

ISSN:2302-3023

fatahillah

Jurnal Pendidikan dan Muamalah

Vol. IX. No. 2. Oktober 2012

**Reinterpretasi Pemahaman Teks Al-Quran Sebagai Sumber Utama Hukum Islam
dalam Perspektif Nashr Hamid Abu Zaid**
Misbah Iskandar

Wakaf Produktif Sebagai Instrumen Investasi Publik
Ahmad Chairul Hadi

**Reformasi Pendidikan Islam: Tinjauan Pembelajaran Pendidikan Islam dengan
Pendekatan Kurikulum Sosial Kultural**
Gunawan Ikhtiono

**Restrukturisasi Pendidikan Agama Islam
Menuju Pendidikan yang Berkualitas**
Ajat Sudrajat

Hubungan Sarana Berfikir Ilmiah Dengan Logika Dan Statistika
Faridl Musyadad

Pengembangan Media Dan Sumber Belajar
Ismanto

**Pengembangan Instrumen Penilaian Tahfidz Al-Qur'an di FITK UNSIQ Wonosobo
Jawa Tengah**
Abdullah Faiz

**Model Instrumen Asesmen Pemahaman Membaca
Skala Dikotomus dan Politomus:
Menjawab Kebutuhan Sistem Penilaian Literasi Siswa**
Beniati Lestyarini

Fatahillah	Vol. 9	No. 2	Hlm. 1-105	Tangerang Selatan Oktober 2012	ISSN 2302-3023
------------	--------	-------	------------	-----------------------------------	-------------------

Fatahillah

Jurnal Pendidikan dan Muamalah
Vol. IX No.2. Oktober 2012

	Halaman
Reinterpretasi Pemahaman Teks Al-Quran Sebagai Sumber Utama Hukum Islam dalam Perspektif Nashr Hamid Abu Zaid Oleh: Misbah Iskandar	3
Wakaf Produktif Sebagai Instrumen Investasi Publik Oleh: A. Chairul Hadi.....	15
Reformasi Pendidikan Islam: Tinjauan Pembelajaran Pendidikan Islam dengan Pendekatan Kurikulum Sosial Kultural Oleh: Gunawan Ikhtiono.....	26
Restrukturisasi Pendidikan Agama Islam Menuju Pendidikan yang Berkualitas Oleh: Ajat Sudrajat.....	35
Hubungan Logika dan Statistika dengan Sarana Berfikir Ilmiah Oleh: Faridl Musyadad.....	47
Pengembangan Media Dan Sumber Belajar Oleh: Ismanto.....	60
Pengembangan Instrumen Penilaian Tahfidz Al-Qur'andi FITK UNSIQ Wonosobo Jawa Tengah Oleh: Abdullah Faiz.....	75
Model Instrumen Asesmen Pemahaman Membaca Skala Dikotomus dan Politomus: Menjawab Kebutuhan Sistem Penilaian Literasi Siswa Oleh: Beniati Lestyarini.....	88

MODEL INSTRUMEN ASESMEN PEMAHAMAN MEMBACA SKALA DIKOTOMUS DAN POLITOMUS:

Menjawab Kebutuhan Sistem Penilaian Literasi Siswa

Oleh: Beniati Lestiyarini*

Abstract

This study attempts to develop an instrument model for assessing the reading comprehension of junior high school students. This study was a research and development (R & D) study employing a procedural model consisting of four main steps, i.e: a) preliminary study, b) prototype product development, c) product validation, and d) product review. The construct of an instrument was developed through FGD. The item specification of the reading comprehension test was developed on the basis of construct instrument and Basic Competences of VII. The developed items matched with the Rasch model (IRT 1 PL) based on the MNSQ INFIT values (0,77-1,30) both in the small and large scale product tryout. The item difficulty level was based on the size of the delta or the threshold in the range of $\pm 2,00$. The test reliability increased from 0,97 to 0,99. The guidelines for using the instrument model was made in order that teacher can independently use the instrument.

A. Pendahuluan

Pendidikan bahasa menjadi salah satu elemen penting dalam upaya membangun karakter bangsa sekaligus meningkatkan kualitas bangsa. Dalam konteks ini, budaya literasi atau baca-tulis menjadi komponen yang harus dikuasai dan dikembangkan oleh seluruh warga negara. Upaya menuju masyarakat madani literasi senantiasa menjadi program dunia yang terus dijalankan. Hal ini tertuang juga dalam program *Education for All* (EFA) atau Pendidikan untuk Semua (PUS) di bawah koordinasi PBB untuk 164 negara di dunia yang ikut serta dalam keanggotaan program. Pada April 2000, Konferensi di Dakar diselenggarakan untuk menyepakati enam tujuan EFA yakni: ... *the expansion of early childhood care and education, the achievement of universal primary education, the development of learning opportunities for youth and adults, the spread of literacy, the achievement of gender parity and gender equality in education, and improvements in education quality* (UNESCO, 2007: 5).

Wujud partisipasi Indonesia untuk mendukung kesuksesan program EFA tertuang secara konseptual dalam Rencana Pembangunan Pendidikan Nasional (RPPN). Saat ini, Rencana strategis (Renstra) Kementerian pendidikan nasional (Kemendiknas) periode 2010-2014 menjadi pedoman bagi semua

* Dosen Fakultas Bahasa dan Sastra Universitas Negeri Yogyakarta

tingkatan pengelola pendidikan di pusat dan daerah dalam merencanakan dan melaksanakan serta mengevaluasi program dan kegiatan pembangunan pendidikan (Kementrian Pendidikan Nasional, 2010: 2). Renstra Kemdiknas 2010-2014 memuat beberapa program Kemdiknas yang didesentralisasikan ke lembaga-lembaga di bawahnya kemudian diaktualisasikan dalam praktik-praktik pendidikan.

Konsep literasi bersinggungan dengan dua aktivitas utama, yakni membaca dan menulis. Perwujudan budaya literasi dalam tulisan ini lebih difokuskan pada kompetensi membaca. Penguasaan kompetensi membaca merupakan *sine qua non* (prasyarat mutlak) untuk meningkatkan kualitas suatu bangsa. Dalam lingkup pembelajaran, kompetensi membaca merupakan subkompetensi berbahasa yang sangat penting untuk menguasai materi berbagai bidang. Tanpa penguasaan kompetensi membaca yang baik, penguasaan terhadap materi khususnya dari bacaan atau teks menjadi kurang optimal.

Kondisi ideal seperti yang diuraikan di atas ironisnya belum menemukan kesesuaiannya dengan kondisi di lapangan. Dari beberapa hasil survei, Indonesia masih menduduki peringkat rendah dalam hal kemampuan memahami bacaan. Survei internasional yang dilakukan oleh PISA tahun 2000 (Greany & Kellaghan, 2008: 126) menempatkan Indonesia pada urutan terakhir dari 40 negara yang disurvei.

