

## **MODEL PENGEMBANGAN BANK SOAL BERBASIS GURU DAN MUTU PENDIDIKAN**

**Pujiati Suyata, Djemari Mardapi, Badrun Kartowagiran, dan Heri Retnawati**

Program Pascasarjana Universitas Negeri Yogyakarta

email: pujiati\_suyata@uny.ac.id

### **Abstrak**

Penelitian tahun kedua ini bertujuan untuk meneliti karakteristik perangkat tes buatan guru, melakukan analisis penyetaraan karakteristik butir antarwilayah sehingga karakteristik butir antarkabupaten berada pada skala yang sama, dan melakukan analisis penyetaraan karakteristik butir soal antartahun. Penelitian ini menggunakan pendekatan *Research and Development*. Data diperoleh dari dokumen, yang dianalisis dengan metode analisis kuantitatif dan kualitatif. Responden adalah guru penyusun soal ujian semester bersama mata pelajaran IPA, Matematika, Bahasa Inggris, dan Bahasa Indonesia yang tergabung dalam MGMP dari dua kabupaten, yaitu Sleman dan Gunungkidul. Hasil penelitian menunjukkan bahwa ditemukan sejumlah butir soal buatan guru yang baik, berdasarkan teori pengukuran klasik maupun teori respons butir, yang dapat dimasukkan ke dalam tabungan bank soal, diketahuinya perbandingan kualitas karakteristik butir soal buatan guru antartahun antara Kabupaten Sleman dan Kabupaten Gunungkidul, terkait dengan organisasi guru dalam MGMP dan MKKS, serta keterlaksanaan ujian akhir semester bersama yang dilakukan pada setiap kabupaten, dapat disusun model pengembangan bank soal melalui penyusunan soal oleh MGMP atau MKKS, serta keterlaksanaan semua ini meningkatkan kualitas pengujian dan pendidikan di Provinsi DIY.

**Kata kunci:** model pengembangan bank soal, kualitas butir soal, karakteristik instrumen test

## **MODEL OF DEVELOPING TEST ITEM BANK BASED ON TEACHER AND QUALITY OF EDUCATION**

### **Abstract**

This second year study aimed at (1) finding the test instrument characteristics made by teacher for test item bank development, (2) doing an equality analysis by calibrating inter-regional test item characteristics and among districts, (3) doing inter year calibration analyses. A research and development (R&D) approach was used. Data from documents were used for this research. Quantitative and qualitative data analyses were carried out. Respondents for this study were teachers who are members of subject matters of mathematics, science, Indonesia, and English groups constructed test items for local semester exams in Sleman and Gunung Kidul districts. The study found: (1) a number of test items made by teachers were good items of the classic and item response theory criteria and are added up to the test item bank. (2) teacher made test items were comparable inter-years in Sleman and Gunung Kidul districts. (3) an item test bank development might be built using the available grouping of MGMP, KKS and their supporting systems of the districts, and (4) it was implied that based on findings (1), (2), and (3) an improvement of the educational quality can be achieved.

**Keywords:** item test bank development model, quality of item test, test instrument characteristic

## **PENDAHULUAN**

Dalam pendidikan, evaluasi memegang peran penting. Melalui evaluasi, kualitas pendidikan dan perkembangannya dapat dievaluasi dari waktu ke waktu (Jahja Umar, et al, 1998:1, Djemari Mardapi, 2001). Penilaian dapat dilakukan di antaranya melalui kegiatan pengukuran. Dalam pengukuran diperlukan alat, dan alat yang digunakan untuk melakukan pengukuran, salah satu di antaranya adalah tes (Mehrens & Lehmann, 1973). Sebagai contoh, tes prestasi belajar siswa digunakan untuk mengetahui pencapaian hasil belajar siswa. Hal ini menunjukkan bahwa kualitas tes yang digunakan akan memengaruhi hasil pengukuran dan kualitas hasil pengukuran mempengaruhi hasil evaluasi dalam pendidikan.

