikan apakah soal tersebut benar-benar paralel. Soal yang telah dirakit menjadi 5 paket soal diujikan pada siswa SMP terpilih sebanyak 640 peserta didik dari SMP N 5 dan SMP N 8 Yogyakarta. Berdasarkan hasil analisis kemudian dilihat kemampuan siswa setelah mengerjakan soal yang telah mengalami perubahan urutan butir dan peletakan jawaban. Prosedur penelitian ini dinyatakan pada Gambar 1.



Gambar 1. Prosedur Penelitian

C. Metode Pengumpulan Data dan Instrumen yang Digunakan

Metode pengumpulan data dilakukan dengan teknik tes. Tes dilakukan dengan menggunakan instrumen tes soal Matematika SMP kelas IX untuk mata pelajaran Matematika sebanyak 40 butir soal pilihan ganda yang sudah dilakukan

pengacakan pada nomor butir soal dan pilihan jawaban. Urutan butir dan penempatan jawaban menjadi 5 paket yakni sebagai berikut.

Tabel 3. Paket Soal dan Jenis Pengacakan

Paket Soal	Jenis Pengacakan
1	Soal tanpa pengacakan
2	Pengacakan nomor 1 -40
3	Pengacakan nomor 1-20 dan 21-40
4	Pengacakan nomor 1-10, 11-20, 21-30 dan 31-40
5	Pengacakan pilihan jawaban

Paket 1 merupakan soal Matematika SMP kelas IX asli tanpa dilakukan pengacakan pada nomor urut soal maupun pilihan jawaban. Paket 2 merupakan soal dengan pengacakan secara menyeluruh dari nomor 1-40, yaitu dengan membalik urutan nomor 40 menjadi nomor 1, nomor 39 menjadi nomor 2, nomor 38 menjadi nomor 3 dan seterusnya. Paket 3 merupakan soal dengan pengacakan pada nomor 1-20 dan 21-40, yaitu dengan membalik nomor soal 20 menjadi nomor 1, nomor 19 menjadi nomor 2, dan seterusnya. Kemudian untuk nomor soal 40 menjadi nomor 21, nomor 39 menjadi nomor 2 dan seterusnya. Paket 4 merupakan soal dengan pengacakan pada nomor 1-10, 11-20, 21-30 dan 31-40, yaitu dengan membalik nomor soal 10 menjadi nomor 1, nomor 9 menjadi nomor 2, dan seterusnya. Kemudian untuk nomor soal 20 menjadi nomor 11, nomor 19 menjadi nomor 2 dan seterusnya. Untuk nomor soal 30 menjadi nomor 21, nomor 29 menjadi 22, demikian juga untuk pengacakan nomor 31-40. Paket 5 merupakan soal dengan pengacakan pilihan jawaban, dimana pilihan jawaban d menjadi a, c menjadi b, b menjadi c dan a menjadi d.

D. Teknik Analisis Data

Hasil pengujian pada peserta didik selanjutnya dilakukan analisis untuk melihat apakah pengacakan nomor soal dan pilihan jawaban benar-benar menghasilkan tes yang paralel. Data berupa respons jawaban siswa dianalisis menggunakan teori tes klasik (*classical test theory*) dan teori respons butir (*item response theory*). Analisis dengan pendekatan tes klasik dilakukan dengan bantuan program *QUEST* untuk melihat tingkat kesukaran butir. Pada analisis teori respons

butir menggunakan pendekatan 1 Parameter Logistik dengan bantuan program *QUEST* untuk mengetahui tingkat kesukaran dan parameter kemampuan siswa.

Selain dilakukan analisis dengan pendekatan teori tes klasik dan teori respons butir, hasil pengujian kemudian dihitung indeks reliabilitasnya. Indeks reliabilitas dilihat berdasarkan output hasil analisis dengan menggunakan program *QUEST*. Untuk melihat apakah ada pengaruh pengacakan nomor urut soal terhadap keparalelan tes, dilakukan dengan Uji beda dengan bantuan program SPSS. Apabila hasil analisis menunjukkan bahwa besarnya nilai Sig < 0,05 maka terdapat pengaruh pengacakan nomor urut soal keparalelan tes terhadap tingkat kesukaran soal.

BAB IV HASIL DAN PEMBAHASAN

A. Hasil Penelitian

1. Reliabilitas Tes

Hasil tes pada kelima paket soal kemudian dilihat indeks reliabilitasnya. Berdasarkan output hasil analisis dengan program *QUEST* diketahui nilai indeks reliabilitas soal paket 1, paket 2 dan paket 3 besarnya sama yaitu 0,96. Paket 4 dan paket 5 memiliki indeks reliabilitas yang tidak jauh berbeda dengan indeks reliabilitas paket 1, 2 dan 3. Besarnya indeks reliabilitas paket 4 dan paket 5 adalah 0,97. Apabila dilihat besarnya indeks reliabilitas dari kelima paket soal, diketahui bahwa kelima paket soal memiliki nilai indeks reliabilitas yang baik.

2. Analisis Soal Berdasarkan Pendekatan Teori Tes Klasik

Berdasarkan pendekatan teori tes klasik, hasil analisis menggunakan program *QUEST* diperoleh nilai parameter item pada masing-masing paket soal sebagai berikut.

Tabel 4. Hasil Analisis Soal Paket 1

No Soal	Kesukaran	No Soal	Kesukaran
1	0.945	21	0.961
2	0.874	22	0.709
3	0.929	23	0.102
4	0.134	24	0.307
5	0.638	25	0.268
6	0.709	26	0.433
7	0.961	27	0.764
8	0.724	28	0.921
9	0.937	29	0.748
10	0.701	30	0.449
11	0.748	31	0.654
12	0.480	32	0.898
13	0.827	33	0.535
14	0.953	34	0.654
15	0.449	35	0.732
16	0.740	36	0.882
17	0.913	37	0.591
18	0.827	38	0.346
19	0.646	39	0.693
20	0.748	40	0.465
Rata-Rata Tingkat Kesukaran 0.675			

Soal paket 1 merupakan soal asli sebelum dilakukan pengacakan pada butir soal. Dari Tabel 4 dapat dilihat bahwa berdasarkan hasil analisis pendekatan teori tes klasik, soal paket 1 memiliki rata-rata tingkat kesukaran 0,675. Tingkat kesukaran yang baik dalam teori tes klasik adalah apabila soal memiliki indeks kesukaran kategori sedang, yaitu antara 0,3 sampai dengan 0,7 (Nitko, 1983). Secara keseluruhan apabila dilihat dari nilai rata-rata tingkat kesukaran, soal paket 1 tergolong dalam kategori baik.

Soal paket 2 merupakan soal pengacakan dimana soal nomor 1 pada paket 1 menjadi soal nomor 40 di paket 2, soal nomor 2 di paket 1 menjadi soal nomor 39 di paket 2 dan seterusnya. Hasil analisis soal paket 2 adalah sebagai berikut.

Tabel 5. Hasil Analisis Soal Paket 2

No Soal	Kesukaran	No Soal	Kesukaran
1	0.323	21	0.398
2	0.718	22	0.694
3	0.653	23	0.863
4	0.629	24	0.895
5	0.895	25	0.750
6	0.815	26	0.435
7	0.480	27	0.847
8	0.718	28	0.742
9	0.863	29	0.306
10	0.734	30	0.637
11	0.460	31	0.637
12	0.726	32	0.902
13	0.935	33	0.702
14	0.766	34	0.903
15	0.331	35	0.508
16	0.782	36	0.508
17	0.839	37	0.129
18	0.266	38	0.903
19	0.661	39	0.839
20	0.919	40	0.952
Rata-Rata Tingkat Kesukaran 0.677			

Tabel 5 menunjukkan hasil analisis pendekatan teori tes klasik pada soal paket 2. Berdasarkan Tabel 5 dapat diketahui bahwa soal paket 2 memiliki rata-rata tingkat kesukaran 0,677. Secara keseluruhan apabila dilihat dari nilai rata-rata tingkat kesukaran, soal paket 2 tergolong dalam kategori baik karena terletak pada rentang 0,3 sampai dengan 0,7 (Nitko, 1983). Nilai rata-rata tingkat kesukaran soal paket 2 apabila

dibandingkan dengan soal paket 1 tidak jauh berbeda dimana soal paket 1 memiliki nilai rata-rata tingkat kesukaran 0,675.

Soal paket 3 merupakan soal pengacakan dimana soal nomor 1 pada paket 1 menjadi soal nomor 20 di paket 3, soal nomor 2 di paket 1 menjadi soal nomor 19 di paket 3, soal nomor 21 di paket 1 menjadi soal nomor 40 di paket 3 dan seterusnya. Hasil analisis soal paket 3 adalah sebagai berikut.

Tabel 6. Hasil Analisis Soal Paket 3

No Soal	Kesukaran	No Soal	Kesukaran
1	0.651	21	0.484
2	0.881	22	0.754
3	0.619	23	0.857
4	0.921	24	0.786
5	0.675	25	0.325
6	0.460	26	0.325
7	0.079	27	0.302
8	0.770	28	0.278
9	0.873	29	0.556
10	0.889	30	0.968
11	0.714	31	0.206
12	0.683	32	0.683
13	0.833	33	0.556
14	0.817	34	0.611
15	0.754	35	0.849
16	0.468	36	0.722
17	0.897	37	0.675
18	0.730	38	0.651
19	0.516	39	0.849
20	0.641	40	0.714
Rata-Rata Tingkat Kesukaran 0.651			

Berdasarkan hasil analisis pendekatan teori tes klasik yang dijelaskan pada Tabel 6, dapat dilihat bahwa soal paket 3 memiliki rata-rata tingkat kesukaran 0,651. Secara keseluruhan apabila dilihat dari nilai rata-rata tingkat kesukaran, soal paket 3 tergolong dalam kategori baik karena terletak pada rentang 0,3 sampai dengan 0,7. Nilai rata-rata tingkat kesukaran soal paket 3 apabila dibandingkan dengan soal paket 1 tidak jauh berbeda dimana soal paket 1 memiliki nilai rata-rata tingkat kesukaran 0,675.

Soal paket 4 merupakan soal pengacakan dimana soal nomor 1 pada paket 1 menjadi soal nomor 10 di paket 4, soal nomor 2 di paket 1 menjadi soal nomor 9 di paket

4, soal nomor 11 di paket 1 menjadi soal nomor 20 di paket 4 dan seterusnya. Hasil analisis soal paket 4 adalah sebagai berikut.

Tabel 7. Hasil Analisis Soal Paket 4

No Soal	Kesukaran	No Soal	Kesukaran
1	0.954	21	0.961
2	0.868	22	0.632
3	0.829	23	0.566
4	0.330	24	0.493
5	0.533	25	0.349
6	0.638	26	0.250
7	0.967	27	0.658
8	0.638	28	0.743
9	0.914	29	0.632
10	0.651	30	0.428
11	0.638	31	0.684
12	0.408	32	0.862
13	0.697	33	0.618
14	0.941	34	0.724
15	0.474	35	0.724
16	0.737	36	0.822
17	0.875	37	0.645
18	0.816	38	0.553
19	0.645	39	0.664
20	0.586	40	0.283
Rata-Rata Tingkat Kesukaran 0.661			

Pada Tabel 7 dapat dilihat bahwa berdasarkan hasil analisis pendekatan teori tes klasik, soal paket 4 memiliki rata-rata tingkat kesukaran 0,661. Secara keseluruhan apabila dilihat dari nilai rata-rata tingkat kesukaran, soal paket 4 tergolong dalam kategori baik karena terletak pada rentang 0,3 sampai dengan 0,7 (Nitko, 1983). Nilai rata-rata tingkat kesukaran soal paket 4 apabila dibandingkan dengan soal paket 1 tidak jauh berbeda dimana soal paket 1 memiliki nilai rata-rata tingkat kesukaran 0,675.

Tabel 8. Hasil Analisis Soal Paket 5

No Soal	Kesukaran	No Soal	Kesukaran
1	0.954	21	0.961
2	0.868	22	0.632
3	0.829	23	0.566
4	0.330	24	0.493
5	0.829	25	0.349
6	0.638	26	0.250
7	0.967	27	0.658
8	0.638	28	0.743
9	0.914	29	0.632
10	0.661	30	0.428
11	0.638	31	0.684
12	0.408	32	0.862
13	0.697	33	0.618
14	0.941	34	0.724
15	0.474	35	0.724
16	0.737	36	0.822
17	0.875	37	0.645
18	0.816	38	0.553
19	0.645	39	0.664
20	0.586	40	0.283
Rata-Rata Tingkat Kesukaran 0.668			

Soal paket 5 merupakan soal dengan pengacakan pada pilihan jawaban, sedangkan urutan nomornya sama dengan urutan soal pada paket 1. Hasil analisis soal paket 5 dapat dilihat pada Tabel 8. Dapat dilihat bahwa berdasarkan hasil analisis pendekatan teori tes klasik, soal paket 5 memiliki rata-rata tingkat kesukaran 0,675. Secara keseluruhan apabila dilihat dari nilai rata-rata tingkat kesukaran, soal paket 5 tergolong dalam kategori baik karena terletak pada rentang 0,3 sampai dengan 0,7(Nitko, 1983). Nilai rata-rata tingkat kesukaran soal paket 5 apabila dibandingkan dengan paket 1 memiliki nilai rata-rata tingkat kesukaran yang tidak jauh berbeda. Nilai rata-rata tingkat kesukaran pada paket 1 adalah 0,668.