Hal lain yang juga terkait dengan rendahnya tingkat kemampuan membaca adalah angka keaksaraan nasional di Indonesia yakni sebesar 98,7% pada tahun 2002 khususnya untuk warga berusia 15-24 tahun (Fasli Jalal dan Nina Sardjunani, 2006: 5). Dengan kata lain, masyarakat Indonesia usia produktif muda belum terbebas dari buta aksara. Hasil survei ini semakin menunjukkan bahwa masih ada banyak hal yang harus dilakukan untuk meningkatkan kualitas pendidikan di Indonesia khususnya tentang membaca. Kemampuan literasi menempati posisi penting dalam upaya meningkatkan kualitas kehidupan bangsa disamping dua kemampuan lain yakni numerasi dan teknologi. Posisi penting tersebut terkait dengan angka keaksaraan nasional yang dijadikan sebagai salah satu parameter dalam menentukan *Human Development Index* (HDI) yang merupakan tolok ukur pencapaian kualitas suatu negara.

Hasil Ujian Nasional (UN) Bahasa Indonesia tahun 2010 juga masih menunjukkan perlunya peningkatan. Skor rerata mata pelajaran Bahasa Indonesia tingkat SMP/MTs sebesar 7,39 yang diklasifikasikan B (data Puspendik, 2010). Skor ini hampir sama dengan mata pelajaran lain yang di-UN-kan. Padahal semestinya pelajaran Bahasa Indonesia yang merupakan bahasa nasional lebih dikuasai oleh siswa. Bahkan untuk tingkat SMA/MAN, skor rerata nasional mata pelajaran Bahasa Indonesia menduduki posisi terendah kedua setelah mata pelajaran Biologi yang menunjukkan skor rerata 7,39 dengan klasifikasi B. Pencapaian skor ini dapat dikatakan jauh lebih

rendah dari skor rerata nasional pelajaran Bahasa Inggris yang menunjukkan hasil sebesar 7,63 yang diklasifikasikan A.

Untuk mengetahui tingkat penguasaan pemahaman membaca tentu saja hal ini tidak dapat dilepaskan dari sistem penilaian dan asesmen pemahaman membaca. Berbagai strategi membaca dikenalkan pada guru-guru dalam kegiatan *workshop*, seminar, pelatihan, maupun kegiatan akademik lainnya. Kegiatan ini sangat positif untuk menambah referensi strategi mengajar guru. Namun, pengetahuan tentang materi dan strategi pembelajaran membaca belum banyak ditindaklanjuti dengan penguatan referensi mengenai bagaimana mengukur pemahaman membaca siswa. Padahal, kegiatan membaca bukan merupakan kegiatan sederhana, namun melibatkan serangkaian kegiatan yang kompleks mulai dari pemaknaan sebuah kata sampai pada penciptaan ide baru sebagai tindak lanjut dari pemahaman bacaan (Caldwell, 2008: 2; Kintsch & Kintsch, 2005: 7). Jika guru kurang memahami bagaimana asesmen pemahaman membaca yang harus dilakukan, aspek-aspek yang semestinya menjadi fokus perhatian dalam upaya pemahaman materi bacaan menjadi kurang terelaborasi dengan baik.

Permasalahan-permasalahan yang telah diuraikan di atas semakin kompleks dengan adanya diskursus mengenai kurikulum khususnya *learning continuum* kompetensi membaca. Penentuan capaian level pemahaman membaca harus disusun sehingga membentuk garis kontinum pembelajaran sesuai dengan jenjang pendidikan (Northwest Evaluation Association, 2003: 4). Upaya ini dilakukan agar kontinum pembelajaran kompetensi membaca tercapai baik secara konten maupun konteks.

Dari hasil diskusi dalam forum *Focus Group Discussion* (FGD) dengan beberapa guru yang tergabung dalam MGMP Bahasa Indonesia SMP Kota Yogyakarta, para guru menginginkan model penilaian membaca yang dapat digunakan untuk mengukur pemahaman membaca siswa. Biasanya para guru hanya menanyakan ide pokok bacaan dan inti bacaan. Aspek-aspek lain yang melibatkan level kognitif tinggi seperti kemampuan berpikir kritis siswa, daya kreatif siswa, dan asertifitas siswa belum banyak dilakukan. Padahal, kemampuan-kemampuan tersebut adalah komponen penting yang dapat digunakan sebagai dasar untuk melakukan asesmen pemahaman membaca siswa.

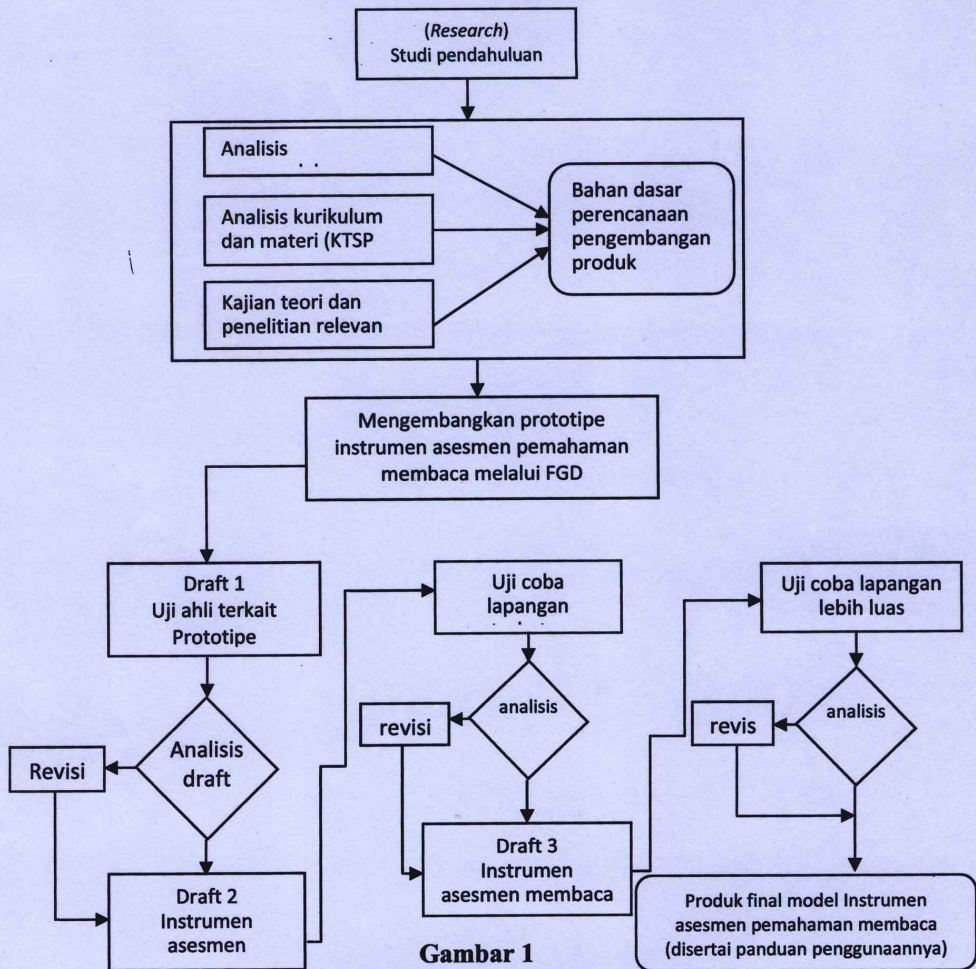
Penelitian ini mencoba untuk mengembangkan model instrumen asesmen pemahaman membaca untuk jenjang SMP yang dapat digunakan guru sebagai salah satu referensi dalam melakukan asesmen pemahaman membaca siswa. Pemahaman yang dimaksudkan adalah penguasaan terhadap bacaan mulai dari level pemahaman paling rendah atau disebut pemahaman faktual sampai pada pemahaman aplikatif, yakni bagaimana siswa mampu mengaplikasikan konsep-konsep yang ada dalam bacaan ke dalam kehidupan nyata. Pemilihan jenjang SMP didasarkan pada asumsi bahwa pada tingkat ini terjadi peralihan cara belajar yang cenderung menuntut level kognitif yang lebih tinggi. Artinya, level pemahaman atau komprehensi yang berupa penguasaan bacaan sudah