Pada era otonomi daerah, pemerintah daerah bertanggung jawab atas pelaksanaan ujian akhir semester beserta kualitasnya. Oleh karena tidak setiap sekolah mampu menyiapkan sendiri ujian akhir ini, pelaksanaannya dilakukan di Provinsi DIY. Soal tes dikembangkan oleh Musyawarah Kerja Kepala Sekolah (MKKS) ataupun Musyawarah Guru Mata Pelajaran (MGMP), yang mempunyai wilayah kerja pada setiap kabupaten/kota.

Ujian akhir semester tingkat menengah pertama dilaksanakan dengan tujuan yang sama, yakni mengukur kompetensi siswa pada akhir pembelajaranyang dikembangkan dari standar kompetensi pokok yang sama. Terkait dengan hal tersebut, tes-tes yang dibuat diharapkan mengukur konstruk yang sama dan memenuhi syarat tes yang baik, terutama ditinjau dari karakteristik butir-butir penyusunannya. Mengingat perangkat ujian akhir semester di Provinsi DIY dibuat oleh MKKS yang berbeda, butir-butir penyusunannya mempunyai tingkat kesulitan yang berbeda pula. Hal tersebut berakibat pada pencapaian sekolah berbeda satu dengan yang lain, dan pemerintah mengalami kesulitan untuk membandingkan mutu antarsekolah (Heri Retnawati, Kana Hidayati, 2007).

Berdasarkan hasil prasarvei yang telah dilakukan, karakteristik perangkat tes yang digunakan dalam ujian akhir semester tingkat menengah pertama belum dianalisis untuk mengetahui karakteristiknya dan belum diketahui apakah tes-tes yang digunakan tersebut telah tersusun atas butir-butir yang baik. Selain itu, bank soal untuk keseluruhan mata pelajaran pada setiap daerah belum ada. Padahal, bank soal sangat diperlukan sebagai rujukan dalam mengembangkan perangkat tes dan dapat dipergunakan untuk membandingkan karakteristik butir yang akan digunakan.

Terkait dengan tugasnya dalam melakukan evaluasi pendidikan, guru mempunyai peran yang sangat besar. Guru merupakan evaluator terdepan dalam memantau keberhasilan pendidikan. Selain itu, guru juga mempunyai potensi untuk dapat meningkatkan kualitas pendidikan. Salah satu di antaranya dengan memperbaiki sistem evaluasi yang selama ini telah digunakan. Potensi guru dalam evaluasi tersebut perlu diberdayakan sehingga dapat menjadi ujung tombak untuk peningkatan kualitas tes yang merupakan parameter keberhasilan pemahaman konsep dan nilai-nilai dalam pelaksanaan pendidikan. Adanya beberapa kesenjangan di atas perlu dijumpati. Pengembangan bank soal berbasis guru di Provinsi DIY menjadi hal urgen untuk dilakukan.

Bank soal yang biasa dikenal guru didefinisikan sebagai kumpulan butir tes. Padahal, bank soal tidak hanya mengacu pada sekumpulan butir. Bank soal mengacu pada proses pengumpulan soal, pemantauan, dan penyimpanan dengan informasi yang terkait sehingga mempermudah pengambilan jika akan merakit soal-soal (Thorndike, 1982). Millman (Jahja Umar, 1999) mendefinisikan bank soal sebagai kumpulan yang relatif besar, yang mempermudah dalam memperoleh pertanyaan-pertanyaan penyusun tes. Mudah memiliki pengertian bahwa pada soal-soal

tersebut diberikan indeks, terstruktur, dan diberikan keterangan sehingga mudah dalam memilih untuk disusun sebagai perangkat tes pada suatu ujian.

Senada dengan pengertian-pengertian di atas, Choppin (dalam J. Umar, 1999) memberikan definisi bahwa bank soal merupakan sekumpulan butir tes yang diorganisasikan dan dikatalogkan untuk mencapai jumlah tertentu berdasarkan isi dan karakteristik butir. Karakteristik butir tersebut meliputi tingkat kesulitan, reliabilitas, validitas, dan ketentuan lain.