Apabila soal paket 2, paket 3, paket 4 dan paket 5 setelah diperoleh tingkat kesukaran tiap butir disesuaikan sesuai dengan urutan soal seperti pada paket 1, maka akan diperoleh nilai tingkat kesukaran seperti pada Tabel 9.

Tabel 9. Tingkat Kesukaran 5 Paket Soal Berdasarkan Teori Tes Klasik

No Soal	Tingkat Kesukaran				
	Paket 1	Paket 2	Paket 3	Paket 4	Paket 5
1	0.945	0.952	0.641	0.651	0.954
2	0.874	0.839	0.516	0.914	0.868
3	0.929	0.903	0.730	0.638	0.829
4	0.134	0.129	0.897	0.967	0.330
5	0.638	0.508	0.468	0.638	0.829
6	0.709	0.508	0.754	0.533	0.638
7	0.961	0.903	0.817	0.330	0.967
8	0.724	0.702	0.833	0.829	0.638
9	0.937	0.902	0.683	0.868	0.914
10	0.701	0.637	0.714	0.954	0.661
11	0.748	0.637	0.889	0.586	0.638
12	0.480	0.306	0.873	0.645	0.408
13	0.827	0.742	0.770	0.816	0.697
14	0.953	0.847	0.079	0.875	0.941
15	0.449	0.435	0.460	0.737	0.474
16	0.740	0.750	0.675	0.474	0.737
17	0.913	0.895	0.921	0.941	0.875
18	0.827	0.863	0.619	0.697	0.816
19	0.646	0.694	0.881	0.408	0.645
20	0.748	0.398	0.651	0.638	0.586
21	0.961	0.919	0.714	0.428	0.961
22	0.709	0.661	0.849	0.632	0.632
23	0.102	0.266	0.651	0.743	0.566
24	0.307	0.839	0.675	0.658	0.493
25	0.268	0.782	0.722	0.250	0.349
26	0.433	0.331	0.849	0.349	0.250
27	0.764	0.766	0.611	0.493	0.658
28	0.921	0.935	0.556	0.566	0.743
29	0.748	0.726	0.683	0.632	0.632
30	0.449	0.460	0.206	0.961	0.428
31	0.654	0.734	0.968	0.283	0.684
32	0.898	0.863	0.556	0.664	0.862
33	0.535	0.718	0.278	0.553	0.618
34	0.654	0.480	0.302	0.645	0.724
35	0.732	0.815	0.325	0.822	0.724
36	0.882	0.895	0.325	0.724	0.822
37	0.591	0.629	0.786	0.724	0.645
38	0.346	0.653	0.857	0.618	0.553
39	0.693	0.718	0.754	0.862	0.664
40	0.465	0.323	0.484	0.684	0.283
Rata-rata	0.675	0.677	0.651	0.661	0.668

Tabel 9 memaparkan nilai kesukaran tiap item soal pada kelima paket soal setelah nomor soal disesuaikan dengan soal paket 1. Berdasarkan Tabel 9 dapat dilihat bahwa rata-rata kesukaran dari kelima paket soal tidak jauh berbeda. Rata-rata tingkat kesukaran pada paket 1 sebesar 0,675, paket 2 sebesar 0,677, paket 3 sebesar 0,651, paket 4 sebesar 0,661 dan paket 5 sebesar 0,668. Hasil tersebut juga dapat dilihat pada hasil analisis deskriptif dengan menggunakan program SPSS sebagai berikut.

Tabel 10. Hasil Analisis Deskriptif Berdasarkan SPSS

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
PAKET_1	40	,6749	,23108	,03654	,6010	,7488	,10	,96
PAKET_2	40	,6766	,21449	,03391	,6080	,7452	,13	,95
PAKET_3	40	,6506	,21399	,03383	,5821	,7190	,08	,97
PAKRT_4	40	,6608	,18879	,02985	,6004	,7211	,25	,97
PAKET_5	40	,6684	,18944	,02995	,6078	,7290	,25	,97
Total	200	,6662	,20632	,01459	,6375	,6950	,08	,97

2. Uji Beda Berdasarkan Pendekatan Teori Tes Klasik

Untuk melihat apakah terdapat pengaruh dilakukannya pengacakan urutan soal pada paket soal, dilakukan analisis uji One Way Anova dengan program SPSS. Hasil analisis dapat dilihat pada Tabel 11.

Tabel 11. Uji Beda 5 Paket Soal Menggunakan Teori Tes Klasik

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	,018	4	,005	,107	,980
Within Groups	8,452	195	,043		
Total	8,471	199			

Berdasarkan hasil analisis SPSS pada Tabel 11, dapat dilihat bahwa nilai Sig adalah sebesar 0,980. Nilai tersebut lebih besar dari 0,05, artinya tidak terdapat perbedaan tingkat kesukaran siswa yang mengerjakan soal paket 1, paket 2, paket 3, paket 4 dan paket 5. Hal ini dapat diambil kesimpulan bahwa pengacakan urutan nomor soal tidak berpengaruh terhadap tingkat kesulitan item.

Lebih lanjut lagi dilakukan analisis untuk melihat apakah ada perbedaan cara pengacakan pada paket 2, paket 3, paket 4 dan paket 5 terhadap soal paket 1 (soal tanpa pengacakan). Hasil analisis dapat dilihat pada Tabel 12.

Tabel 12. Hasil Analisis Uji Beda Tiap Paket Soal

(I) PAKET	(J) PAKET	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
PAKET_1	PAKET_2	-,00170	,04655	1,000	-,1299	,1265
	PAKET_3	,02432	,04655	,985	-,1039	,1525
	PAKRT_4	,01413	,04655	,998	-,1141	,1423
	PAKET_5	,00648	,04655	1,000	-,1217	,1347

Tabel 12 dapat dilihat bahwa pada soal paket 2 dimana dilakukan pengacakan soal, dimana soal nomor 1 pada paket 1 menjadi soal nomor 40 pada paket 2, soal nomor 2 pada paket 1, menjadi soal nomor 39 pada paket 2 tidak terdapat perbedaan tingkat kesukaran. Artinya tidak ada pengaruh pengacakan soal pada paket 2 dengan soal tanpa pengacakan pada paket 1. Hal ini dapat dilihat berdasarkan nilai sig. pada Tabel 12 sebesar 1,000.

Pengacakan soal pada paket 3, dimana soal nomor 1 pada paket 1 menjadi soal nomor 20 pada paket 3, soal nomor 21 pada paket 1 menjadi soal nomor 40. Pada paket 3 juga tidak terdapat perbedaan tingkat kesukaran. Artinya tidak ada pengaruh pengacakan soal pada paket 3 dengan soal tanpa pengacakan pada paket 1. Hal ini dapat dilihat berdasarkan nilai sig. pada Tabel 12 sebesar 0,985.

Pengacakan soal pada paket 4, dimana dilakukan pengacakan soal nomor 1 pada paket 1 menjadi soal nomor 10 pada paket 4, soal nomor 11 pada paket 1 menjadi nomor 20 pada paket 4, soal nomor 21 pada paket 1 menjadi nomor 30 pada paket 4, dan soal nomor 31 pada paket 1 menjadi nomor 40 pada paket 5. Pada paket 4 tidak terdapat perbedaan tingkat kesukaran. Artinya tidak ada pengaruh pengacakan soal pada paket 4 dengan soal tanpa pengacakan pada paket 1. Hal ini dapat dilihat berdasarkan nilai sig. pada Tabel 12 sebesar 0,998.

Pengacakan soal paket 5, dimana dilakukan pengacakan pilihan jawaban setiap butir soal. pilihan jawaban A pada paket 1 menjadi D pada paket 5. Pada Paket 5 tidak terdapat perbedaan tingkat kesukaran. Artinya tidak ada pengaruh pengacakan pilihan jawaban pada pada paket 5 dengan soal tanpa pengacakan pilihan jawaban pada paket 1. Hal ini dapat dilihat berdasarkan nilai sig. pada Tabel 12 sebesar 1,000.

3. Analisis Soal Berdasarkan Pendekatan Teori Respons Butir

Setelah dilakukan analisis dengan pendekatan teori tes klasik, kelima paket soal kemudian dilakukan analisis dengan pendekatan teori respons butir. Analisis pendekatan teori respons butir dilakukan dengan bantuan program QUEST. Hasil analisis dengan pendekatan teori respons butir pada soal paket 1 diperoleh nilai parameter tingkat kesukaran masing-masing item soal seperti pada Tabel 13.

Tabel 13. Hasil Analisis Soal Paket 1

No Soal	Kesukaran	No Soal	Kesukaran
1	0.390	21	0.460
2	0.270	22	0.200
3	0.350	23	0.300
4	0.270	24	0.200
5	0.190	25	0.210
6	0.200	26	0.190
7	0.460	27	0.220
8	0.210	28	0.340
9	0.370	29	0.210
10	0.200	30	0.190
11	0.210	31	0.200
12	0.190	32	0.300
13	0.240	33	0.190
14	0.420	34	0.200
15	0.190	35	0.210
16	0.210	36	0.280
17	0.320	37	0.190
18	0.240	38	0.200
19	0.190	39	0.200
20	0.210	40	0.190
Rata-rata Tingkat Kesukaran = 0.250			

Soal paket 1 merupakan soal asli sebelum dilakukan pengacakan pada butir soal. Dari Tabel 13 dapat dilihat bahwa berdasarkan hasil analisis pendekatan teori tes klasik, soal paket 1 memiliki rata-rata tingkat kesukaran 0,250. Tingkat kesukaran yang baik dalam teori respons butir adalah apabila soal memiliki indeks kesukaran antara -2 sampai dengan +2 (Hambleton & Swaminathan: 1985). Secara keseluruhan apabila dilihat dari nilai rata-rata tingkat kesukaran, soal paket 1 tergolong dalam kategori baik yaitu memiliki nilai rata-rata 0,250.

Soal paket 2 merupakan soal pengacakan dimana soal nomor 1 pada paket 1 menjadi soal nomor 40 di paket 2, soal nomor 2 di paket 1 menjadi soal nomor 39 di paket 2 dan seterusnya. Hasil analisis soal paket 2 adalah sebagai berikut.

Tabel 14. Hasil Analisis Soal Paket 2

No Soal	Kesukaran	No Soal	Kesukaran
1	0.200	21	0.230
2	0.210	22	0.200
3	0.200	23	0.270
4	0.190	24	0.300
5	0.300	25	0.220
6	0.240	26	0.190
7	0.420	27	0.260
8	0.210	28	0.210
9	0.270	29	0.200
10	0.210	30	0.200
11	0.190	31	0.200
12	0.210	32	0.310
13	0.370	33	0.200
14	0.220	34	0.310
15	0.200	35	0.190
16	0.230	36	0.190
17	0.250	37	0.280
18	0.210	38	0.310
19	0.200	39	0.250
20	0.340	40	0.420
Rata-Rata Tingkat Kesukaran 0.245			

Dari Tabel 14 dapat dilihat bahwa berdasarkan hasil analisis pendekatan teori tes klasik, soal paket 2 memiliki rata-rata tingkat kesukaran 0,245. Secara keseluruhan apabila dilihat dari nilai rata-rata tingkat kesukaran, soal paket 2 tergolong dalam kategori baik karena terletak pada rentang -2 sampai dengan $+2$ (Hambleton & Swaminathan:1985). Nilai rata-rata tingkat kesukaran soal paket 2 apabila dibandingkan dengan soal paket 1 tidak jauh berbeda dimana soal paket 1 memiliki nilai rata-rata tingkat kesukaran 0,50.

Soal paket 3 merupakan soal pengacakan dimana soal nomor 1 pada paket 1 menjadi soal nomor 20 di paket 3, soal nomor 2 di paket 1 menjadi soal nomor 19 di paket 3, soal nomor 21 di paket 1 menjadi soal nomor 40 di paket 3 dan seterusnya. Hasil analisis soal paket 3 adalah sebagai berikut.