34

harus dimiliki oleh siswa meskipun pemahaman membaca sudah mulai ditekankan sejak kelas IV. Tuntutan ini semestinya diimbangi dengan upaya komprehensif termasuk dalam pengembangan asesmen pemahaman siswa karena asesmen yang dilakukan sangat penting untuk melihat proyeksi kemajuan siswa pada setiap level yang dicapai.

Pengembangan model instrumen asesmen pemahaman membaca dalam penelitian ini mengadaptasi beberapa teori asesmen pemahaman membaca yang berlaku secara internasional seperti PISA dan PIRLS (Olivert, 2007: 1; Lawton, 2006: 3; Abernathy, 2007: 2) serta dari berbagai ahli membaca (Bloom, 1956: 18; Anderson & Krathwohl, 2001: 23; Dalton, 2003: 5; Parinas, 2009: 14, Dettmer, 2006: 70) maupun yang dikembangkan oleh beberapa negara. Model instrumen dikembangkan dengan menggabungkan skala dikotomis dan politomis serta dirumuskan dengan mempertimbangkan kebutuhan dan konteks dunia pendidikan di Indonesia supaya dapat lebih mudah dipahami dan diaplikasikan oleh guru. Melalui forum diskusi yang terfokus, berbagai masukan ditampung dan dijadikan bahan pertimbangan untuk menyusun desain penelitian dan pengembangan produk. Pengembangan ini disesuaikan dengan kurikulum dan target pencapaian level pemahaman membaca sehingga dapat digunakan secara efektif oleh guru khususnya guru Bahasa Indonesia sesuai dengan konteks pembelajaran bahasa Indonesia. Jenjang SMP dipilih dengan pertimbangan bahwa siswa pada jenjang ini dituntut untuk lebih aktif terlibat dalam aktivitas kognitif pemahaman membaca yang lebih tinggi daripada jenjang sebelumnya. Tingkat kelas, yakni kelas VII dipilih karena penelitian ini menyesuaikan Standar Kompetensi dan Kompetensi Dasar membaca kelas VII. Pada tingkat ini, kegiatan membaca lebih menitikberatkan pada membaca untuk memahami bacaan. Beberapa permasalahan dalam penelitian ini ialah bagaimanakah prosedur pengembangan instrumen asesmen pemahaman membaca siswa jenjang SMP kelas VII di Kota Yogyakarta? Bagaimanakah prototipe instrumen asesmen pemahaman membaca siswa jenjang SMP kelas VII di Kota Yogyakarta? Bagaimanakah kualitas produk akhir instrumen asesmen pemahaman membaca siswa jenjang SMP kelas VII di Kota Yogyakarta?

Tulisan ini merupakan ringkasan hasil penelitian dan pengembangan (*Research and Development/R & D*) yang bertujuan untuk mengembangkan produk pembelajaran berupa model instrumen asesmen pemahaman membaca. Seperti yang dinyatakan oleh Borg dan Gall (1989: 772) bahwa "*Educational Research and Development (R & D) is a process used to develop and validate educational product*". Model pengembangan yang digunakan dalam penelitian ini adalah model prosedural. Prosedur pengembangan secara visual adalah sebagai berikut.



Gambar 1
Prosedur Pengembangan Penelitian

Produk dalam penelitian dan pengembangan ini meliputi tiga jenis, yakni konstruk instrumen asesmen pemahaman membaca, kisi-kisi soal tes asesmen pemahaman membaca, dan soal tes asesmen pemahaman membaca disertai panduan penggunaannya untuk siswa jenjang SMP kelas VII. Produk yang diujicobakan secara empiris kepada siswa dalam penelitian ini adalah produk ketiga, yakni soal tes pemahaman membaca. Adapun konstruk instrumen dan kisi-kisi soal telah melalui uji ahli yang dilakukan dalam FGD.

Tahap uji coba produk dilakukan sesudah kegiatan perencanaan produk yang dikembangkan yakni prototipe model instrumen asesmen membaca.

1. Desain Uji Coba

Uji coba produk dalam penelitian ini dilakukan melalui dua tahap yaitu, a) Uji coba lapangan terbatas (*preliminary fields test*), dan b) Uji coba lapangan lebih luas (*main field test*).

2. Subjek Coba

Subjek coba dalam penelitian dan pengembangan ini dipilih dengan teknik *cluster proporsional random sampling*. Teknik ini dipilih karena populasi sangat banyak dengan letak geografis yang berbeda (Dattalo, 2008: 5). Berdasarkan teknik pemilihan subjek penelitian tersebut, maka subjek coba yang dilibatkan meliputi:

- a. Siswa pada 1 SMPN di Kota Yogyakarta pada uji coba lapangan terbatas yakni SMPN 12 Yogyakarta (dengan jumlah ± 164 orang)
- b. Siswa pada lima (5) SMPN di Kota Yogyakarta pada uji coba lapangan lebih luas yakni SMPN 8 Yogyakarta, SMPN 2 Yogyakarta, SMPN 4 Yogyakarta, SMPN 14 Yogyakarta, dan SMPN 11 Yogyakarta (tiap sekolah melibatkan ± 165 siswa sehingga jumlahnya ± 711 orang)

Selanjutnya instrumen yang digunakan dalam penelitian ini antara lain sebagai berikut.

- a. Kuesioner, ditujukan untuk para ahli dan guru
- b. Pedoman observasi yang dilengkapi dengan Catatan Anecdotal, Observasi dilakukan selama uji validasi ahli dalam FGD.
- b. Tes pemahaman membaca, untuk melihat keberhasilan dan kelayakan produk yang dikembangkan

Dalam penelitian ini, uji validasi terdiri dari dua macam kegiatan, yakni uji validasi secara logis yang merupakan validasi kerangka konseptual menyangkut konten produk dan uji validasi secara empiris atau prosedural. Validitas logis dicapai melalui *Focus Group Discussion* (FGD) dengan melibatkan ahli kurikulum, ahli pembelajaran bahasa, ahli evaluasi dan guru Bahasa Indonesia SMP sebagai praktisi. Validitas logis menyangkut konten instrumen yang meliputi konstruk instrumen, kisi-kisi soal, dan soal tes pemahaman membaca. Untuk menguji kelayakan konstruk instrumen dilakukan analisis faktor yakni dengan *exploratory factor analysis* melalui program SPSS 16.00.