Ide pengembangan bank soal terkait dengan kebutuhan merakit tes lebih mudah, cepat, dan efisien. Selain itu juga adanya tuntutan kualitas butir soal yang baik pada penyusunan tes. Dengan adanya bank soal, kualitas butir soal pada penyusunan tes dapat dijamin kualitasnya. Van der Linden (Jahja Umar, 1999) menyatakan bahwa pengembangan bank soal merupakan praktik baru dalam pengembangan tes, sebagai hasil dari pengenalan teori respons butir dan kegunaan ekstensif dari pengetahuan komputer pada masyarakat modern.

Pada suatu bank soal yang dikembangkan dengan teori respons butir, program tes dapat dibuat lebih fleksibel dan sesuai. Hal ini disebabkan karena karakteristik butir perangkat tes pada teori respons butir tidak tergantung pada karakteristik peserta tes pada saat kalibrasi. Selain itu, kemampuan siswa peserta tes dapat diketahui dan dapat dibandingkan, karena parameter kemampuan dapat diestimasi pada skala yang sama (Jahja Umar, 1999). Terkait dengan perkembangan ilmu dan teknologi, pengembangan bank soal berdasarkan teori respons butir dapat diset untuk dikembangkan menjadi *computerized adaptive testing* (Hambleton, Swaminathan, dan Rogers, 1991).

Keuntungan-keuntungan yang dapat diperoleh dengan adanya pengembangan bank soal sangat besar. Beberapa di antaranya adalah kebijakan desentralisasi

pada program tes nasional dapat dikenalkan tanpa mengorbankan hasil tes dan dapat; biaya dan waktu yang diperlukan pada kegiatan konstruksi tes dapat direduksi; makin besar jumlah butir soal yang terdapat pada bank soal, permasalahan keamanan menjadi lebih terjamin; kualitas program tes dapat ditingkatkan, dengan adanya butir-butir dalam bank soal yang telah diketahui karakteristiknya; pendidik dapat mendesain perangkat tes yang akan digunakan, dengan memanfaatkan butir-butir yang baik dalam bank soal; guru dapat mengonsentrasikan diri pada usaha untuk meningkatkan kualitas pembelajaran, tanpa harus membelanjakan banyak waktu untuk penyusunan perangkat tes (Jahja Umar, 1999).

Choppin (dalam Jahja Umar, 1999) berpendapat bahwa keuntungan dalam pengembangan bank soal dapat dikelompokkan menjadi empat kategori. *Pertama*, kategori ekonomi. Dengan adanya sistem bank soal, memungkinkan adanya penggunaan butir-butir soal yang baik secara berulang. *Kedua*, dengan adanya bank soal, panjang tes dapat disesuaikan dengan kebutuhan, yang merupakan kategori fleksibilitas. *Ketiga*, kategori konsistensi. Dengan adanya bank soal, dapat dikembangkan tes yang paralel dan hasil tes dapat diperbandingkan karena kemampuan peserta tes dapat diketahui dengan skala yang sama. *Keempat*, keamanan. Dengan adanya bank soal, pengembang tes dapat menyusun beberapa tes alternatif untuk menjaga kebocoran soal pada tes yang bertujuan sangat penting. Selain itu, terdapat beberapa kegiatan penting dalam pengembangan bank soal, yakni penulisan butir soal, validasi dan kalibrasi butir soal, penyimpanan dan pengamanan soal, pengaitannya dengan butir-butir baru dalam bank soal (*equiting*), dan mempertahankan bank soal (Jahja Umar, 1999).

Proses penulisan butir soal merupakan hal penting dalam pengembangan bank soal. Penulisan butir soal ini bukan merupakan

suatu hal yang mudah. Pada penulisan butir soal, diperlukan rekrutmen dan pelatihan (*training*) bagi penulis, yang memerlukan biaya besar.

Pada pengembangan bank soal, penulis butir soal terlebih dahulu memperhatikan tujuan tes yang akan dikembangkan dengan menggunakan butir-butir bank soal. Apakah tes yang akan dikembangkan tersebut untuk seleksi, tes penalaran, ataukah tes prestasi belajar. Tujuan pengembangan tes perlu diperhatikan, mengingat sifat-sifat tes tersebut berbeda-beda.