Tabel 15. Hasil Analisis Soal Paket 3

No Soal	Kesukaran	No Soal	Kesukaran
1	0.200	21	0.190
2	0.290	22	0.220
3	0.200	23	0.270
4	0.340	24	0.230
5	0.200	25	0.210
6	0.190	26	0.210
7	0.360	27	0.210
8	0.220	28	0.220
9	0.280	29	0.190
10	0.290	30	0.520
11	0.210	31	0.240
12	0.210	32	0.210
13	0.250	33	0.190
14	0.240	34	0.200
15	0.220	35	0.260
16	0.190	36	0.210
17	0.300	37	0.200
18	0.220	38	0.200
19	0.190	39	0.260
20	0.200	40	0.210
Rata-Rata Tingkat Kesukaran 0.236			

Dari Tabel 15 dapat dilihat bahwa berdasarkan hasil analisis pendekatan teori tes klasik, soal paket 3 memiliki rata-rata tingkat kesukaran 0,236. Nilai rata-rata tingkat kesukaran soal paket 3 apabila dibandingkan dengan soal paket 1 tidak jauh berbeda dimana soal paket 1 memiliki nilai rata-rata tingkat kesukaran 0,250.

Soal paket 4 merupakan soal pengacakan dimana soal nomor 1 pada paket 1 menjadi soal nomor 10 di paket 4, soal nomor 2 di paket 1 menjadi soal nomor 9 di paket 4, soal nomor 11 di paket 1 menjadi soal nomor 20 di paket 4 dan seterusnya. Hasil analisis soal paket 4 adalah sebagai berikut.

Tabel 16. Hasil Analisis Soal Paket 4

No Soal	Kesukaran	No Soal	Kesukaran
1	0.390	21	0.420
2	0.250	22	0.180
3	0.220	23	0.180
4	0.470	24	0.170
5	0.170	25	0.180
6	0.180	26	0.200
7	0.460	27	0.180
8	0.180	28	0.200
9	0.300	29	0.180
10	0.180	30	0.180
11	0.180	31	0.190
12	0.180	32	0.240
13	0.190	33	0.180
14	0.350	34	0.190
15	0.170	35	0.190
16	0.190	36	0.220
17	0.250	37	0.180
18	0.220	38	0.170
19	0.180	39	0.180
20	0.180	40	0.190
Rata-Rata Tingkat Kesukaran 0.222			

Dari Tabel 16 dapat dilihat bahwa berdasarkan hasil analisis pendekatan teori tes klasik, soal paket 4 memiliki rata-rata tingkat kesukaran 0,222. Secara keseluruhan apabila dilihat dari nilai rata-rata tingkat kesukaran, soal paket 4 tergolong dalam kategori baik karena terletak pada rentang -2 sampai dengan +2 (Hambleton & Swaminathan:1985). Nilai rata-rata tingkat kesukaran soal paket 4 apabila dibandingkan dengan soal paket 1 tidak jauh berbeda dimana soal paket 1 memiliki nilai rata-rata tingkat kesukaran 0,250.

Soal paket 5 merupakan soal dengan pengacakan pada pilihan jawaban, sedangkan urutan nomornya sama dengan urutan soal pada paket 1. Hasil analisis soal paket 5 dapat dilihat pada Tabel 17. Dapat dilihat bahwa berdasarkan hasil analisis pendekatan teori respons butir, soal paket 5 memiliki rata-rata tingkat kesukaran 0,222.

Secara keseluruhan apabila dilihat dari nilai rata-rata tingkat kesukaran, soal paket 5 tergolong dalam kategori baik karena terletak pada rentang -2 sampai dengan +2(Hambleton& Swaminathan:1985). Nilai rata-rata tingkat kesukaran soal paket 5 apabila dibandingkan dengan paket 1 memiliki nilai rata-rata tingkat kesukaran yang tidak jauh berbeda. Nilai rata-rata tingkat kesukaran pada paket 1 adalah 0,250.

Tabel 17. Hasil Analisis Soal Paket 5

No Soal	Kesukaran	No Soal	Kesukaran
1	0.390	21	0.420
2	0.250	22	0.180
3	0.220	23	0.180
4	0.470	24	0.170
5	0.170	25	0.180
6	0.180	26	0.200
7	0.460	27	0.180
8	0.180	28	0.200
9	0.300	29	0.180
10	0.180	30	0.180
11	0.180	31	0.190
12	0.180	32	0.240
13	0.190	33	0.180
14	0.350	34	0.190
15	0.170	35	0.190
16	0.190	36	0.220
17	0.250	37	0.180
18	0.220	38	0.170
19	0.180	39	0.180
20	0.180	40	0.190
Rata-Rata Tingkat Kesukaran 0.222			

Apabila soal paket 2, paket 3, paket 4 dan paket 5 setelah diperoleh tingkat kesukaran tiap butir disesuaikan sesuai dengan urutan soal seperti pada paket 1, maka akan diperoleh nilai tingkat kesukaran seperti pada Tabel 18.

Tabel 18. Tingkat Kesukaran 5 Paket Soal Berdasarkan Teori Respon Butir

No Soal	Tingkat Kesukaran				
	Paket 1	Paket 2	Paket 3	Paket 4	Paket 5
1	0.390	0.420	0.200	0.180	0.390
2	0.270	0.250	0.190	0.300	0.250
3	0.350	0.310	0.220	0.180	0.220
4	0.270	0.280	0.300	0.460	0.470
5	0.190	0.190	0.190	0.180	0.170
6	0.200	0.190	0.220	0.170	0.180
7	0.460	0.310	0.240	0.470	0.460
8	0.210	0.200	0.250	0.220	0.180
9	0.370	0.310	0.210	0.250	0.300
10	0.200	0.200	0.210	0.390	0.180
11	0.210	0.200	0.290	0.180	0.180
12	0.190	0.200	0.280	0.180	0.180
13	0.240	0.210	0.220	0.220	0.190
14	0.420	0.260	0.360	0.250	0.350
15	0.190	0.190	0.190	0.190	0.170
16	0.210	0.220	0.200	0.170	0.190
17	0.320	0.300	0.340	0.350	0.250
18	0.240	0.270	0.200	0.190	0.220
19	0.190	0.200	0.290	0.180	0.180
20	0.210	0.230	0.200	0.180	0.180
21	0.460	0.340	0.210	0.180	0.420
22	0.200	0.200	0.260	0.180	0.180
23	0.300	0.210	0.200	0.200	0.180
24	0.200	0.250	0.200	0.180	0.170
25	0.210	0.230	0.210	0.200	0.180
26	0.190	0.200	0.260	0.180	0.200
27	0.220	0.220	0.200	0.170	0.180
28	0.340	0.370	0.190	0.180	0.200
29	0.210	0.210	0.210	0.180	0.180
30	0.190	0.190	0.240	0.420	0.180
31	0.200	0.210	0.520	0.190	0.190
32	0.300	0.270	0.190	0.180	0.240
33	0.190	0.210	0.220	0.170	0.180
34	0.200	0.420	0.210	0.180	0.190
35	0.210	0.240	0.210	0.220	0.190
36	0.280	0.300	0.210	0.190	0.220
37	0.190	0.190	0.230	0.190	0.180
38	0.200	0.200	0.270	0.180	0.170
39	0.200	0.210	0.220	0.240	0.180
40	0.190	0.200	0.190	0.190	0.190
Rata-rata	0.250	0.245	0.236	0.222	0.222

Tabel 18 memaparkan nilai kesukaran tiap item soal pada kelima paket soal setelah nomor soal disesuaikan dengan soal paket 1. Dapat dilihat bahwa rata-rata kesukaran dari kelima paket soal tidak jauh berbeda. Rata-rata tingkat kesukaran pada paket 1 sebesar 0,250, paket 2 sebesar 0,245, paket 3 sebesar 0,236, paket 4 sebesar 0,222 dan paket 5 sebesar 0,222. Hasil tersebut juga dapat dilihat pada hasil analisis deskriptif dengan menggunakan SPSS seperti berikut.

Tabel 19. Hasil Analisis Deskriptif Berdasarkan SPSS

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
PAKET_1	40	,2503	,07934	,01254	,2249	,2756	,19	,46
PAKET_2	40	,2453	,06198	,00980	,2254	,2651	,19	,42
PAKET_3	40	,2363	,06196	,00980	,2164	,2561	,19	,52
PAKET_4	40	,2222	,08103	,01281	,1963	,2482	,17	,47
PAKET_5	40	,2222	,08103	,01281	,1963	,2482	,17	,47
Total	200	,2353	,07380	,00522	,2250	,2455	,17	,52

4. Uji Beda Berdasarkan Teori Respons Butir

Untuk melihat apakah terdapat pengaruh dilakukannya pengacakan urutan soal pada paket soal, dilakukan analisis uji One Way Anova dengan program SPSS. Hasil analisis dapat dilihat pada Tabel 20.

Tabel 20. Uji Beda 5 Paket Soal Menggunakan Teori Respon Butir

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	,027	4	,007	1,225	,302
Within Groups	1,057	195	,005		
Total	1,084	199			

Berdasarkan hasil analisis SPSS pada Tabel 20, dapat dilihat bahwa nilai Sig adalah sebesar 0,302. Nilai tersebut lebih besar dari 0,05, artinya tidak terdapat perbedaan kesukaran siswa yang mengerjakan soal paket 1, paket 2, paket 3, paket 4 dan paket 5. Hal ini dapat diambil kesimpulan bahwa pengacakan urutan nomor soal tidak berpengaruh terhadap tingkat kesulitan item.

Lebih lanjut lagi dilakukan analisis untuk melihat apakah ada perbedaan cara pengacakan pada paket 2, paket 3, paket 4 dan paket 5 terhadap soal paket 1 (soal tanpa pengacakan). Hasil analisis dapat dilihat pada Tabel 21.

Tabel 21. Hasil Analisis Uji Beda Tiap Paket Soal

(I) PAKET	(J) PAKET	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
PAKET_1	PAKET_2	,00500	,01646	1,000	-,0417	,0517
	PAKET_3	,01400	,01646	1,000	-,0327	,0607
	PAKET_4	,02800	,01646	,906	-,0187	,0747
	PAKET_5	,02800	,01646	,906	-,0187	,0747

Tabel 12 dapat dilihat bahwa pada soal paket 2 dimana dilakukan pengacakan soal, dimana soal nomor 1 pada paket 1 menjadi soal nomor 40 pada paket 2, soal nomor 2 pada paket 1, menjadi soal nomor 39 pada paket 2 tidak terdapat perbedaan tingkat kesukaran. Artinya tidak ada pengaruh pengacakan soal pada paket 2 dengan soal tanpa pengacakan pada paket 1. Hal ini dapat dilihat berdasarkan nilai sig. pada Tabel 21 sebesar 1,000.

Untuk pengacakan soal pada paket 3, dimana soal nomor 1 pada paket 1 menjadi soal nomor 20 pada paket 3, soal nomor 21 pada paket 1 menjadi soal nomor 40. Pada paket 3 juga tidak terdapat perbedaan tingkat kesukaran. Artinya tidak ada pengaruh pengacakan soal pada paket 3 dengan soal tanpa pengacakan pada paket 1. Hal ini dapat dilihat berdasarkan nilai sig. pada Tabel 21 sebesar 1,000.

Pengacakan soal pada paket 4, dimana dilakukan pengacakan soal nomor 1 pada paket 1 menjadi soal nomor 10 pada paket 4, soal nomor 11 pada paket 1 menjadi nomor 20 pada paket 4, soal nomor 21 pada paket 1 menjadi nomor 30 pada paket 4, dan soal nomor 31 pada paket 1 menjadi nomor 40 pada paket 5. Pada paket 4 tidak terdapat perbedaan tingkat kesukaran. Artinya tidak ada pengaruh pengacakan soal pada paket 4 dengan soal tanpa pengacakan pada paket 1. Hal ini dapat dilihat berdasarkan nilai sig. pada Tabel 21 sebesar 0,906.

Pada pengacakan soal paket 5, dimana dilakukan pengacakan pilihan jawaban setiap butir soal. pilihan jawaban A pada paket 1 menjadi D pada paket 5. Pada Paket 5 tidak terdapat perbedaan tingkat kesukaran. Artinya tidak ada pengaruh pengacakan pilihan jawaban pada pada paket 5 dengan soal tanpa pengacakan pilihan jawaban pada paket 1. Hal ini dapat dilihat berdasarkan nilai sig. pada Tabel 21 sebesar 0,906.