Validitas empiris produk atau prosedural dilakukan melalui uji coba terbatas dan uji coba lebih luas. Produk yang diujicobakan adalah soal tes pemahaman membaca. Untuk melihat validitas produk ini dilakukan pengujian model pengukuran serta kalibrasi. Model pengukuran yang digunakan adalah *Rasch Model* atau dalam teori pengukuran modern sama dengan *Item Respons Theory* (IRT) satu parameter (1 PL). *Partial Credit Model* atau PCM yang merupakan perkembangan dari *Rasch Model* juga digunakan untuk data dengan skala politomus (Ostini & Nering, 2006: 26). Kalibrasi dilakukan untuk menentukan karakteristik butir soal yang dilihat dari tingkat kesukaran butir soal sesuai dengan *Rasch Model*. Analisis butir soal dilakukan dengan menggunakan program Quest mengingat butir soal yang dikembangkan menggunakan kombinasi skala dikotomus dan politomus.

Sementara itu, reliabilitas instrumen dicapai melalui reliabilitas tes atau yang juga sering disebut indeks sparasi person/testi (Keeves & Master, 1999: 96). Estimasi reliabilitas berdasarkan person/testi sama kedudukannya dengan reliabilitas menurut tes teori klasik (*Classical Test Theory/CTT*) yakni Alpha Cronbach untuk data politomus dan Kuder-Richardson-20 untuk data dikotomus. Semakin tinggi indeks sparasi person semakin konsisten setiap *item* pengukur digunakan untuk mengukur testi yang bersangkutan. Karena butir soal yang dikembangkan dalam penelitian ini menggunakan dua skala yakni skala dikotomus dan politomus, maka baik Alpha Cronbach maupun Kuder-Richardson-20 digunakan untuk menentukan reliabilitas.

Selanjutnya ialah kegiatan teknik analisis data yang dilakukan menggunakan teknik analisis deskriptif kualitatif dan kuantitatif. Data dari observasi dianalisis secara deskriptif kualitatif. Data dari kuesioner dianalisis dengan menggunakan teknik analisis deskriptif kuantitatif. Begitu juga data dari tes pemahaman membaca dianalisis secara deskriptif kuantitatif. Telaah butir soal tes pemahaman membaca dianalisis menggunakan *Rasch Model (RM)* dan pengembangannya yakni *Partial Credit Model (PCM)* sesuai dengan skala yakni skala dikotomus dan politomus. Analisis butir soal menggunakan program Quest kombinasi skala dikotomus dan politomus. Sementara itu, analisis faktor dilakukan untuk menguji kelayakan konstruk instrumen yang dikembangkan.

B. Pembahasan

1. Data Uji Coba

a. Data Prapengembangan

Sebelum produk instrumen asesmen kompetensi pemahaman membaca dikembangkan, diskusi dengan para ahli dan guru dilakukan dalam bentuk *Focus Group Discussion (FGD)* ke-I. Ahli yang dilibatkan dalam diskusi ini dipilih dengan mempertimbangkan kebutuhan penelitian. Ahli yang dilibatkan antara lain ahli kurikulum, ahli pembelajaran bahasa, dan ahli evaluasi, serta guru sebagai praktisi.

Langkah pertama yang dilakukan sebagai data prapengembangan adalah melakukan FGD ke-I dengan agenda utama membahas *learning continuum* keterampilan membaca. Dalam penyusunan kurikulum, dua hal yang menjadi pertimbangan penting adalah kerangka konseptual materi yang disesuaikan dengan perkembangan teori serta kesesuaian kompetensi yang disusun untuk tiap level pendidikan. Kedua hal ini menghasilkan kompetensi pada masing-masing level yang saling berurutan menurut garis kontinum pembelajaran sehingga *learning continuum* yang disusun tercapai baik secara *level by content* maupun *level by context*.

Sebagaimana yang dikemukakan dalam *Idaho State Learning Continuum* (Northwest Evaluation Association/NWEA, 2001: 3) bahwa "The Learning Continuum enhances teacher's ability to provide targeted instruction for individual student or groups of students". Pada setiap level pendidikan, kompetensi yang

32

diharapkan dapat dikuasi oleh siswa harus jelas dan ditempatkan pada posisi yang sesuai. Hal ini akan dapat menghindari adanya kompetensi yang saling tumpang tindih (*overlap competences*) maupun yang terlewatkan (*missing competences*). Semakin tinggi level pendidikan tentu saja semakin tinggi pula kompetensi yang harus dicapai.

b. Data Pengembangan Prototipe Model Instrumen

a. Penyusunan Konstruk Awal Instrumen Asesmen Pemahaman Membaca

Konstruk instrumen asesmen pemahaman membaca mata pelajaran Bahasa Indonesia jenjang SMP kelas VII disusun berdasarkan kajian terhadap *learning continuum* dalam kurikulum, kerangka konseptual pemahaman membaca, berbagai taksonomi membaca oleh para ahli dan standar penilaian kemampuan membaca di berbagai negara dan lembaga internasional, serta pertimbangan relevansi produk dengan konteks pendidikan dan pembelajaran di Indonesia.

b. Penyusunan Kisi-Kisi Soal Awal Tes Pemahaman Membaca

Kisi-kisi soal awal tes kemampuan pemahaman membaca disusun berdasarkan konstruk instrumen yang sudah dikembangkan di atas. Kisi-kisi soal memuat kompetensi pemahaman yang merupakan rincian dari konstruk awal instrumen dan dilengkapi dengan keterangan butir soal yang mencerminkan soal untuk kompetensi tersebut.

c. Penyusunan Soal Awal Tes Pemahaman Membaca

Konstruk dan kisi-kisi instrumen awal asesmen pemahaman membaca dijadikan dasar dalam penyusunan soal awal tes pemahaman membaca yang diujikan ke siswa SMP kelas VII di Yogyakarta. Kompetensi Dasar yang menjadi tuntutan materi dalam kurikulum serta aspek-aspek kompetensi yang dikembangkan berdasarkan tinjauan dari berbagai teori level kognitif pemahaman membaca menjadi pertimbangan penting dalam penyusunan soal. Hal ini juga berlaku sebagai pertimbangan pemilihan bahan bacaan yang dicantumkan dalam soal.