Hal lain yang perlu diperhatikan pada penulisan butir soal untuk pengembangan bank soal adalah lingkup materi pelajaran. Dengan memperhatikan ruang lingkup atau cakupan materi sebagai bahan tes, diharapkan butir soal tidak terlalu mudah atau tidak terlalu sulit. Butir soal seperti ini yang dapat membedakan peserta tes berdasarkan kemampuan. Terkait dengan hal ini, pembuatan kisi-kisi menjadi strategi yang tepat untuk mempermudah penulisan butir soal.

Langkah selanjutnya adalah validasi dan kaliberasi. Pada tahap ini, terlebih dahulu butir-butir soal yang ada disusun menjadi perangkat tes kemudian diujicobakan. Ujicoba disesuaikan dengan peserta tes yang akan merespons perangkat tes. Pada pengembangan bank soal berdasarkan teori tes klasik, peserta ujicoba harus berasal dari berbagai strata siswa secara proporsional. Hal ini disebabkan pada teori tes klasik, karakteristik peserta ujicoba memengaruhi karakteristik butir soal yang diujicobakan. Jika menggunakan pendekatan teori respons butir, yang perlu diperhatikan adalah jumlah peserta ujicoba, mengingat model parameter berbeda akan memerlukan ukuran peserta ujicoba yang berbeda pula agar karakteristik butirnya stabil (Hambleton dan Swaminathan, 1985).

Validasi merupakan proses menentukan validitas perangkat tes. Validitas ini dapat

diketahui dari isi, konstruk, maupun dikorelasikan dengan kriteria lain. Adapun kaliberasi merupakan proses yang digunakan untuk menentukan karakteristik butir soal. Pengembangan bank soal berdasarkan teori tes klasik diestimasi tingkat kesulitan, daya beda, dan reliabilitas, sedangkan pengembangan soal berdasarkan teori respons butir yang diestimasi adalah parameter butirnya. Pada model satu parameter, butir soal yang diestimasi adalah tingkat kesulitan, nilai fungsi informasi, dan kesalahan pengukuran. Agar lebih mudah dilakukan, kaliberasi ini dapat dilakukan dengan bantuan komputer, dengan program Itean, Ascal, Rascal, Bigstep, Bilog, Multilog, atau yang lain. Dari hasil kaliberasi, dapat ditentukan butir-butir soal yang baik. Butir soal yang baik ini merupakan bank soal yang terjadi.

Langkah selanjutnya adalah mengaitkan butir-butir soal yang ada dengan butir soal yang baru (*linking new items*). Langkah ini bertujuan agar butir-butir baru yang ditambahkan dalam bank soal terkait dengan butir-butir yang lama berdasarkan kaliberasi yang telah dilakukan. Proses tersebut dinamai penyetaraan (*equiting*), yang bertujuan untuk memastikan kualitas butir soal dan mengestimasi konstanta hubungan dengan perangkat tes yang lama (Dorans, 2004).

## METODE

Penelitian ini merupakan tahap revisi atas penelitian model bank soal berbasis guru yang telah dihasilkan sebelumnya (Pujiati Suyata, 2009), dengan indikator kinerja sebagai berikut. *Pertama*, terevisinya model pengembangan bank soal berbasis guru pada penelitian sebelumnya. *Kedua*, bertambahnya tabungan butir-butir soal buatan guru yang baik. *Ketiga*, dilakukannya penyetaraan horisontal karakteristik butir tes buatan guru antarkabupaten. *Keempat*, terselenggaranya penyetaraan vertikal karakteristik butir tes buatan guru antartahun (tahun 2009 dan tahun 2010).

Dengan menggunakan butir-butir perangkat tes buatan guru berikut respons siswa terhadap tes tersebut, dapat dikembangkan bank soal di wilayah provinsi DIY. Setelah melalui analisis kualitatif oleh ahli dan analisis kuantitatif berdasarkan respons peserta terhadap tes tersebut, dapat diketahui sekumpulan butir tersebut baik atau tidak. Terhadap butir-butir yang baik, dilakukan penyetaraan (*equating*) antarwilayah dan antarsatuan waktu untuk memperoleh skala parameter butir dan parameter peserta agar berada pada ukuran yang sama. Setelah tahap ini, model pengembangan bank soal mulai dari butir-butir ini diorganisasikan, disimpan, dijaga keamanannya, ditemukannya sistem guru untuk mengakses dan menambahkan butir baik dari tahun ke tahun perlu. Kesemuanya itu dimaksudkan untuk meningkatkan kualitas perangkat tes yang digunakan untuk ujian akhir semester.