B. Pembahasan

1. Teori Tes Klasik

Bermutu atau tidaknya butir tes dapat diketahui dari tingkat kesukaran yang dimiliki oleh masing-masing butir tersebut. Butir tes dikatakan baik jika butir tersebut tidak terlalu sukar dan tidak pula terlalu mudah, atau dengan kata lain tingkat kesukaran butir adalah sedang atau cukup. Tingkat kesukaran butir biasanya dikaitkan dengan tujuan tes (Mehrens & Lehmann, 1973: 195; Retnawati, 2012). Dari hasil analisis diketahui untuk paket 1 dari 40 butir beradapada rentang 0,102 sampai 0,961. Paket 2 dari 40 butir secara keseluruhan dalam kategori baik karena berada pada rentang 0,129 sampai 0,952. Paket 3 dari 40 butir berada pada rentang 0,206 sampai 0,968. Paket 4 dari 40 butir berada pada rentang 0,250 sampai 0,961. Paket 5 dari 40 butir berada pada rentang 0,250 sampai 0,967. Sesuai dengan pendapat (Azwar, 2002: 137), besarnya tingkat kesukaran item terletak antara 0 sampai 1. Semakin besar tingkat kesukaran, maka semakin mudah butir tes tersebut. Sebaliknya semakin kecil tingkat kesukaran, maka butir tes akan semakin sukar. Hal senada juga diutarakan oleh Anastasi & Urbina (1997: 173) bahwa pada pengukuran butir tes, tingkat kesukaran butir berhubungan dengan presentase orang-orang yang dapat menjawab soal dengan benar. Tingkat kesukaran ini merupakan proporsi jumlah peserta tes yang menjawab benar dari suatu butir soal, yaitu perbandingan antara jumlah peserta tes seluruhnya.

Analisis Uji beda dapat dilihat Berdasarkan hasil analisis SPSS nilai Sig adalah sebesar 0,980. Nilai tersebut lebih besar dari 0,05, artinya tidak terdapat perbedaan tingkat kesukaran siswa yang mengerjakan soal paket 1, paket 2, paket 3, paket 4 dan paket 5. Hal ini dapat diambil kesimpulan bahwa pengacakan urutan nomor soal tidak berpengaruh terhadap tingkat kesulitan item. Pada soal paket 2 tidak ada pengaruh pengacakan soal pada paket 2 dengan soal tanpa pengacakan pada paket 1. Hal ini dapat dilihat berdasarkan nilai sig. sebesar 1,000. Untuk pengacakan soal pada paket 3, dimana tidak terdapat perbedaan tingkat kesukaran. Artinya tidak ada pengaruh pengacakan soal pada paket 3 dengan soal tanpa pengacakan pada paket 1. Hal ini dapat dilihat berdasarkan nilai sig. sebesar 0,985. Pengacakan soal pada paket 4, tidak terdapat perbedaan tingkat kesukaran. Artinya tidak ada pengaruh pengacakan soal pada paket 4 dengan soal tanpa pengacakan pada paket 1. Hal ini dapat dilihat berdasarkan nilai sig. sebesar 0,998. Pada pengacakan soal paket 5, tidak terdapat perbedaan tingkat kesukaran.

Artinya tidak ada pengaruh pengacakan pilihan jawaban pada pada paket 5 dengan soal tanpa pengacakan pilihan jawaban pada paket 1. Hal ini dapat dilihat berdasarkan nilai sig. sebesar 1,000.

2. Teori Respon Butir

Matematika merupakan salah satu mata pelajaran yang diujikan dalam ujian nasional sekolah menengah pertama (SMP). Menurut Syukrul Hamdi dkk (2018) kemampuan matematika siswa dapat digunakan untuk menyelesaikan berbagai masalah dan kesulitan yang mereka hadapi dalam mempelajari berbagai ilmu, terutama ilmu alam. Hal itu menjadi dasar pentingnya matematika, sehingga menjadi salah satu mata pelajaran yang diujikan dalam ujian nasional. Soal tes yang diujikan dalam ujian nasional matematika terdiri dari beberapa paket soal. Paket tersebut disusun dengan butir soal yang sama, namun dilakukan pengacakan nomor urut soal dan pilihan jawaban untuk membedakan antar paket soal.

Penelitian ini dilakukan dengan menganalisis lima paket soal yang berbeda berdasarkan pengacakan soal. Apabila didasarkan pada hasil analisis dengan pendekatan teori respons butir, semua soal pada paket 1, paket 2, paket 3, paket 4 dan paket 5 tergolong pada kategori soal yang baik. Hal itu dapat dilihat dari nilai tingkat kesukaran (b) yang terletak antara -2 sampai dengan 2. Hal ini senada dengan Anesa, S. & Aman. (2016), dalam penelitiannya menyatakan bahwa tingkat kesukaran soal dikatakan baik apabila memiliki nilai antara -2 sampai 2. Dari 40 butir soal pada kelima paket soal memiliki nilai tingkat kesukaran yang termasuk kategori baik.

Untuk melihat apakah terdapat pengaruh pengacakan pada kelima paket soal, dilakukan uji beda dengan menggunakan SPSS. Berdasarkan hasil analisis SPSS dihasilkan nilai Sig sebesar 0,302. Nilai tersebut lebih besar dari 0,05, artinya tidak terdapat perbedaan kesukaran siswa yang mengerjakan soal paket 1, paket 2, paket 3, paket 4 dan paket 5. Hal ini dapat diambil kesimpulan bahwa pengacakan urutan nomor soal tidak berpengaruh terhadap tingkat kesulitan item pada pendekatan teori respons butir.

BAB V KESIMPULAN

A. Kesimpulan

Berdasarkan hasil penelitian dan pembahasan, penelitian ini menyimpulkan bahwa tidak ada pengaruh pengacakan urutan nomor butir soal dan penempatan pilihan jawaban terhadap tingkat kesukaran butir. Secara rinci kesimpulan tersebut dapat dituliskan sebagai berikut.

1. Tidak ada pengaruh pengacakan urutan nomor butir soal terhadap tingkat kesukaran butir. Hasil analisis menggunakan analisis varians menunjukkan bahwa tidak ada perbedaan yang signifikan tingkat kesukaran butir antara **paket 1** (soal tanpa pengacakan), **paket 2** (pengacakan nomor butir 1-40), **paket 3** (pengacakan nomor butir 1-20 dan 21-40), dan **paket 4** (pengacakan nomor 1-10, 11-20, 21-30 dan 31-40).
2. Tidak ada pengaruh penempatan pilihan jawaban terhadap tingkat kesukaran butir. Hasil analisis menggunakan analisis varians menunjukkan bahwa tidak ada perbedaan yang signifikan tingkat kesukaran butir antara **paket 1** soal tanpa pengacakan) dan **paket 5** (pengacakan penempatan pilihan jawaban).

B. Saran

1. Apabila ingin dibuat tes paralel maka butir tersebut harus dibuat berdasarkan indikator yang sama.
2. Apabila diinginkan tes paralel maka dapat dibuat dengan cara pengacakan pada nomor butir soal, dengan syarat tidak mengurutkan nomor butir soal berdasarkan tingkat kesukarannya.
3. Apabila diinginkan tes paralel maka dapat dibuat dengan cara melakukan pengacakan penempatan pilihan jawaban, dengan syarat tidak mengurutkan nomor butir soal berdasarkan tingkat kesukarannya.

DAFTAR PUSTAKA

- Allen, M.J., & Yen, W.M. (1979). *Introduction to measurement theory*. California: Brooks/Cole Publishing Company Wardsworth, Inc.
- Anastasi, A. & Urbina, S. (2008). *Psychological testing*. Indiana: Prentice Hall, Inc.
- Anesa, S. & Aman. (2016). Developing formative authentic assessment instrument based on learning trajectory for elementary school. *Research and Evaluation in Education*, 2, 13-24
- Badrun Kartowagiran. (Oktober 2012). *Penulisan butir soal*. Makalah disajikan dalam Pelatihan penulisan dan analisis butir soal bagi Sumber daya PNS Dik-Rekinpeg, di Hotel Kawanua Aerotel, Jakarta.
- Crocker, L., & Algina, J. (2008). *Introduction to classical and modern test theory*. New York: Holt, Reinhart, and Winston, Inc.
- Djemari Mardapi. (2008). *Teknik penyusunan instrumen tes dan nontes*. Yogyakarta: Mitra Cendekia.
- _____. (2014). *Pengukuran, penilaian, dan evaluasi pendidikan*. Yogyakarta: Nuha Litera.
- Fernandes, H.J.X. (1984). *Testing and measurement*. Jakarta: National Education Planning, Evaluation, and Curriculum Development
- Heri Retnawati. (2014). *Teori respon butir dan penerapannya (untuk peneliti, praktisi pengukuran, dan pengujian mahasiswa pascasarjana)*. Yogyakarta: Nuha Medika.
- Linn, R.L. (1989). *Education measurement (3th ed.)*. New York: MacMillan Publishing Company.
- Mehrens, W.A., & Lehmann, J.L. (1973). *Measurement and evaluation in education and psychology*. New York: Holt, Rinehart, and Winston, Inc.
- Moore, B., Stanly, T. 2010. *Critical thinking and formative assessments*. Larchmount, NY: Eye On Education, Inc
- Nitko, A. (1983). *Educational test and measurement: an introduction*. New York: Harcourt Brace Jovanovich, Inc.
- Rahma Zulaiha. (2008). *Bagaimana menganalisis soal dengan program ITEMAN*. Jakarta: Depdiknas-Balitbang-Puspendik.

Reynolds, C.R., Livingston, R.B., & Wilson, V. (2009). *Measurement and assessment in education (2nd ed)*. Boston: Pearson Education Inc.

Saifuddin Azwar. (2013). *Reliabilitas dan validitas (Edisi keempat)*. Yogyakarta: Pustaka Pelajar.

Sumarna Surapranata. (2005a). *Panduan penulisan tes tertulis (penilaian berbasis kelas)*. Bandung: Remaja Rosdakarya

Syukrul Hamdi, dkk (2018). Developing a testlet model for mathematics at elementary level. *International Journal of Instruction*, 11, 375-390.

Thomas, A. dan Thorne, G. (2007). *Higher Order Thinking*. Center for Development and learning. Diambil dari CDL pda tanggal 6 Agustus 2011.

BAB VI
PERSONALIA PENELITIAN

Personalia dan atau organisasi tim peneliti dalam penelitian ini adalah sebagai berikut.

No	Nama	Jabatan dalam Tim	Tugas Penelitian
	NIP/NIM	Alokasi Waktu: Jam /Minggu	
1	Prof. Dr. Badrun Kartowagiran	Ketua Pelaksana Kegiatan	Koordinator, pengembang instrumen, analisis, dan pelaporan
	NIP.19530725 197811 1 001	10 jam/minggu	
2	Prof. Djemari Mardapi, Ph.D	Anggota Peneliti	Pengembang instrumen analisis
	NIP.19470101 197412 1 001	8 jam/minggu	
3	Dr. Harun Al-Rasyid, M.Pd	Anggota Peneliti	Pengembang, Pengambil Data, pengurus administrasi.
	NIP. 19560727 198503 1 024	8 jam/minggu	
4	Danang Hari Kristiyanto	Anggota Mahasiswa	pengambilan data, entry data.
	NIM. 16701261017	5 jam/minggu	
5	Kriswantoro	Anggota Mahasiswa	pengambilan data, entry data.
	NIM. 17701261018	5 jam/minggu	
6	Dian Normalitasari Purnama	Anggota Mahasiswa	pengambilan data, entry data.
	NIM. 17701261020	5 jam/minggu	

BAB VIII
JADWAL PENELITIAN

Jadwal kerja penelitian hibah pascasarjana adalah sebagai berikut.

No	JENIS AKTIVITAS	Bulan, 2018											
		1	2	3	4	5	6	7	8	9	10	11	12
1.	Persiapan												
2.	Mengkaji prosedur yang sudah ada												
3.	Pembuatan proposal												
4.	Membuat instrumen dan rancangan analisis												
5.	Merevisi instrumen dan rancangan analisis												
6.	Ujicoba												
7.	Analisis data												
8.	Penyusunan laporan												
9.	Seminar dan penyerahan laporan												

Lampiran 1. Lembar Telaah Butir Soal Pilihan Ganda

a)	Aspek materi
	(1) Soal sesuai dengan indikator; (2) Distraktor berfungsi; (3) Hanya ada satu kunci jawaban yang paling tepat
b)	Aspek konstruksi
	(1) Pokok soal dirumuskan dengan singkat, jelas dan tegas; (2) Rumusan pokok soal dan pilihan jawaban merupakan pertanyaan yang diperlukan (3) Pokok soal tidak memberi petunjuk ke kunci jawaban; (4) Pokok soal bebas dari pernyataan yang bersifat negatif ganda; (5) Gambar, grafik, tabel, diagram, wacana, dan sejenisnya yang terdapat pada soal jelas dan berfungsi; (6) Panjang pilihan jawaban relatif sama; (7) Pilihan jawaban tidak menggunakan pernyataan “Semua jawaban di atas salah” atau “Semua pilihan jawaban di atas benar” dan sejenisnya; (8) Pilihan jawaban yang berbentuk angka atau waktu harus disusun berdasarkan urutan besar kecilnya angka tersebut atau kronologis; (9) Butir-butir soal tidak bergantung pada jawaban soal sebelumnya;
c)	Aspek bahasa/budaya
	(1) Menggunakan bahasa yang sesuai dengan kaidah bahasa Indonesia; (2) Menggunakan bahasa yang komunikatif; (3) Tidak menggunakan bahasa yang berlaku setempat (bias budaya); (4) Pilihan jawaban tidak mengulang kata/kelompok kata yang sama.