Ada sejumlah tiga puluh (30) soal yang dibuat dengan berdasar pada kisi-kisi yang telah dikembangkan. Jenis soal meliputi soal pilihan ganda (20 butir) dan uraian singkat (10 butir). Jumlah bacaan yang disediakan adalah sebelas (11) teks pendek dengan jenis bacaan sesuai dengan kisi-kisi soal yang telah disusun. Pertimbangan penentuan jumlah soal, jenis soal, dan jumlah teks didasarkan pada masukan dari para guru dan waktu yang disediakan sekolah untuk uji coba soal. Tes sebaiknya dilaksanakan dalam satu kali pertemuan atau 80 menit. Dengan penyediaan waktu tersebut maka diasumsikan waktu untuk persiapan 5 menit, waktu untuk mengerjakan soal pilihan ganda 20 menit (1menit/butir), soal uraian 20 menit (2 menit/butir), membaca teks 33 menit (3 menit/teks), dan pengumpulan jawaban 2 menit sehingga total waktu yang diperlukan 80 menit.

c. Data Uji Coba Terbatas

Dari *output* data dapat diartikan bahwa ada data sebanyak 164 testi yang dianalisis dengan item sebanyak tiga puluh (30). Peluang (*probability level*) sebesar 0,5 sesuai dengan prinsip *Likelihood Maximum*. Tidak ada *case* (testi), item, maupun anchor yang dihapus atau tidak disertakan dalam analisis.

Nilai reliabilitas dari hasil *ouput* pada gambar 2 yakni 0,97. Nilai reliabilitas berdasarkan estimasi item dinamakan reliabilitas sampel. Semakin tinggi nilai reliabilitas sampel semakin meyakinkan bahwa sampel uji coba sesuai dengan item yang diujikan. Sebaliknya, semakin rendah nilai reliabilitas sampel maka semakin banyak sampel uji coba yang tidak memberikan informasi yang diharapkan.

Mean INFIT MNSQ menunjukkan hasil 1,00 dengan SD 0,04, artinya secara keseluruhan item sesuai dengan model Rasch. Nilai INFIT MNSQ berdasarkan model Rasch adalah sebesar 0,77-1,30. Dari data *ouput* dapat diketahui bahwa nilai INFIT MNSQ data uji kecil berada dalam rentang besar INFIT MNSQ menurut teori. Masing-masing item yang diujikan harus memenuhi syarat atau cocok dengan model yang digunakan yakni model Rasch atau IRT 1 PL

d. Data Uji Coba Lebih Luas

Dengan mempertimbangkan beberapa kekurangan pada ujicoba terbatas, maka ujicoba instrumen kompetensi pemahaman membaca skala luas dilakukan. Ujicoba ini melibatkan lima (5) SMP Negeri di Kota Yogyakarta yang sebelumnya sudah ditentukan melalui *cluster proportional random sampling*.

Dari data *output*, ada sebanyak 711 testi dengan L atau panjang item 30 dan peluang 0,50 sesuai dengan prinsip *likelihood maximum*. *Case* atau *testi*, *item*, maupun *anchor* yang dihapus atau tidak disertakan tidak ada. Artinya, semua data layak uji. Nilai reliabilitas adalah 0,99. Pada hasil item estimates uji kecil nilai reliabilitasnya adalah 0,97. Artinya, ada peningkatan nilai reliabilitas dari uji terbatas ke uji lebih luas sebesar 0,02. Hal ini mengindikasikan kualitas instrumen yang dikembangkan lebih baik karena sampel uji coba sesuai dengan item yang diujikan.

Mean INFIT MNSQ menunjukkan hasil 1,00 dengan SD 0,08. Artinya, secara keseluruhan item sesuai dengan model Rasch. Nilai mean INFIT MNSQ menurut teori adalah 0,77-1,30. Reliabilitas berdasarkan case atau testi juga penting untuk melihat reliabilitas instrumen secara keseluruhan. Hasil *output* yang menunjukkan nilai reliabilitas tes hasil uji coba skala luas sesuai dengan model Rasch yakni mean INFIT MNSQ 1,00 dengan SD sebesar 0,20. Item yang paling sulit adalah item 12, sedangkan item yang paling mudah adalah item 3. Setiap tanda x mewakili 3 siswa. Indeks kesukaran masing-masing item dapat dicermati dari hasil keluaran tes klasik yakni pada *Item Analysis Result for Observed Responses*. Karena memakai model IRT 1 PL, data yang dianalisis hanya tingkat kesukaran saja yang ditunjukkan dengan nilai *thresholds*.

2. Analisis Data

1. Analisis Data Prapengembangan

Beberapa temuan penting dalam FGD sebagaimana hasil observasi dapat diuraikan sebagai berikut. *Pertama*, perlu dilakukan kajian untuk menjernihkan permasalahan Standar Kompetensi dan Kompetensi Dasar yang sering membingungkan para praktisi pendidikan. Beberapa KD tampak hanya merupakan jaban yang begitu sederhana dari SK yang justru menjadikan Kompetensi Dasar

tumpang tindih (*overlap*), misalnya saja kelas VII pada KD 7.1 menceritakan kembali cerita anak dan KD 7.2 mengomentari cerita anak (lihat lampiran 2.2).

Kedua, terdapat konsep yang kontradiktif terkait Standar Kompetensi dan Kompetensi Dasar dengan bagaimana seharusnya pembelajaran bahasa dilaksanakan di satuan pendidikan. Materi keterampilan berbahasa tidak dapat dilakukan hanya dalam sekali saja, namun prosesnya semestinya dilaksanakan seperti spiral, diulang secara berjenjang dengan tuntutan tingkat kemampuan yang semakin tinggi.

Ketiga, kurikulum Bahasa Indonesia dipandang atau dikatakan ada di bawah standar yang berlaku internasional. Hal ini dapat dilihat misalnya pada kompetensi penguasaan unsur intrinsik cerpen. Di Indonesia, kompetensi ini mulai dipelajari pada kelas VII sedangkan di luar negeri mulai diajarkan di kelas II. Namun kenyataan di lapangan masih saja menunjukkan bahwa banyak kesulitan yang dihadapi dan tuntutan kurikulum dirasa sulit.

Keempat, ada beberapa kesalahan konsep membaca. Hal ini terkait kemampuan membaca yang dituntut pada Kompetensi Dasar, apakah pada level ingatan, pemahaman, atau mungkin sampai pada aplikasi. Revisi kurikulum harus segera dilaksanakan agar kerancuan ini tidak terus menerus membingungkan para praktisi di lapangan.

Selain itu, secara umum membaca dapat dibagi menjadi dua jenis, yakni membaca nyaring dan membaca dalam hati yang masing-masing memiliki tujuan yang berbeda. Namun ada beberapa Standar Kompetensi yang membingungkan, apakahuntutannya membaca nyaring atau membaca dalam hati karena tidak sinkron dengan Kompetensi Dasar. Misalnya saja, pada SK kelas III, semester 1, memahami teks harusnya sudah masuk pada membaca pemahaman, namun dalam KD masih dengan membaca nyaring. Beberapa KD sampai pada kelas VIII juga memiliki kesalahan konsep yang sama, yakni kurang tegas membedakan kompetensi membaca nyaring atau membaca dalam hati karena membaca nyaring lebih bertujuan untuk kemampuan komunikatif, sedangkan membaca dalam hati lebih menekankan pada pemahaman teks bacaan.