Penelitian ini merupakan penelitian *research and development* dalam rangka merevisi model pengembangan bank soal yang dilakukan sebelumnya, menambah tabungan butir-butir soal yang baik, serta melakukan penyetaraan karakteristik butir antarwilayah dan antartahun. Penelitian dilakukan di Daerah Istimewa Yogyakarta, yang meliputi dua kabupaten yakni Kabupaten Gunungkidul dan Sleman

Secara garis besar metode pengumpulan data penelitian ini menggunakan metode dokumentasi. Dokumen yang berupa lembar

jawaban ujian semester siswa digunakan untuk mengumpulkan data karakteristik butir soal tes. Data yang terkumpul dianalisis dengan metode analisis kuantitatif dan kualitatif.

#### HASIL PENELITIAN DAN PEMBAHASAN

Secara umum, hasil penelitian ini dapat dirangkum sebagai berikut. *Pertama*, tes ujian akhir semester sudah disusun oleh MGMP/MKKS setiap kabupaten/kota. Namun, tingkat kesulitan soal tidak sama. Padahal, soal dikembangkan dengan tujuan dan standar kompetensi yang sama pula.

Hal tersebut mengakibatkan pencapaian sekolah berbeda-beda sehingga mutu antarsekolah sulit dibandingkan. Oleh karena itu, perlu dilakukan penyetaraan karakteristik butir antarwilayah sehingga estimasi butir berada pada skala yang sama. *Kedua*, karakteristik butir soal antarwilayah yang dapat dimasukkan ke dalam bank soal harus berada pada kategori sedang. Berdasarkan teori tes klasik, butir soal dengan indeks kesulitan antara 0,25-0,75 (berkategori “sedang”) dengan daya beda asal bukan negatif dan teori respons dengan tingkat kesulitan (berkategori “mudah” jika  $-2$ , “sedang” jika berada di antara  $-2$  dan  $2$ , dan “sulit” jika  $+2$ .) dapat disimpan dalam bank soal. Berikut ini disajikan rekap butir mata pelajaran. Berdasarkan Tabel 1 yang tergolong ke dalam butir yang baik adalah butir dalam kategori sedang.

Tabel 1. Karakteristik Butir Soal Mata Pelajaran Matematika

Kabupaten	Indeks Kesulitan		Daya Beda		Respons Butir	
Sleman	Mudah	9	Baik	40	Mudah	5
	Sedang	29	Tidak baik	0	Sedang	35
	Sulit	2			Sulit	0
Gunung Kidul	Mudah	1	Baik	38	Mudah	1
	Sedang	20	Tidak baik	2	Sedang	22
	Sulit	19			Sulit	17

Catatan: Menurut teori tes modern: mudah =  $\leq 2$ ; sedang =  $-2$  dan  $2$ ; sulit =  $\geq 2$

Menurut teori tes klasik: sedang = IK 0,25-0,75; DB asal tidak negatif

Tabel 2. Karakteristik Butir Soal Mata Pelajaran IPA

Kabupaten	Indeks Kesulitan		Daya Beda		Respons Butir	
Sleman	Mudah	11	Baik	40	Mudah	7
	Sedang	26	Tidak baik	0	Sedang	32
	Sulit	3			Sulit	1
Gunung Kidul	Mudah	4	Baik	37	Mudah	4
	Sedang	28	Tidak baik	3	Sedang	32
	Sulit	8			Sulit	4

Tabel 3. Karakteristik Butir Soal Mata Pelajaran B. Indonesia

Kabupaten	Indeks Kesulitan		Daya Beda		Respons Butir	
Sleman	Mudah	26	Baik	45	Mudah	26
	Sedang	18	Tidak baik	5	Sedang	18
	Sulit	6			Sulit	6
Gunung Kidul	Mudah	12	Baik	43	Mudah	16
	Sedang	29	Tidak baik	7	Sedang	23
	Sulit	9			Sulit	11