LAMPIRAN 2. BIODATA KETUA DAN ANGGOTA

1. Curriculum Vitae Ketua Tim Peneliti

I. IDENTITAS DIRI

1.1	Nama Lengkap dan Gelar	Prof. Dr. Badrun Kartowagiran/Kartowagiran		
1.2	Jabatan Fungsional/Gol	Guru Besar, TMT 1 Agust 2012/ IV d, TMT: 1 April 2015		
1.3	NIP/NIK	19530725 197811 1 001		
1.4	NIDN	0025075303		
1.5	Tempat dan Tanggal Lahir	Bantul, 25 Juli 1953		
1.6	Alamat Rumah	Gejayan, JL. Mangga, Gang Apel 101, RT 07, RW 31, Condong Catur, Depok, Sleman, Yogyakarta 55283		
1.7	Nomor Telp/Fax	(0274) 881523		
1.8	Nomor HP	08122781548		
1.9	Alamat Kantor	Universitas Negeri Yogyakarta (UNY), Karangmalang, Yogyakarta, 55281		
1.10	Nomor Telp/Fax	(0274) 520326/550835		
1.11	Alamat e-mail	badrunkw@yahoo.com		
1.12	Lulusan yang telah dihasilkan	S1= 56 orang	S2= 30orang	S3= 12 orang
1.13	Matakuliah yang diampu	1. Asesmen pembelajaran 2. Evaluasi program pendidikan 3. Statistik		

II. RIWAYAT PENDIDIKAN

2.1	Program	S1	S2	S3
2.2	Nama PT	IKIP Yogyakarta	IKIP Jakarta	UGM
2.3	Bidang Ilmu/ Keahlian	Pendidikan Teknik Mesin	Penelitian dan Evaluasi Pendidikan	Psikologi/ Psikometri
2.4	Tahun Masuk	1973	1987	1998
2.5	Tahun Lulus	1977	1992	2005
2.6	Judul Skripsi/Tesis/Disertasi	Hasil belajar Mata pelajaran Pewartat Pengangkat siswa STM Muhammadiyah Bantul	Kemampuan mengajar lulusan FPTK IKIP Yogyakarta	Perbandingan berbagai metode deteksi bias

2.7	Nama Pembimbing/Promotor	Dr. Sukamto	1. Prof. Dr. Sutrisno Hadi, MA 2. Dr. Sukamto	1. Prof. Dr. Sumadi Suryabrata 2. Prof. Djemari Mardapi, PhD 3. Jahja Umar, PhD
-----	--------------------------	-------------	--	---

III. PENGALAMAN MENGAJAR

Tahun	Mata Kuliah	Jenjang	Semester	Prodi/Fakultas/PT
1978 – 2004	Praktik Bengkel Teknik Mesin	S1	Gasal danganap	Diknik Mesin FT UNY
1978 – 2004	Mekanika Teknik	S1	Gasal dan genap	Diknik Mesin FT UNY
2005 - sekarang	Statistik	D3	Genap	Diknik Mesin FT UNY
2005 - sekarang	Statistik	S1	Genap	Diknik Mesin FT UNY
2005 - sekarang	Metodologi Penelitian Pendidikan	S1	Gasal	Diknik Mesin FT UNY
2011 - sekarang	Evaluasi Pembelajaran	S1	Genap	Diknik Mesin FT UNY
2005 - Sekarang	Metodologi Penelitian Pendidikan	S2	Gasal	PEP-PPS UNY
2005 - 2013	Konstruksi Instrumen	S2	Genap	PEP-PPS UNY
2007 – sekarang	Evaluasi Pembelajaran	S2	Gasal dan Genap	Dikdas dan PTK PPS UNY
2007 – sekarang	Evaluasi Program	S2	Genap	PEP-PPS UNY
2007 – sekarang	Praktik Evaluasi	S2	Gasal	PEP-PPS UNY
2011 - sekarang	Evaluasi Kebijakan	S2	Gasal	PEP-PPS UNY
2007 – sekarang	Statistik	S2	Gasal	Dikdas PPS-UNY
2011 – sekarang	Statistik	S2	Genap	PLS PPS-UNY
2007 – sekarang	Statistik: SEM	S3	Gasal	PEP-PPS UNY
2007 – sekarang	Konstruksi Instrumen	S3	Genap	PEP-PPS UNY
2007 – 2010	Praktik Evaluasi	S3	Gasal	PEP-PPS UNY
2007 – sekarang	Statistik	S2	Gasal	Psikologi, Pascasarjana UAD
2007 – sekarang	Konstruksi instrumen	S2	Gasal	Psikologi, Pascasarjana UGM
2011 - sekarang	Seminar Psikometrik	S3	Genap	Psikologi,

				Pascasarjana UGM
--	--	--	--	---------------------

IV. PENGALAMAN PENELITIAN

No	Tahun	Judul	Status dalam penelitian	Sumber Dana
1	2015	Model asesmen autentik untuk menilai hasil belajar siswa SMP (Tahun ke 1)	Ketua	Dikti
2	2015	Kajian peluang merekrut guru produktif dari industri dan profesional berdasarkan KKNI	Ketua	Dit PSMK
3	2015	Monitoring dan evaluasi pendampingan kurikulum di SMK	Ketua	Dit PSMK
4	2015	Penyusunan Naskah Ujian	Ketua	
6	2014	Monev implementasi kurikulum tahun 2013 di SMK	Ketua	Dit PSMK
7	2014	Peminatan SMK	Ketua	Dit PSMK
8	2014	Model evaluasi pembelajaran di SD yang menerapkan kurikulum tahun 2013	Ketua	PPS UNY
9	2013	Pengembangan rambu-rambu pelaporan hasil belajar siswa SMK	Ketua	Dit PSMK
10	2013	Pengembangan rambu-rambu penyusunan perangkat penilaian hasil belajar di SMK	Ketua	Dit PSMK
11	2013	Hibah Pascasarjana Tahun ke 3: Model Evaluasi Kinerja Guru	Ketua	Dikti
12	2013	Hibah Pascasarjana Tahun ke 3: Model Penjaminan mutu sekolah	Anggota	Dikti
13	2013	Evaluasi kesiapan SMP di DI Y dalam mengimplementasikan kurikulum tahun 2013	Ketua	Pascasarjana UNY
14	2012	Hibah Pascasarjana Tahun ke 2: Model Evaluasi Kinerja Guru	Ketua	Dikti
15	2012	Hibah Pascasarjana Tahun ke 2: Model Penjaminan mutu sekolah	Anggota	Dikti
16	2012	Evaluasi standar Proses Pembelajaran dan Standar Penilaian di SMK D.I. Yogyakarta	Ketua	Pascasarjana UNY
17	2011	Hibah Pascasarjana Tahun ke 1: Model Evaluasi Kinerja Guru	Ketua	Dikti
16	2011	Hibah Pascasarjana Tahun ke 1: Model Penjaminan mutu sekolah	Anggota	Dikti

17	2011	Evaluasi Dampak Ujian Nasional terhadap kualitas pendidikan	Ketua	Pascasarjana UNY
----	------	---	-------	------------------

18	2011	Pemetaan daerah berdasar daya serap UN	Anggota	Litbang, Diknas
19	2010	Hibah Pascasarjana Tahun ke 2: Pengembangan bank soal berbasis guru	Anggota	Dikti
20	2010	Hibah Pascasarjana Tahun ke 2: Pengembangan soal yang baku dan nir bias	Anggota	Dikti
	2010	Evaluasi kinerja lulusan Prodi Penelitian dan Evaluasi Pendidikan	Anggota	UNY

V. PENGALAMAN PENGABDIAN KEPADA MASYARAKAT

No	Tahun	Judul Kegiatan	Sumber
1	2015	Workshop pengembangan perangkat penilaian otentik di SD di Kab Sleman	PPS-UNY
2	2014	Workshop pengembangan perangkat penilaian otentik di SMPN 3 Wonosari	SMPN 3 Wonosari
3	2014	Sosialisasi tentang sertifikasi bagi guru dalam jabatan: Pelaksanaan PLPG tahun 2014 di Universitas Lambung Mangkurat (Unlam) Banjarmasin	Universitas Jambi
4	2014	Workshop pengembangan perangkat pembelajaran pada SMK dengan kurikulum 2013	Dikspora DIY
5	2014	Workshop penyusunan soal bagi pengembang soal Badan Kepegawaian Negara Jakarta	BKN Jkt
6	2013	Pelatihan penyusunan soal bagi guru SD di D.I. Yogyakarta	UNY
7	2013	Workshop standar isi psikometrik	Fak Psikologi UGM
8	2013	Pelatihan penulisan butir soal UN SMK	Dit PSMK JKT
9	2012	Pelatihan penulisan butir soal UN SMK	Dit PSMK JKT
10	2012	Pelatihan penulisan kisi-kisi soal UN SMK	Dit PSMK JKT
11	2011	Pelatihan penulisan soal pilihan ganda bagi guru SD	UNY
12	2011	Penyegaran penyusunan soal bagi dosen Farmasi UGM	Fak Farmasi UGM
13	2010	Pelatihan penyusunan kisi-kisi dan butir soal bagi guru Matematik SMP	UNY
14	2010	Pelatihan penulisan soal bagi guru SD di	UNY

		Kabupaten Sleman	
15	2010	Kiat meningkatkan skor Ujian Nasional	UNY

VI. PENGALAMAN PENULISAN ARTIKEL ILMIAH DALAM JURNAL

No	Tahun	Judul Artikel Ilmiah	Volume/ Nomor	Nama Jurnal
1	2015	Evaluation model for Islamic education learning in Junior High School and its significance to students' behaviours	Vol 3, No. 8, 990-995	American Journal of Educational Research, 2015
2	2014	Pengembangan model evaluasi program layanan pendidikan anak usia dini	Jurnal HEPI/ Tahun 18 Nomor 1, 2014	Jurnal HEPI/ Pascasarjana UNY
3	2014	Pengembangan instrumen penilaian kinerja guru Sekolah Dasar	Jurnal HEPI/ Tahun 18 Nomor 2, 2014	Jurnal HEPI/ Pascasarjana UNY
4	2011	Pengembangan instrumen pengukur hasil belajar NIR bias dan terskala baku	Jurnal HEPI/ Tahun 15 Nomor 2, 2011	Jurnal HEPI/ Pascasarjana UNY
5	2011	Rintisan bank soal berbasis kinerja guru untuk mata pelajaran Bahasa Indonesia SMP di D.I. Yogyakarta	Cakrawala Pendidikan November 2011, Ta hun XXX, No. 2	Cakrawala Pendidikan/ Lembaga Penelitian
6	2011	Kinerja guru profesional (pasca sertifikasi)	Cakrawala Pendidikan November 2011, Ta hun XXX, No. 3	Cakrawala Pendidikan/ Lembaga Penelitian
7	2008	Uji unidimensionalitas soal UAN Matematika SMP Tahun 2007	Volume 40, Nomor 1, Tahun 2010	Jurnal HEPI/ Pascasarjana UNY
8	2007	Validitas prediktif tes masuk SMP di D.I. Yogyakarta		Jurnal HEPI/ Pascasarjana UNY

9	2006	Hubungan subtes kemampuan verbal, kuantitatif, penalaran dengan TPA untuk calon mahasiswa non-reguler		Jurnal HEPI/ Pascasarjana UNY
10	2005	Analisis kritis terhadap ujian akhir nasional		Dinamika/ Diknik Mesin
11	1999	Pendekatan kualitatif dan kuantitatif dalam penelitian tindakan di bidang psikologi		Kontroversi/ Univ Malang