Kelima, ada beberapa kompetensi membaca yang seharusnya perlu diajarkan namun belum masuk pada kurikulum. Padahal, kompetensi tersebut sangat penting untuk dikuasai dan relevan dengan konteks saat ini. Misalnya saja, kemampuan membaca petunjuk, memahami maksud iklan, memahami isi paragraf sebelum dan sesudah, serta membedakan tabel sebagaimana yang tercantum pada lampiran hasil observasi. Kemampuan-kemampuan tersebut seharusnya sudah diberikan pada level SMP. Pada beberapa tes internasional, kemampuan ini sudah menjadi kompetensi wajib yang harus dikuasai siswa. Implikasi dari hal ini yakni pada revisi terhadap kurikulum yang harus mempertimbangkan perkembangan pengetahuan dan paradigma sistem atau model penilaian yang berlaku baik dalam lingkup nasional maupun internasional.

Secara umum, FGD mengenai telaah *learning continuum* menghasilkan kesepakatan bahwa harus ada beberapa Standar Kompetensi dan Kompetensi Dasar yang semestinya direvisi. Penyusunan konstruk instrumen yang dikembangkan

akan didasarkan pada revisi ini. Namun dari pertimbangan bahwa mau tidak mau guru harus patuh pada kurikulum yang berlaku, maka hasil telaah permasalahan dalam kurikulum terkait *learning continuum* yang harus direvisi dijadikan sebagai rekomendasi bagi penentu kebijakan dalam merumuskan kurikulum baru dan juga tentunya menjadi kekayaan pengetahuan bagi para ahli, guru, dan pengembang model instrumen. Instrumen tetap dikembangkan sesuai kurikulum yang berlaku namun dengan lebih memperjelas dan memperkaya ragam teks serta kompetensi pemahaman yang sesuai dengan konteks berdasarkan masukan dari para guru.

2. Analisis Data Pengembangan Prototipe Model Instrumen

a. Analisis Dasar Pengembangan Prototipe Model Instrumen

Prototipe yang dikembangkan dalam penelitian dan pengembangan ini meliputi tiga jenis, yakni konstruk instrumen asesmen kompetensi pemahaman membaca, kisi-kisi soal tes kompetensi pemahaman membaca, serta soal uji kompetensi pemahaman membaca. Pengembangan produk ini bukan suatu hal yang mudah karena disamping harus didasarkan pada kajian konseptual yang kuat mengenai pemahaman membaca, produk yang dihasilkan juga harus dapat digunakan untuk kepentingan praktis di lapangan. Artinya, produk tersebut secara substansi sesuai dengan teori yang relevan, dapat diterima oleh guru serta dapat diaplikasikan dalam pembelajaran membaca khususnya yang berkaitan dengan membaca pemahaman.

Konstruk instrumen asesmen kompetensi pemahaman membaca dikembangkan dengan mengkaji berbagai teori pendukung mengenai pemahaman membaca, teori mengenai level pemahaman membaca, teori mengenai asesmen dan asesmen pemahaman membaca. Disamping itu, kajian terhadap kurikulum juga dilakukan karena instrumen penilaian memuat indikator-indikator kompetensi yang harus dikuasai siswa yang termuat dalam kurikulum. Dengan demikian, produk instrumen ini dilandaskan pada kebutuhan pendidikan yang sesuai dengan tuntutan kurikulum dan kebutuhan sistem penilaian untuk kepentingan praktis.

Konstruk memuat level-level kognitif yang dimulai dari level paling rendah menuju level pengetahuan yang tinggi. Level-level ini disusun dengan mengadopsi taksonomi kognitif dari Bloom yang sudah direvisi oleh Dettmer (2006) yang memuat tiga klasifikasi belajar, yakni *basic learning*, *applied learning*, dan *ideational learning*. *Basic learning* berkaitan dengan kemampuan kognitif dasar yang mencakup *to know* (mengetahui) dan *to understand* (memahami). Pada level ini, kompetensi membaca meliputi kompetensi membaca dasar seperti identifikasi ide atau informasi yang tercantum secara eksplisit serta penarikan hubungan dan kesimpulan antarinformasi. Keterampilan yang terlibat meliputi kemampuan menguasai kata dan frase serta *setting* baik waktu, tempat, suasana. Jika ditinjau dari level pemahaman membaca, keterampilan ini termasuk dalam tingkat pemahaman literal. Sementara itu, kemampuan menarik hubungan dan kesimpulan antarinformasi sudah melibatkan level interpretasi atau dapat juga disebut pemahaman inferensial karena siswa dituntut untuk membuat inferensi dari ide-ide pokok, ide penjelas, urutan informasi, serta hubungan sebab akibat antara hal satu dengan hal lain yang termuat dalam teks bacaan.

Level kognitif *applied learning* berkaitan dengan kemampuan mengaplikasikan (*to apply*), menganalisis (*to analyze*), dan mengevaluasi (*to evaluate*) ide-ide atau informasi dalam teks bacaan. Pada level ini kompetensi pemahaman yang terlibat meliputi kemampuan menginterpretasi dan mengintegrasikan ide pokok dan informasi serta kemampuan memberikan respons kritis terhadap isi, bahasa, pemecahan masalah, serta pribadi penulis. Level pemahaman yang terlibat untuk beberapa subketerampilan sudah merupakan tingkat pemahaman aplikatif karena siswa dituntut untuk menerapkan suatu konsep atau pengetahuan pada dunia nyata seperti pada subkompetensi menemukan aplikasi/relevansi ide teks dalam dunia nyata, menilai relevansi isi teks, menilai kejelasan/kelengkapan teks, merespons secara kritis solusi yang diberikan penulis, merespons secara kritis kepribadian dan motivasi penulis, serta merespons secara kritis perspektif penulis.

Pada level *ideational learning*, kemampuan kognitif yang terlibat meliputi hal-hal yang berkenaan dengan menciptakan ide atau gagasan baru. Kemampuan ini meliputi mensintesis (*to synthesis*), mengimajinasi (*to imagine*), dan menciptakan (*to create*). Pada level ini, pembaca tidak hanya menerima dan mengkonstruksi ide dari teks bacaan namun juga sudah menciptakan ide atau gagasan baru dari hasil membaca. Kompetensi yang terlibat meliputi kemampuan mensintesis keseluruhan isi teks serta merefleksikan dan merencanakan aktualisasi nilai-nilai teks dalam kehidupan. Kompetensi-kompetensi ini jika dilihat dari tingkat pemahaman termasuk dalam pemahaman literal, interpretatif, dan aplikatif.