Tabel 4. Karakteristik Butir Soal Mata Pelajaran B. Inggris

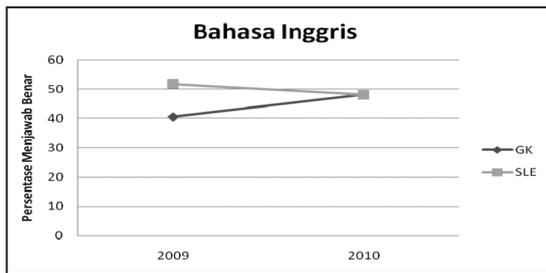
Kabupaten	Indeks Kesulitan		Daya Beda		Respons Butir	
Sleman	Mudah	11	Baik	49	Mudah	5
	Sedang	29	Tidak baik	1	Sedang	42
	Sulit	10			Sulit	3
Gunung Kidul	Mudah	11	Baik	48	Mudah	8
	Sedang	25	Tidak baik	2	Sedang	35
	Sulit	14			Sulit	7

Ketiga, perbandingan tingkat kesulitan antartahun digunakan pendekatan teori tes klasik dan teori respons butir dengan *teknik concurrent calibration* memanfaatkan program BILOGMG. Teknik ini dilakukan untuk mendapatkan hasil estimasi parameter butir berada pada skala yang sama. Hasil selengkapnya disajikan dalam Gambar 1.

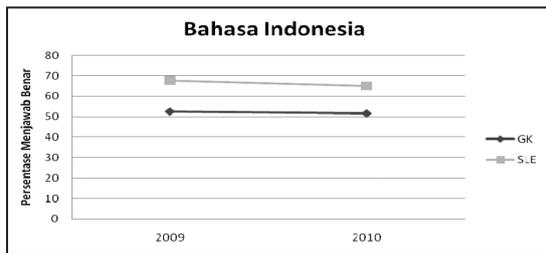
Jika dianalisis dengan teori tes klasik, pada mata pelajaran Bahasa Indonesia persentase menjawab benar di Kabupaten Sleman lebih tinggi dibandingkan dengan di Kabupaten Gunungkidul. Ini berarti bahwa rerata tingkat kesulitan butir perangkat tes di Kabupaten Gunungkidul lebih tinggi dibandingkan dengan di Kabupaten Sleman.

Jika dibandingkan antartahun, rerata tingkat kesulitan tahun 2009 lebih tinggi dibandingkan dengan tahun 2010. Pada mata pelajaran Matematika, rerata tingkat kesulitan pada tahun 2009 hampir sama di Kabupaten Sleman dan Gunungkidul. Pada tahun 2009 dan 2010, butir-butir di Sleman lebih mudah dibandingkan dengan butir soal di Kabupaten Gunungkidul. Hasil selengkapnya disajikan pada Gambar 2, Gambar 3, dan Gambar 4.

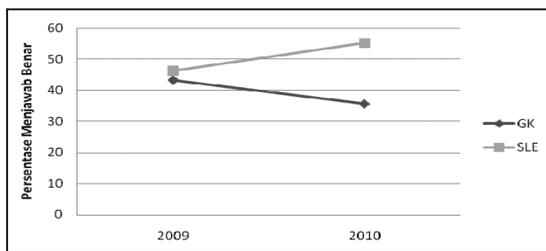
Pada mata pelajaran IPA tahun 2009 dan 2010, butir-butir perangkat IPA di Kabupaten Gunungkidul lebih sulit dibandingkan dengan di Kabupaten Sleman. Hal ini ditunjukkan dengan rerata proporsi menjawab benar di Kabupaten Sleman lebih



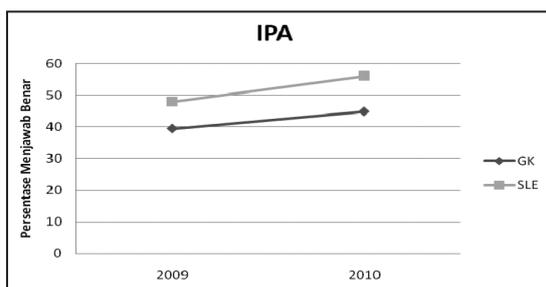
Gambar 1. Perbandingan Tingkat Kesulitan Antartahun Bhs. Inggris