VII. SEMINAR

No	Tahun	Judul Kegiatan	Penyelenggara	Panitia/Peserta/ Pembicara
1	2015	Konferensi internasional: <i>The Evaluation model of the Islamic study learning in Junior High School and its correlation to students' behavior</i>	Higher Education Forum: Tokyo, Jepang	Pembicara
2	2014	Konferensi internasional: <i>An evaluation model of certified teachers' performance</i>	Burapha University: Thailand	Pembicara
3	2014	Seminar Nasional: Pengembangan instrumen untuk menilai kinerja guru praktikum SMA	HEPI, Makasar	Pembicara
4	2013	Seminar Nasional: Pengembangan instrumen penilaian kinerja guru praktikum	HEPI, Manado	Pembicara
5	2012	Seminar Nasional: Model penilaian kinerja guru	Pascasarjana UNY	Pembicara
6	2012	Seminar Nasional: Strategi peningkatan kompetensi guru	Fakultas Teknik UNY	Pembicara
7	2011	Seminar Nasional: Materi sulit pada soal Ujian Nasional mata pelajaran Bahasa Indonesia SMP	Lemlit UNY	Pembicara
8	2010	Seminar Regional: Peningkatan kualitas soal uji kompetensi guru	Unnes, Semarang	Pembicara
9	2010	Seminar Regional: Peningkatan kualitas pembelajaran dalam PLPG melalui peningkatan kualitas soal uji kompetensi guru	Uhamka, Jakarta	Pembicara
10	2010	Seminar Nasional: Revitalisasi guru melalui sertifikasi guru	Teacher Development	Pembicara

			Centre (TDC) Surakarta	
--	--	--	---------------------------	--

11	2010	Seminar Nasional: Penjaminan dan peningkatan mutu pendidik dan tenaga kependidikan	LPMP Kalimantan Tengah	Pembicara
12	2010	Seminar Nasional: Identifikasi Bias Butir Perangkat UN Matematika SMP 2003 Menggunakan Teori Respons Multidimensi	HEPI Jakarta	Pembicara

VIII. PELATIHAN PROFESIONAL

No	Tahun	Jenis Pelatihan (Dalam/Luar Negeri)	Penyelenggara	Jangka Waktu
1	2014	Metodologi penelitian kualitatif (luar negeri, Auckland University, New Zealand)	Auckland University, New Zealand	6 hari
2	2012	Bank Soal (Luar Negeri, Pearson New York, USA)	Pearson New York, USA)	6 hari
3	2009	Psikometri (dalam negeri, Pelatih dr Utrech University, Belanda)	Pascasarjana, UNY bekerja sama dengan Utrech University, Belanda	4 hari
4	2008	Psikometri (dalam negeri, Pelatih dr Massachusetts University, USA)	Pascasarjana, UNY bekerja sama dengan Massachusetts University, USA	4 hari
5	2008	Research by project (luar negeri, RMIT Melbourne, Australia)	RMIT Melbourne, Australia	7 hari
6	2004	Statistik Lanjut: Structural Equation Modeling (SEM), dalam negeri	Unair, Surabaya	3 hari
7	1997	Penelitian Tindakan (luar negeri, Deakin University, Melbourne, Australia)	Deakin University Boorwud, Australia	3 bulan

8	1993	Penelitian dan Evaluasi Pendidikan (luar negeri, Innotech, SEAMEO, Philipina)	Innotech, SEAMEO, Philipina)	3 bulan
---	------	---	------------------------------	---------

IX. PENGALAMAN PENULISAN BUKU

No	Tahun	Judul Buku	Jumlah Halaman	Penerbit
1	-	-	-	-

X. PENGALAMAN PEROLEHAN HKI

No	Tahun	Judul/Tema HKI	Jenis	No Pendaftaran/ Setifikat
1	-	-	-	-

XI. PENGALAMAN RUMUSAN KEBIJAKAN PUBLIK/REKAYASA SOSIAL LAINNYA

No	Tahun	Judul/Tema/Jenis Rekayasa Sosial Lainnya yang Telah Diterapkan	Tempat Penerapan	Respon Masyarakat
1	2006	Sebagai anggota Tim yang menyusun: 1. Buku 2: Petunjuk Teknis Pelaksanaan Sertifikasi 2. Buku 3: Pedoman Penyusunan Portofolio 3. Buku 4:Rambu-Rambu Pelaksanaan Pendidikan dan Latihan Profesi Guru (PLPG)	1. Digunakan oleh Universitas penyelenggara sertifikasi guru di Indonesia 2. Digunakan oleh guru se Indonesia	Sangat memerlukan buku-buku panduan sertifikasi itu
2	2007	Sama dengan tahun 2006	Sama dengan tahun 2006	Sama dengan tahun 2006
		Sebagai Ketua Tim Adhoc BSNP yang menyusun draf:Permendiknas Nomor 20 Tahun 2007 Tentang Standar Penilaian	Seluruh sekolah di seluruh Indonesia	Sangat memerlukan
3	2008	Sama dengan tahun 2006	Sama dengan tahun 2006	Sama dengan tahun 2006
4	2009	Sama dengan tahun 2006	Sama dengan tahun 2006	Sama dengan tahun 2006
5	2010	Sama dengan tahun 2006	Sama dengan	Sama dengan

			tahun 2006	tahun 2006
--	--	--	------------	------------

No	Tahun	Judul/Tema/Jenis Rekayasa Sosial Lainnya yang Telah Diterapkan	Tempat Penerapan	Respon Masyarakat
6	2011	Sebagai KETUA Tim yang menyusun: 1. Buku 2: Petunjuk Teknis Pelaksanaan Sertifikasi 2. Buku 3: Pedoman Penyusunan Portofolio 3. Buku 4:Rambu-Rambu Pelaksanaan Pendidikan dan Latihan Profesi Guru (PLPG)	1. Digunakan oleh Universitas penyelenggara sertifikasi guru di Indonesia 2. Digunakan oleh guru se Indonesia	Sangat memerlukan buku-buku panduan sertifikasi itu
7	2012	Sama dengan tahun 2011	Sama dengan tahun 2011	Sama dengan tahun 2011
8	2013	Sama dengan tahun 2006	Sama dengan tahun 2006	Sama dengan tahun 2006
		Sebagai Anggota Tim Adhoc BSNP yang menyusun draf:Permendikbud Nomor 65 Tahun 2014 Tentang Standar Penilaian	Seluruh sekolah di seluruh Indonesia	Sangat memerlukan
		Sebagai Ketua Tim yang menyusun: Rambu-rambu penulisan perangkat penilaian hasil belajar siswa SMK	Seluruh SMK di seluruh Indonesia	Sangat memerlukan
9	2014	Sama dengan tahun 2006	Sama dengan tahun 2006	Sama dengan tahun 2006
		Sebagai KETUA TIM yang menyusun: Rambu-rambu peminatan SMK	Seluruh SMK di seluruh Indonesia	Sangat memerlukan

XII. PAKAR/KONSULTAN/STAF AHLI/NARASUMBER

No	Nama Kegiatan	Nama Lembaga	Waktu	Tingkat (lokal,nasional, internasional)
1	Workshop: Penyusunan kisi-kisi dan soal tes Ujian Nasional (UN) SMK	P4TK Matematika	2013	Nasional
2	Workshop: Evaluasi	Sekolah Tinggi	2013	Nasional

	Kurikulum Sekolah Tinggi Agama Budha	Agama Budha		
--	--------------------------------------	-------------	--	--

No	Nama Kegiatan	Nama Lembaga	Waktu	Tingkat (lokal, nasional, internasional)
3	Workshop Penyusunan RPP pada kurikulum 2013 di SMKN 3 Magelang	SMKN 3 Magelang	2013	Lokal
4	Workshop: Penyusunan soal bagi karyawan Badan Kepegawaian Negara (BKN)	Badan Kepegawaian Negara (BKN), Jakarta	2013	Nasional
5	Workshop: Penyusunan Standar Isi Fakultas Psikologi UGM	Fakultas Psikologi UGM	2013	Lokal
6	Workshop: Penilaian berbasis kriteria	Fakultas Psikologi UGM	2012	Lokal
7	Workshop:Pemanfaatan hasil penilaian	Fakultas Teknik UNY	2012	Lokal
8	Workshop: Metodologi penelitian	Fakultas Psikologi UGM	2012	Nasional
9	Workshop: Pengembangan instrumen evaluasi berbasis tes	FKIP Universitas Ahmad Dahlan	2012	Lokal
10	Workshop: Sistem penilaian di RSBI	Universitas Ahmad Dahlan	2011	Lokal
11	Workshop: Penyusunan bahan ajar pada PLPG sertifikasi guru dalam jabatan	Universitas Islam Negeri (UIN) Sunan Kalijaga Yogyakarta	2011	Lokal
12	Workshop: Pengembangan kurikulum diklat	PPPPTK Kesenian	2010	Nasional
13	Workshop penyiapan uji kompetensi dalam Pendidikan Profesi Guru di Universitas PGRI Yogyakarta	Universitas PGRI Yogyakarta, Hotel Saphir Yogyakarta	2010	Lokal
14	Workshop penyiapan Uji kompetensi dalam Pendidikan Profesi Guru di Universitas Negeri	Universitas Negeri Yogyakarta	2010	Lokal

	Yogyakarta			
--	------------	--	--	--

XIII. KEANGGOTAAN PADA ORGANISASI/ASOSIASI PREFESI DAN ILMIAH

NO	Nama Organisasi	Periode	Keanggotaan	Keterangan
1	Ikatan Sarjana Pendidikan Indonesia	1980 - sekarang	Anggota	Nasional
2	Ikatan Alumni Pasca Sarjana IKIP YK	1994- sekarang	Anggota	Nasional
3	Perhimpunan Indonesia untuk Pengembangan Kreativitas (PIPK)	1995 - sekarang	Anggota	Nasional
4	Himpunan Evaluasi Pendidikan Indonesia (HEPI)	2000 – 2010/ sekarang	Sekjen/anggota	Nasional
5	Asosiasi Mahasiswa dan Alumni Program Pascasarjana UNY	2000 - 2008	Ketua	Nasional
6	KAGAMA	2005 - sekarang	Anggota	Nasional

XIV. VISITING PROFESSOR/LECTURING

No	Nama Dosen	Nama Lembaga	Waktu	Tingkat (lokal,nasional, internasional)
1	Badrun Kartowagiran, Prof. Dr	Sekolah Pascasarjana Universitas Gadjah Mada Yogyakarta	September 2015 s/d Jan 2016	Nasional
2	Badrun Kartowagiran, Prof. Dr	Sekolah Pascasarjana Universitas Ahmad Dahlan Yogyakarta	September 2015 s/d Jan 2016	Nasional

XV. PENCAPAIAN PRESTASI/REPUTASI DOSEN

No	Nama Dosen	Prestasi yang dicapai	Waktu Pencapaian	Tingkat (lokal,nasional, internasional)
1	Badrun Kartowagiran	Satya Lancana Kesetiaan 20 Tahun dari Presiden R.I.	2003	Nasional
2	Badrun Kartowagiran	Satya Lancana Karya Satya 30 Tahun dari Presiden R.I.	2009	Nasional

Semua data yang saya isikan dan tercantum dalam biodata ini adalah benar dan dapat dipertanggungjawabkan secara umum. Apabila dikemudian hari ternyata dijumpai ketidaksesuaian dengan kenyataan, saya sanggup menerima resikonya.

Demikian biodata ini saya buat dengan sebenarnya.