Untuk mengembangkan soal yang akan digunakan dalam melakukan asesmen kompetensi pemahaman membaca, kisi-kisi soal harus disusun terlebih dahulu. Kisi-kisi soal ini didasarkan pada konstruk instrumen yang sudah dikembangkan. Kisi-kisi soal memuat kompetensi serta butir soal yang menjadi alat untuk mengukur kompetensi tersebut. Kisi-kisi soal harus memuat level-level pemahaman kompetensi membaca yang diukur

Dari kisi-kisi soal, pengembang kemudian menyusun soal yang akan diujicobakan ke sekolah. Penyusunan soal harus mempertimbangkan bentuk soal serta hal-hal yang berkaitan dengan administrasi tes. Untuk bentuk soal, soal campuran antara pilihan ganda dan esai atau uraian singkat dipilih karena kedua bentuk soal dianggap dapat menjadi cermin untuk melihat kompetensi pemahaman siswa. Pengembang juga mempertimbangkan cara menganalisis soal yang merupakan soal campuran. Maka, analisis soal akan menggunakan program Quest yang dapat digunakan untuk soal kombinasi dengan skala dikotomis dan politomis.

b. Analisis Prototipe Asesmen Kompetensi Pemahaman Membaca

Prototipe instrumen asesmen kompetensi pemahaman membaca seperti yang telah diuraikan pada bagian sebelumnya meliputi draf tiga produk utama, yakni konstruk instrumen asesmen kompetensi pemahaman membaca, kisi-kisi soal tes kompetensi pemahaman membaca, serta soal tes kompetensi pemahaman membaca. Produk ini kemudian dilengkapi dengan kunci jawaban yang menjadi pedoman dalam penilaian. Untuk tujuan praktis, produk instrumen yang

dikembangkan disusun dalam sebuah panduan singkat penggunaan instrumen yang dilengkapi dengan petunjuk penggunaannya.

c. Analisis Data Telaah Ahli dan Guru

Hasil telaah ahli dan guru mempertimbangkan aspek substansi dan kemudahan dalam aplikasi di lapangan. Beberapa hasil telaah yang telah diuraikan di atas menjadi pertimbangan penting dalam upaya merevisi produk. Karena untuk mencapai validitas logis produk ini dilakukan melalui forum FGD, maka kesepakatan lebih mudah tercapai. Hal ini dapat dipahami karena antara peneliti atau pengembang, ahli, dan guru bertemu secara langsung sehingga terjadi komunikasi yang lebih efektif. Berdasarkan kesepakatan dari hasil telaah maka revisi terhadap produk dilakukan.

3. Analisis Data Uji Coba Terbatas

Ujicoba terbatas dilakukan untuk melihat kelayakan instrumen yang dikembangkan. Dari hasil ujicoba ini dapat diketahui kualitas instrumen tes kompetensi pemahaman membaca yang dilihat dari kelayakan konstruk instrumen, kecocokan model, serta kualitas butir soal yang dilihat dari tingkat kesulitas butir soal. Pada bagian berikut akan diuraikan hasil analisis ketiga hal tersebut.

a. Uji Kelayakan Konstruk Instrumen

Konstruk instrumen yang dikembangkan sebelumnya telah didiskusikan oleh para ahli, praktisi, dan peneliti melalui FGD. Tahap ini merupakan bentuk pencapaian validitas logis secara kualitatif yang dilakukan melalui *expert judgment*. Produk yang dikembangkan secara kualitatif dapat dianggap valid karena sudah didiskusikan dalam forum diskusi ilmiah melalui kesepakatan bersama. Selain itu, respons guru yang melakukan ujicoba terbatas juga memberikan indikasi bahwa instrumen yang diujicobakan dapat diterima baik oleh guru maupun siswa.

Disamping telaah kualitatif, konstruk instrumen ditelaah juga secara kuantitatif melalui analisis faktor dengan program SPSS 16.00 untuk melihat apakah konstruk layak menurut kecukupan sampel dan diteruskan untuk analisis selanjutnya. Hasil telaah ini dapat mencerminkan penerimaan terhadap konstruk dan sebagai indikator keterbacaan instrumen yang dikembangkan. Kelayakan sampel dilakukan melalui uji *Bartlett's test of sphericity* atau Kaiser-Meyer-Olkin (KMO). Korelasi antara variabel dengan pengujian keseluruhan matriks lewat uji Barlett (*Barlett's test of sphericity*) adalah signifikan apabila harga KMO-nya 0,50 dan dikategorikan cukup (Singgih Santoso, 2010: 73). Besarnya nilai KMO dan Bartlett tes yakni 0,546 dengan taraf signifikansi sebesar 0,000. Secara teori, nilai KMO harus lebih besar dari 0,5 sehingga dari hasil ini dapat diketahui bahwa instrumen yang dikembangkan memenuhi syarat kelayakan model dan dapat dilakukan analisis selanjutnya.

b. Uji Kecocokan Model

Dalam analisis dengan program Quest, suatu item atau *testi/case/person* dikatakan fit atau cocok dengan model dengan mengetahui besarnya nilai *INFIT Mean of Square* (INFIT MNSQ). Besarnya rentang nilai INFIT MNSQ dari 0,77-1,30 (Adam & Khoo, 1996: 30 & 90). Ada pula peneliti yang menggunakan batas yang lebih ketat yakni dengan kisaran 0,83 sampai 1,20 dan ada pula yang menguji

berdasarkan nilai INFIT t yang menggunakan kisaran nilai $t \pm 2,0$ (pembulatan dari $\pm 1,96$) jika taraf kesalahan atau alpha sebesar 5% (Bond & Fox, 2007: 43).

Pedoman yang dipakai untuk mengetahui uji kecocokan model dalam penelitian ini yakni dengan INFIT MNSQ dengan rentang nilai 0,77-1,30. dari hasil *output* program Quest, nilai INFIT MNSQ dari masing-masing item memenuhi syarat kecocokan model, yakni berada pada rentang 0,77 sampai 1,30. Hal ini dapat dilihat pada lampiran hasil *output* program Quest pada ujicoba terbatas baik pada *output item fit, Item Estimates (Category Deltas) In input Order*, serta *Case Estimates In input Order*.

c. Analisis Indeks Kesulitan Butir Soal dengan Rasch Model/IRT 1 Parameter

Hasil *output* program Quest dapat menunjukkan besar indeks kesulitan butir soal dalam CTT (*Classical Test Theory*) maupun IRT (*Item Respons Theory*). Jika analisis menggunakan CTT, indeks kesulitan butir soal tercermin dari besarnya percent pada *output Item Analysis Results for Observed Responses*. Jika analisis menggunakan Rasch Model maka nilai indeks kesulitan butir soal terlihat dari *threshold* atau delta dengan rentang nilai $\pm 2,00$.

Hasil *output* program Quest untuk nilai delta atau *threshold* ujicoba skala kecil dapat dilihat pada lampiran *Item Estimates (Thresholds), Item Estimates (Category Deltas) In input Order, Item Analysis Results for Observed Responses*. Item yang menunjukkan delta atau *thresholds* di luar rentang $\pm 2,00$ yakni item 2 (2,26), item 10 (2,62), dan item 12 (2,80). Selain ketiga item tersebut, semua item memenuhi rentang indeks kesulitan butir soal $\pm 2,00$. Nilai delta atau *thresholds* pada ketiga item yakni item 2, 10, dan 12 menunjukkan bahwa item memiliki tingkat kesulitan yang tinggi.