Gambar 2. Perbandingan Tingkat Kesulitan Antartahun B. Indonesia



Gambar 3. Perbandingan Tingkat Kesulitan Antartahun Matematika



Gambar 4. Perbandingan Tingkat Kesulitan Antartahun IPA

tinggi dibandingkan dengan di Kabupaten Gunungkidul. Perbandingan tingkat kesulitan antartahun dengan pendekatan teori respons butir untuk tahun 2009 dan 2010 disajikan

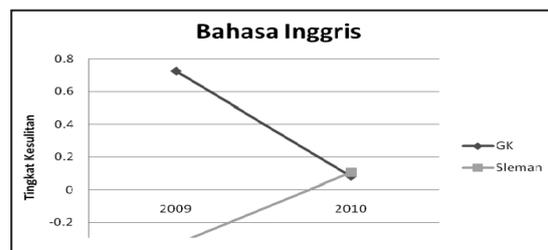
pada Gambar 5. Hasil ini menunjukkan bahwa selama tahun 2009 dan 2010, tingkat kesulitan butir di Kabupaten Gunungkidul untuk mata pelajaran Bahasa Indonesia lebih sulit jika dibandingkan dengan tingkat kesulitan di Kabupaten Sleman.



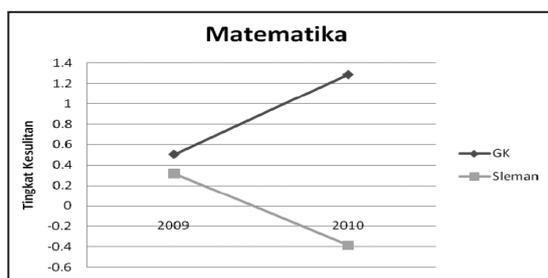
Gambar 5. Perbandingan Tingkat Kesulitan Antartahun Bahasa Indonesia

Untuk mata pelajaran bahasa Inggris di Kabupaten Gunungkidul rerata tingkat kesulitan tahun 2009 lebih tinggi dibandingkan dengan tahun 2010. Namun, di Kabupaten Sleman yang terjadi adalah kebalikannya, rerata tingkat kesulitan tahun 2009 lebih rendah dibandingkan dengan tahun 2010. Di tahun 2010, tingkat kesulitan di kedua kabupaten hampir sama. Hasil analisis ditunjukkan pada Gambar 6

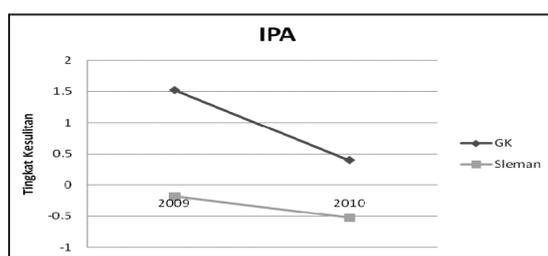
Untuk mata pelajaran Matematika pada tahun 2009, rerata tingkat kesulitan lebih rendah dibandingkan dengan tahun 2010 untuk Kabupaten Gunungkidul. Namun, untuk kabupaten Sleman, terjadi sebaliknya. Rerata tingkat kesulitan tahun 2010 di Kabupaten Gunungkidul jauh lebih tinggi dibandingkan dengan di Kabupaten Sleman. Hasil selengkapnya disajikan pada Gambar 7.