Yogyakarta, 6 April 2016

Dosen,



Badrun Kartowagiran

NIP. 19530725 197811 1 001

2. Curriculum Vitae Anggota Tim Peneliti

I. IDENTITAS DIRI

1.1	Nama Lengkap dan Gelar	Prof. Djemari Mardapi, Ph.D.		
1.2	Jabatan Fungsional	Guru Besar		
1.3	NIP/NIK	19470101 197412 1 001		
1.4	NIDN	0001014705		
1.5	Tempat dan Tanggal Lahir	Binjai, 1 Januari 1947		
1.6	Alamat Rumah	Jalan Garuda No. 13, Gejayan, Condongcatur, Depok, Sleman, Yogyakarta		
1.7	Nomor Telp/Fax	0274 880928		
1.8	Nomor HP	08122952895		
1.9	Alamat Kantor	Jalan Kolombo No.1 Karangmalang Yogyakarta		
1.10	Nomor Telp/Fax	0274 – 550835		
1.11	Alamat e-mail	djemarimardapi@gmail.com		
1.12	Lulusan yang telah dihasilkan	S1= 120 orang	S2= 100 orang	S3= 30 orang
1.13	Matakuliah yang diampu	<ol style="list-style-type: none"> 1. Statistik 2. Teori Statistik dan Sampling 3. Teori Tes Klasik 4. Teori Respons Butir 5. Statistika Multivariat 6. Analisis Faktor 7. Analisis Regresi 8. Pengembangan Teori Pengukuran 		

II. RIWAYAT PENDIDIKAN

2.1	Program	S1	S2	S3
2.2	Nama PT	IKIP Yogyakarta	IKIP Jakarta di Yogyakarta	The University of Iowa, Iowa City USA
2.3	Bidang Ilmu/Keahlian	Teknik Listrik	Penelitian dan Evaluasi Pendidikan	Educational Measurement and Statistics
2.4	Tahun Masuk	1967	1982	1984
2.5	Tahun Lulus	1973	1984	1988
2.6	Judul Skripsi/Tesis/Disertasi	Amplitudo Modulation dan Pembelajrannya	Faktor-faktor yang mempengaruhi hasil belajar mahasiswa	<i>Validity Generalization for Indonesian University Selection Test</i>

2.7	Nama Pembimbing/ Promotor	Drs. Matoyib	Dr. Gary Theisen	Dr. Stephen Dunbar
-----	------------------------------	--------------	------------------	--------------------

III. PENGALAMAN PENELITIAN

No	Tahun	Judul Penelitian	Pendanaan	
			Sumber	Jumlah (Rp)
1.	2012	Pengembangan Sistem Pengujian Hasil Belajar Berbantuan Komputer (<i>Computerized Adaptive Testing</i>)	DIPA UNY	25 juta
2.	2012	Penentuan Kriteria Ketuntasan Minimal (KKM) di Sekolah Menengah Kejuruan (SMK)	PPs UNY	15 juta
3.	2013	<i>Standard setting</i> matematika untuk siswa SMP	PPs UNY	25 juta
4.	2013	Model Penentuan Kriteria Ketuntasan Minimal (KKM) Berbasis Peserta Didik	Hibah Penelitian Pascasarjana	90 juta
5.	2013	Evaluasi kurikulum program studi penelitian dan evaluasi pendidikan program S3 Pascasarjana UNY	S3	15 juta
6.	2014	Rintisan ujian sekolah dengan <i>computer based testing</i>	PPs UNY	15 juta
7.	2014	Kajian Awal Pelaksanaan Ujian Nasional Teori Kejuruan Sekolah Menengah Kejuruan Secara Online	Direktorat SMK	50 juta
8.	2014	Peminatan SMK	Dit PSMK	200 juta
9.	2015	Pengembangan instrumen nontes Penilaian karya kreasi seni rupa Terintegrasi dengan karakter peserta Didik sekolah menengah kejuruan untuk Kesiapan kerja	Penelitian Unggulan Perguruan Tinggi	100 juta

IV. PENGALAMAN PENGABDIAN KEPADA MASYARAKAT

No	Tahun	Judul pengabdian pada masyarakat	Pendanaan	
			Sumber	Jumlah (Rp)
1	2015	Pelatihan penyusunan soal ujian sekolah SMP	FT UNY	5 juta
2	2015	Pelatihan penyusunan soal matematika ujian sekolah SMP	PPs UNY	10 juta
3	2014	Workshop pengembangan perangkat pembelajaran pada SMK dengan kurikulum 2013	Disdikpora DIY	5 juta

4	2014	Workshop penyusunan soal bagi pengembang soal Badan Kepegawaian Negara Jakarta	BKN	90 juta
5	2014	Pembinaan Kemampuan Melakukan Penilaian Otentik bagi Guru SMP di Kabupaten Gunung Kidul		12,5 juta
6	2013	Pelatihan penyusunan soal bagi guru SMP di Kabupaten Sleman Yogyakarta	DIPA UNY	12,5 juta
7	2012	Pelatihan penyusunan soal ujian nasional bagi guru SMP	DIPA UNY	12,5 juta

V. PENGALAMAN PENULISAN ARTIKEL ILMIAH DALAM JURNAL

No	Tahun	Judul Artikel Ilmiah	Volume/ Nomor	Nama Jurnal
1	2015	Menentukan Kriteria Ketuntasan Minimal Berbasis Peserta Didik	Vol 19, No 1, Juni 2015	Jurnal Penelitian dan Evaluasi Pendidikan
2	2014	Pengembangan Butir Soal Matematika SD di Kabupaten Lombok Timur Sebagai Upaya dalam Pengadaan Bank Soal	Vol 44, No 1, Mei 2014	Jurnal Kependidikan (Penerbit: LPPM UNY bekerjasama dengan MPPI)
3	2014	The Accuracy of The Fixed Parameter Calibration Method: Study of Mathematics National Examination Test	Vol 18, No 2 (2014)	Jurnal Penelitian dan Evaluasi Pendidikan
4	2014	Developing The Assessment of Learning Outcomes for The Student of Physical, Sports, and Health Education in Volleyball Game for Senior High Schools	Vol 18, No 1 (2014)	Jurnal Penelitian dan Evaluasi Pendidikan
5	2014	The Standard Setting Method for The National Examination in The Elementary School	Vol 18, No 1 (2014)	Jurnal Penelitian dan Evaluasi Pendidikan
6	2014	Developing An Evaluation Instruments of The Office Administration Expertise Competency Test in Vocational High Schools	Vol 18, No 1 (2014)	Jurnal Penelitian dan Evaluasi Pendidikan
7	2014	An Analysis of Method of Cheating on Large Test Scale	Vol 18, No 1 (2014)	Jurnal Penelitian dan Evaluasi Pendidikan
8	2014	A Model for Assessment of Principal Instructional Leadership of Basic Education	Vol 18, No 1 (2014)	Jurnal Penelitian dan Evaluasi Pendidikan

9	2014	Developing Higher Order Thinking Skill Test of Physics (PhysTHOTS) for Senior High School Students	Vol 18, No 1 (2014)	Jurnal Penelitian dan Evaluasi Pendidikan
10	2014	An Evaluation of “School of Quality” in Primary Education in Lao PDR	Vol 5, No 1, 2014	The Educational Measurement and Evaluation Review
11	2013	Scaling Classical Theory of Multiple Intelligences Classical Instrument Type Thurstone and Likert	Vol 17, No 2 (2013)	Jurnal Penelitian dan Evaluasi Pendidikan
12	2013	Comparison of Standard Setting Method for Determining Minimum Mastery Criteria	Vol 17, No 2 (2013)	Jurnal Penelitian dan Evaluasi Pendidikan
13	2013	Pengembangan Model Penilaian Akhlak Peserta Didik Madrasah Aliyah	Vol. 17, No. 2, 2013	Jurnal Penelitian dan Evaluasi Pendidikan
14	2012	Pengujian Hasil Belajar dan Penilaian Pendidikan Berbantuan Komputer	Vol 42, No 2, 2012	Jurnal Kependidikan (Penerbit: LPPM UNY bekerjasama dengan MPPI)

VI. SEMINAR

No	Tahun	Judul Kegiatan	Penyelenggara	Panitia/Peserta/ Pembicara
1	2015	<i>International Conference on Education, Psychology, and Society (ICEEPS)</i>	Tokyo, Jepang	Pembicara
2	2015	Workshop Nasional “Pengembangan Standar, Kurikulum, Sertifikat KOMPETENSI Pendidik Pendidikan Teknologi dan Vokasional”	FT UNY bekerjasama dengan BELMAWA DIKTI	Narasumber
3	2015	Seminar Nasional Konferensi HEPI Tahun 2015 “Peranan Asesmen dan Ujian dalam Peningkatan Mutu Pendidikan Nasional”	Makassar, Sulawesi Selatan	Pemakalah
4	2014	Seminar Nasional Kimia dan Pendidikan Kimia VI dengan tema “Pemantapan Riset Kimia dan Asesmen dalam Pembelajaran Berbasis Pendekatan Saintifik”	Universitas Sebelas Maret Surakarta	Pembicara
5	2014	Seminar Hasil Penelitian Puspendik Tahun 2014: Meningkatkan Mutu	Kemdikbud Balitbang	Peserta

		Pendidikan dengan Penilaian Berkualitas	Puspendik	
6	2014	Konferensi Ilmiah HEPI “Asesmen untuk Pendidikan Berkualitas”	HEPI UKD Bali	Pemakalah
7	2014	<i>The 27th ICSEI Conference: Redefining Education, Learning, and Teaching in the 21st. Century: The Past, Present and Future of Sustainable School Effectiveness</i>	UNY	Peserta
8	2014	Seminar Nasional “Pengembangan 5M dalam Pembelajaran sebagai Implementasi Kurikulum 2013”	Kerjasama UNY dengan LPPMP Kalimantan Selatan	Pembicara
9	2014	Workshop Peningkatan Kompetensi SDM Pengelola Bank Soal	Badan Kepegawaian Negara	Pembicara
10	2014	Workshop Pengayaan Soal dengan Perguruan Tinggi Negeri	Badan Kepegawaian Negara	Pembicara
11	2014	<i>The 22nd International Conference on Research and Measurement of Thailand: New Conception of Educational Research, Measurement, and Statistics for ASEAN</i>	Burappa University Bangkok Thailand	Pembicara
12	2013	Seminar Nasional Matematika dan Pendidikan Matematika dengan tema “Pengembangan Kompetensi Guru Matematika dalam Rangka Menyongsong Implementasi Kurikulum 2013”	Universitas Sebelas Maret Surakarta	Pembicara
13	2013	Konferensi Ilmiah Nasional HEPI “Evaluasi Implementasi Kurikulum 2013 dan Sistem Penilaiannya”	HEPI	Pemakalah
14	2013	Seminar Satu Dasawarsa APSI dengan tema “Kita Bangun Jati Diri Pengawas Sekolah/Madrasah dan Kita Songsong Berlakunya Kurikulum 2013”	Asosiasi Pengawas Sekolah Indonesia (APSI) DIY	Pembicara
15	2013	Seminar dan FGD Penyelenggaraan Pendidikan Guru: Dari Rintisan Menuju Pembakuan	Dirjen Dikti Kemdikbud	Peserta
16	2013	<i>Focused Group Discussion (FGD) Penentuan Batas Lulus UN Berbasis Peserta Didik</i>	PPs UNY	Pembicara
17	2012	International Seminar of “Reformulation of Technology and	APTEKINDO	Pembicara

		Vocational Education Paradigms”		
18	2012	Seminar Regional Pendidikan “Strategi Meningkatkan Profesionalisme Guru”	Pusaka Pendidikan dan UKMF KMIP FIP UNY	Pembicara

VII. PELATIHAN PROFESIONAL

No	Tahun	Jenis Pelatihan (Dalam/Luar Negeri)	Penyelenggara	Jangka Waktu
1	2015	<i>Quality Assurance Monitor TIMSS-R Study</i>	TIMSS and PIRLS International Study Center	16 hari

VIII. PENGALAMAN PENULISAN BUKU

No	Tahun	Judul Buku	Jumlah Halaman	Penerbit
1	2007	Teknik Penilaian Instrumen Tes dan Nontes	166	Mitra Cendekia Press
2	2012	Pengukuran Penilaian dan Evaluasi Pendidikan	247	Nuha Medika Yogyakarta

IX. PAKAR/ KONSULTAN/ STAF AHLI/ NARASUMBER

No	Nama Kegiatan	Nama Lembaga	Waktu	Tingkat (lokal,nasional, internasional)
1.		Badan Standar Nasional Pendidikan	2005 sampai 2014	Nasional
2.	Seminar Regional Pendidikan : Strategi Meningkatkan Profesionalisme Guru	UKMF KMIP FIP UNY	2012	Lokal
3.	International Seminar of “REformulation of Technology and Vocational Education Paradigms	The Indonesian Association of Technology Vocational Education	2012	International
4.		Universitas	2013	Nasional

		Negeri Padang		
5.	Seminar Satu Dasawarsa APSI “Kita Bangun Jatidiri Pengawas Sekolah/Madrasah dan Kita Songsong Berlakunya Kurikulum 2013”	Asosiasi Pengawas Sekolah Indonesia DIY	2013	Nasional
6.	Focus Group Discussion (FGD) Penentuan Batas Lulus UN berbasis Peserta Didik	PEP PPS UNY	2013	Lokal
7.		Universitas Muhammadiyah Surakarta	2014	Nasional
8.		Universitas Negeri Surakarta	2014	Nasional
9.	Seminar Assessment Otentik: Implementasi Asesmen Otentik	PEP PPS Universitas Negeri Semarang	2014	Nasional
10.		Universitas Hamka Jakarta	2014	Nasional
11.	Seminar Nasional: Pengembangan 5M dalam Pembelajaran sebagai Implementasi Kurikulum 2013	LPMP Kalimantan Selatan	2014	Nasional
12.	Workshop Peningkatan Kompetensi SDM pengelola Bank Soal	Kantor Regional I BKN Yogyakarta	2014	Nasional
13.	Workshop Pengayaan Soal dengan Perguruan Tinggi Negeri	UNY	2014	Nasional
14.		Badan Pemeriksa Keuangan RI Pusat Pendidikan dan Pelatihan	2014	Nasional
15.		BAN PT	2014	Nasional
16.		Universitas Sarjanawiyata Tamansiswa	2014	Nasional
17.		Universitas Muhammadiyah	2015	Nasional

		Prof. Dr. Hamka		
18.		FT UNY	2015	Nasional
19.		Balitbang Kemdikbud	2015	Nasional
20.		HEPI di Makassar	2015	Nasional

X. KEANGGOTAAN PADA ORGANISASI/ASOSIASI PREFESI DAN ILMIAH

No	Nama Dosen	Nama Lembaga	Waktu	Tingkat (lokal,nasional, internasional)
1	Djemari Mardapi	American Education Research Association (AERA)	2005 sampai sekarang	Internasional
2	Djemari Mardapi	Himpunan Evaluasi Pendidikan Indonesia (HEPI)	2005 sampai sekarang	Nasional
3	Djemari Mardapi	World Bank Consultant for Strategic Planning for Education in Indonesia	2005 sampai sekarang	Internasional
4	Djemari Mardapi	Anggota badan Standar Nasional Pendidikan (BSNP)	2005 sampai sekarang	Nasional
5	Djemari Mardapi	<i>TIMSS and PIRLS International Study Center</i>	2003 sampai sekarang	Internasional

XI. VISITING PROFESSOR/LECTURING

No	Nama Dosen	Nama Lembaga	Waktu	Tingkat (lokal,nasional, internasional)
1	Djemari Mardapi	Universitas Islam Riau	1 hari	Nasional
2	Djemari Mardapi	Universitas Airlangga	1 hari	Nasional
3	Djemari Mardapi	Faculty of Education, Burapha University	1 hari	Internasional

XII. PENCAPAIAN PRESTASI/REPUTASI DOSEN

No	Nama Dosen	Prestasi yang dicapai	Waktu Pencapaian	Tingkat (lokal,nasional, internasional)
1	Djemari Mardapi	Piagam Tanda Kehormatan:Satyalancana Karya Satya 30 Tahun dari Presiden RI DR. H. Susilo Bambang Yudoyono	2010	Nasional
2	Djemari Mardapi	Sumadi Suryabrata Lifetime Achievement Award	2015	Nasional

Semua data yang saya isikan dan tercantum dalam biodata ini adalah benar dan dapat dipertanggungjawabkan secara umum. Apabila dikemudian hari ternyata dijumpai ketidaksesuaian dengan kenyataan, saya sanggup menerima resikonya.

Yogyakarta, 6 April 2017



Prof. Djemari Mardapi, Ph.D.
NIP. 19470101 197412 1 001

3. Curriculum Vitae Anggota Tim Peneliti

I. IDENTITAS DIRI

1.1	Nama Lengkap	Harun Rasyid, Drs., M.Pd., Dr.		
1.2	Jabatan Fungsional	Lektor Kepala		
1.3	NIP/NIK	195607271985031024		
1.4	NIDN	0027075610		
1.5	Tempat dan Tanggal Lahir	Ciamis, 20 Juli 1956		
1.6	Alamat Rumah	Perum Griya Mulya Asri H3 Mayungan, Potorono, Banguntapan Bantul		
1.7	Nomor Telp/Fax	---		
1.8	Nomor HP	08122697092		
1.9	Alamat Kantor	Kampus FIP UNY Karangmalang Yogyakarta		
1.10	Nomor Telp/Fax	--		
1.11	Alamat e-mail	harun_pgpauduny@yahoo.com		
1.12	Lulusan yang telah dihasilkan	S1= orang	S2= orang	S3= orang
1.13	Matakuliah yang diampu	4. Metode Penelitian Pendidikan 5. Evaluasi Pembelajaran 6. Evaluasi Pendidikan 7. Asesmen Perkembangan anak usia dini		

II. RIWAYAT PENDIDIKAN

2.1	Program	S1	S2	S3
2.2	Nama PT	IKIP Yogyakarta	IKIP Yogyakarta	UNY
2.3	Bidang Ilmu	Adm. Pendidikan	PEP	PEP
2.4	Tahun Masuk	1978	1995	2004
2.5	Tahun Lulus	1983	1998	2010
2.6	Judul Skripsi/Tesis/ Disertasi	Proyeksi kebutuhan alat pelajaran ketrampilan Pelita IV berdasarkan kurikulum SPG 1976 dan kebutuhan guru SD/TK di Kab. Ciamis	Deyterminan pemerataan kesempatan memperoleh pendidikan SLTP bagi anak usia 13-15 tahun di Daerah Perbatasan Kalimantan Barat-Sarawak Malaysia	Pengembangan Instrumen Asesmen Konsentrasi dan Perkembangan Anak Usia Dini di TK Berbasis Model Bermain
2.7	Nama Pembimbing/Promotor	Dr. Noeng Muhadjir	Prof. Dr. Sunarto Sumarno, Ph.D.	Prof. Suharsimi AK Prof. Endang Ekowarni Sumarno, Ph.D

III. PENGALAMAN MENGAJAR

Tahun	Mata Kuliah	Jenjang	Semester	Prodi/Fakultas
2011	1. Evaluasi Pendidikan 2. Evaluasi Pembelajaran 3. Asesmen Perkembangan AUD 4. Penilaian Pendidikan	S1	Sem. Genap Sem. Ganjil	PGPAUD FIP
2012	1. Evaluasi Pendidikan 2. Evaluasi Pembelajaran 3. Asesmen Perkembangan AUD 4. Penilaian Pendidikan	S1	Sem. Genap Sem. Ganjil	PGPAUD FIP
2013	1. Evaluasi Pendidikan 2. Evaluasi Pembelajaran 3. Asesmen Perkembangan AUD 4. Penilaian Pendidikan	S1	Sem. Genap Sem. Ganjil	PGPAUD FIP
2014	1. Evaluasi Pendidikan 2. Evaluasi Pembelajaran 3. Asesmen Perkembangan AUD 4. Penilaian Pendidikan	S1	Sem. Genap Sem. Ganjil	PGPAUD FIP
2015	1. Evaluasi Pendidikan 2. Evaluasi Pembelajaran 3. Asesmen Perkembangan AUD 4. Penilaian Pendidikan	S1 S2 S1	Sem. Genap Sem. Ganjil	PGPAUD FIP PPs UNY

IV. PENGALAMAN PENELITIAN

No	Tahun	Judul Penelitian	Pendanaan	
			Sumber	Jumlah (Rp)
1	2015	Pengembangan Model Pendidikan Karakter untuk Membangun Harmoni Sosial Potensi Anak Usia Dini di Daerah Intimewa Yogyakarta	DP4M	120jt
2	2014	--	--	--
3	2013	--	--	--
4	2012	Model pendidikan kecakapan hidup bagi remaja putus sekolah dalam usaha hidup	DP4M	73.2jt

		mandiri melalui pelatihan kewirausahaan kerajinan cetak sablon		
5	2011	Model pendidikan kecakapan hidup bagi remaja putus sekolah dalam usaha hidup mandiri melalui pelatihan kewirausahaan kerajinan cetak sablon	DP4M	73.2jt

V. PENGALAMAN PENGABDIAN KEPADA MASYARAKAT

No	Tahun	Judul Penelitian	Pendanaan	
			Sumber	Jumlah (Rp)
1	2015	--	--	--
2	2014	Pemakalah: Seminar Nasional Evaluasi Pendidikan (SNEP) II yang diselenggarakan Program Studi S2 PEP Pascasarjana UNES Semarang tanggal 22 November 2014--	--	--
3	2013	Pendidikan dan Pelatihan Dasar bagi Pendidik PAUD di Kabupaten Gunungkidul	Diknas GK	3.5jt
4	2012	Penyaji Materi dalam Diklat PLPG guru TK/SD se DIY	Dijen Dikti	-
5	2012	Pembicara pada lokakarya analisis kebutuhan model mobile learning untuk PAUD	BPM Semarang	-
6	2011	Mempersiapkan anak secara sosial-emosional menghadapi lingkungan baru, di KB Aisyiyah Taman Fitria Balangan Sleman	--	--

VI. PENGALAMAN PENULISAN ARTIKEL ILMIAH DALAM JURNAL

No	Tahun	Judul Artikel Ilmiah	Volume/ Nomor	Nama Jurnal
1	2013	Model Pendidikan Kecakapan Hidup Bagi Remaja Miskin Putus Sekolah dalam usaha Hidup Mandiri Melalui Pelatihan Kewirausahaan Kerajinan Cetak Sablon	22/ 3 November 2013	LPPM Univet Bangun Nusantara
2	2008	Pengembangan Strategi Pembelajaran untuk meningkatkan kemampuan memahami teks pada siswa SD dengan menggunakan <i>K-W-L Teaching Model</i>	1/XI/ Juni 2008 (Terakredit asi)	HEPI-PPS UNY
3	2008	Potret guru taman kanak-kanak profisonal	Vol. 6 No.1 September 2008	Jurnal cakrawala kependidika n UnivTanjung pura Pontianak
4	2008	Bermain untuk membangun Center of Mind anak usia dini	Volume 1, nomor2, Agustus 2008	Jurnal Pengemban gan ilmu ke- TK-an.
5	2008	Bermain dan perkembangan sosial emosional anak usia dini	Vol.6 No.1 Meret 2008	Jurnal Cakrawala Kependidika n Univ. Tanjungpura Pontianak
6	2004	Guru Agama Islam profesional di era global	04/2004(te rakreditasi)	Khazanah: Majalah Ilmiah Keagamaan dan Kemasyarak atan
7	2002	Pengembangan Strategi Pembelajaran Kolaboratif dalam Tim Mahasiswa Kalimantan Barat	2/XXXII November 2002	Jurnal Kependidika n UNY

VII. SEMINAR

No	Tahun	Judul Kegiatan	Penyelenggara	Panitia/Peserta/ Pembicara
1	2014	Pemakalah: Seminar Nasional Evaluasi Pendidikan (SNEP) II yang diselenggarakan Program Studi S2 PEP Pascasarjana UNES Semarang tanggal 22 November 2014	PPs UNES	Pembicara
2	2014	Peserta Seminar nasional: pendidikan untuk perubahan masyarakat bermartabat di FIP UN Y	FIP UNY	Peserta
3	2012	Seminar Internasional tentang pengembangan bahasa untuk anak usia dini di Universitas Tanjungpura Pontianak (desember 2012)	UNTAN Pontianak	Pembicara
4	2012	Pelatihan dan pembinaan Guru TK di Dinas Pendidikan Pemuda dan Olah Raga Kab. Selemen, tanggal 30 Oktober 2012	Dinas Pendidikan Pemuda dan Olah Raga Sleman	Pembicara
5	2012	Analisis Kebutuhan Model Mobile Learning PAUD 2012 di Hotel Solo Inn.	BPM Semarang	Pembicara
6	2012	Seminar : tanggungjawab Ilmu Pendidikan Terhadap Generasi Bangsa. Di FIP UNY	FIP UNY	Peserta

VIII. PELATIHAN PROFESIONAL

No	Tahun	Jenis Pelatihan (Dalam/Luar Negeri)	Penyelenggara	Jangka Waktu
1	2011	Pelatihan penyusunan soal untuk calon peserta PPG UNY	UNY	1 hari
2	2010	Pelatihan TOT Karya Tulis Ilmiah On Line (TOT-KTI) TK PLB, Batam, 2 hari, P4TK TK PLB	P4TK TK PLB Ditjen PMPTK Diknas	2 hari
3	2008	Pelatihan Penyusunan Instrument Tes Modern	PPs UNY	1 hari
4	2008	Pelatihan penyusunan proposal Hibah Kompetensi	LP- UNY	1 hari

IX. PENGALAMAN PENULISAN BUKU

No	Tahun	Judul Buku	Jumlah Halaman	Penerbit
1	2014	Pengembangan Instrumen Asesmen Anak Usia dini	140	CV. Insan Madani Yk.
2	2014	Pengantar Pendidikan Anak Usia Dini	143	CV. Insan Madani Yk.
3	2012	Asesmen Perkembangan Anak Usia Dini	312	Gama Media Yogyakarta

X. PENGALAMAN PEROLEHAN HKI

No	Tahun	Judul/Tema HKI	Jenis	No Pendaftaran/ Setifikat
-	-	----	---	---
-	--	----	---	---
-	--	----	---	---

XI. PENGALAMAN RUMUSAN KEBIJAKAN PUBLIK/REKAYASA SOSIAL LAINNYA

No	Tahun	Judul/Tema/Jenis Rekayasa Sosial Lainnya yang Telah Diterapkan	Tempat Penerapan	Respon Masyarakat
-	--	---	--	--
-	--	---	--	--
-	--	---	--	--

Semua data yang saya isikan dan tercantum dalam biodata ini adalah benar dan dapat dipertanggungjawabkan secara umum. Apabila dikemudian hari ternyata dijumpai ketidaksesuaian dengan kenyataan, saya sanggup menerima resikoanya.

Demikian biodata ini saya buat dengan sebenarnya.

Yogyakarta, 6 April 2017
Dosen,

Dr. Harun Rasyid, M.Pd.
NIP. 195607271985031024