Gambar 6. Perbandingan Tingkat Kesulitan Antartahun Bhs. Inggris



Gambar 7. Perbandingan Tingkat Kesulitan Antartahun Matematika



Gambar 8. Perbandingan Tingkat Kesulitan Antartahun IPA

Di Kabupaten Gunungkidul dan Sleman, tingkat kesulitan untuk mata pelajaran IPA di tahun 2009 lebih tinggi dibandingkan dengan tahun 2010. Pada kedua tahun ini, tingkat kesulitan di Kabupaten Sleman jauh lebih rendah dibandingkan dengan di Kabupaten Gunungkidul

Tabel 5 memperlihatkan bahwa menurut teori tes klasik, butir soal Kabupaten Sleman hampir sama dengan di Kabupaten Gunungkidul. Untuk mata pelajaran matematika dan bahasa Inggris, jumlah butir baik Kabupaten Sleman lebih banyak daripada

Kabupaten Gunungkidul. Sementara itu, jumlah butir soal IPA dan bahasa Indonesia baik terdapat di Kabupaten Gunungkidul. Hal itu menginformasikan bahwa kualitas butir soal untuk kedua kabupaten relatif sama.

## SIMPULAN

Berdasarkan hasil penelitian, dapat disimpulkan sebagai berikut. *Pertama*, terkait dengan organisasi guru dalam MGMP dan MKKS dan keterlaksanaan ujian akhir semester bersama yang dilakukan di tiap kabupaten dan tiap mata pelajaran, dapat disusun model pengembangan bank soal berbasis guru di Provinsi DIY. Pada tahun II penelitian ini (2010), telah dilakukan revisi model awal bank soal berbasis guru di DIY. *Kedua*, model bank soal yang telah direvisi (2010), butir-butir soalnya telah disetarakan baik antarwilayah maupun antartahun.

Dengan demikian, butir soal antarwilayah maupun antartahun telah berada pada skala yang sama. *Ketiga*, dengan model tersebut, telah dihasilkan sejumlah butir soal baik, baik dalam mata pelajaran bahasa Indonesia, bahasa Inggris, Matematika, maupun IPA yang dapat dijadikan tabungan bagi bank soal di wilayah-wilayah. *Keempat*, dengan dihasilkannya butir-butir baik, menurut teori klasik maupun teori respons butir, keberadaan bank soal berbasis guru di wilayah dapat meningkatkan mutu pengujian di wilayah-wilayah tersebut.

Tabel 5. Jumlah Butir Soal Baik Teori Tes Klasik

Kabupaten	Matematika	IPA	Bhs. Indonesia	Bhs. Inggris
Gunungkidul	20	28	29	25
Sleman	29	26	18	29

Tabel 6. Jumlah Butir Soal Baik Teori Respons Butir

Kabupaten	Matematika	IPA	Bhs. Indonesia	Bhs. Inggris
Gunungkidul	22	32	23	35
Sleman	35	32	18	42

**DAFTAR PUSTAKA**

- Dorans, N.J. 2004. "Equating, Concordance and Expectation". *Jurnal Applied Psychological Measurement*, Vol. 28. No. 4, July 2004, p. 219-226.
- Hambleton, R.K., Swaminathan, H., & Rogers, H.J. 1991. *Fundamental of Item Response Theory*. Newbury Park, CA: Sage Publication Inc.
- Hambleton, R.K. & Swaminathan, H. 1985. *Item Response Theory*. Boston, MA: Kluwer Inc.
- Retnawati, Heri & Kana Hidayati. 2007. "Perbandingan Metode *Concordance* Berdasarkan Teori Tes Klasik". *Laporan Penelitian*. Yogyakarta: Lembaga Penelitian UNY.
- Mardapi, Djemari. 2001. "Ebtanas dalam Tinjauan Evaluasi Pendidikan". *Bahan Kuliah Umum Mahasiswa Baru Pascasarjana Universitas Muhammadiyah Prof. Dr. Hamka*, tanggal 8 September 2001.
- Mehrens, W.A. & Lehmann, I.J. 1973. *Measurement and Evaluation in Education and Psychology*. New York: Hold, Rinehart and Wiston, Inc.
- Suyata, Pujiati. 2009. "Identifikasi *Need Assessment* Model Bank Soal Berbasis Guru di DIY". *Jurnal Kependidikan*. 3(1).24-37.
- Thorndike. 1982. *Measurement and Evaluation in Psychology and Education*, 7<sup>th</sup>ed. New Jersey: Pearson Education INC.
- Umar, Jahja. 1999. "Item Banking". Dalam Masters, G.N. dan Keeves, J.P. (Ed). *Advances in Measurement in Educational Research and Assessment*. New York: Pergamon.