4. Analisis Data Uji Coba Lebih Luas

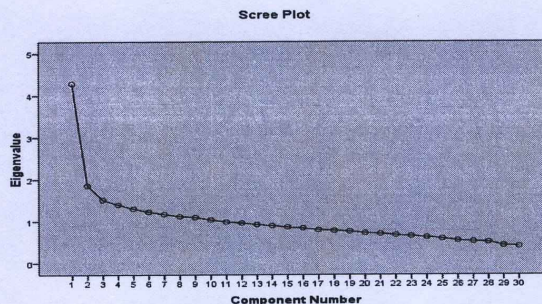
Pada bagian analisis data hasil ujicoba lebih luas diuraikan tiga hal menyangkut validitas konstruk instrumen yang dilihat dari hasil analisis program SPSS, kecocokan model, dan analisis untuk item atau soal yang dikembangkan.

a. Uji Kelayakan Konstruk Instrumen Lebih Luas

Seperti yang telah diuraikan pula pada hasil uji konstruk instrumen terbatas, kualitas konstruk instrumen dapat dilihat dari hasil FGD yang didasarkan pada kesepakatan para ahli bahwa konstruk instrumen sudah layak untuk dikembangkan. Secara empiris, kelayakan konstruk instrumen ini ditunjukkan dari hasil analisis data SPSS untuk melihat kelayakan model dengan uji *Bartlett's test of spherity* atau *Keiser-Meyer-Olkin (KMO)* serta *scree plot* yang ditampilkan sebagai gambaran keunidimensionian tes. Nilai KMO atau Bartlett test dianggap sesuai model jika di atas 0,50. Besarnya nilai KMO dan Bartlett test adalah sebesar 0,791 dengan taraf signifikansi sebesar 0,000 yang artinya layak model. Jika dibandingkan dengan hasil nilai KMO ujicoba terbatas dan uji coba lebih luas, terlihat adanya perbaikan karena nilainya semakin tinggi, yakni 0,546 pada uji coba skala kecil dan 0,791 pada uji coba skala luas.

Hambleton, Swaminathan & Rogers (1991: 9) mengemukakan bahwa asumsi yang melandasi teori respon butir adalah unidimensi, independen lokal dan fungsi karakteristik butir atau kurva karakteristik butir. Unidimensi artinya dimensi

karakteristik peserta yang diukur oleh tes itu tunggal. Tes yang telah diukur diharapkan hanya mengukur satu karakter atau kemampuan saja. Keunidimensian tes tergambar dari gambar *scree plot* yang tampak dari gambar berikut.



Gambar 2
Scree Plot Hasil Uji Coba Lebih Luas

Dari gambar *scree plot* di atas tampak penurunan yang sangat tajam dari *component number* satu ke *component number* dua kemudian diikuti oleh penurunan kecil antara *component number* dua dan seterusnya. Hal ini menunjukkan sifat tes yang dikembangkan telah unidimensi.

b. Uji Kecocokan Model Uji Coba Skala Luas

Dari hasil *output* program Quest, nilai INFIT MNSQ dari masing-masing item memenuhi syarat kecocokan model, yakni berada pada rentang 0,77 sampai 1,30. Hal ini dapat dilihat pada lampiran hasil *output* program Quest pada ujicoba skala kecil baik pada *output item fit*, *Item Estimates (Category Deltas) In input Order*, serta *Case Estimates In input Order*.

c. Analisis Soal dengan Rasch Model Hasil Uji Coba Lebih Luas

Karena model analisis yang digunakan adalah *Rasch Model* atau IRT dengan 1 parameter/1PL, maka analisis karakteristik soal dilakukan dengan hasil *output* program Quest untuk nilai delta atau *threshold* ujicoba lebih luas dapat dilihat berdasarkan *Item Estimates (Thresholds)*, *Item Estimates (Category Deltas) In input Order*, *Item Analysis Results for Observed Responses*.

3. Kajian Produk Pengembangan

Kajian produk pengembangan dapat diuraikan sebagai berikut.

- a. Validitas logis penelitian dicapai melalui FGD sebanyak dua kali dan sudah menghasilkan instrumen asesmen kompetensi pemahaman membaca sesuai dengan rancangan penelitian. Soal yang diujicobakan sudah dinyatakan layak uji coba oleh para ahli dan guru.
- b. Konstruk instrumen yang dikembangkan dilihat kelayakannya dengan menggunakan analisis faktor. Kelayakan konstruk instrumen terlihat dari besarnya indeks KMO dari Tes Bartlett ..
- c. Reliabilitas instrumen dapat dilihat melalui reliabilitas tes dari *output* program Quest. Reliabilitas tes ini pada Teori Tes Klasik setara dengan Kuder

Richardson-20 untuk butir soal dengan dikotomus dan Alpha Cronbach untuk butir soal dengan skala politomus.

- d. Telaah butir soal secara kuantitatif dapat menggunakan program Quest dengan kombinasi skala dikotomus dan politomus.
- e. Model analisis yang digunakan adalah *Rasch Model*/RM/ IRT I PL. Sementara itu, hasil tes uraian menggunakan PCM (*Partial Credit Model*) yang merupakan perluasan dari *Rasch Model*. Pada soal pilihan ganda, *scoring* yang dipakai yakni 0 untuk jawaban salah dan 1 untuk jawaban benar. Pada soal uraian singkat, *scoring* yang dipakai yakni 0 untuk jawaban salah, 1 untuk jawaban mendekati benar, dan 2 untuk jawaban benar. Pada analisis tes dengan PCM, tingkat kesukaran dari suatu tahapan kategori di bawahnya ke kategori di atasnya tidak sama antaritem sehingga besarnya delta untuk suatu tahapan kategori di bawahnya dan delta untuk tahapan kategori di atasnya tidak sama antaritem satu dengan yang lain. PCM yang merupakan perluasan dari model IRT I PL/*Rasch Model* dapat ditulis persamaannya sebagai berikut.

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

C. Penutup

A. Simpulan

Kesimpulan yang dapat ditarik dari hasil penelitian dan pengembangan ini adalah sebagai berikut:

1. Prosedur pengembangan instrumen asesmen pemahaman membaca siswa jenjang SMP kelas VII dilaksanakan dengan model prosedural yang meliputi: a) studi pendahuluan, b) pengembangan prototipe model instrumen asesmen pemahaman membaca, c) validasi produk, serta d) kajian produk akhir.
2. Prototipe produk merupakan draf instrumen asesmen pemahaman membaca yang dikembangkan melalui *Focus Group Discussion* (FGD) dengan melibatkan ahli kurikulum, ahli pembelajaran bahasa, ahli evaluasi, serta guru Bahasa Indonesia SMP.
3. Kualitas produk akhir instrumen tercermin dari kualitas konstruk instrumen asesmen pemahaman membaca, kualitas kisi-kisi soal tes pemahaman membaca, dan kualitas soal tes pemahaman membaca yang dilengkapi dengan panduan penggunaannya. Konstruk instrumen dirumuskan dalam FGD dan dikembangkan berdasarkan teori asesmen pemahaman membaca yang relevan dengan kebutuhan pembelajaran. Kisi-kisi soal dikembangkan berdasarkan konstruk instrumen dan disesuaikan dengan Kompetensi Dasar pada jenjang SMP kelas VII. Soal tes yang dikembangkan cocok dengan model *Rasch* (IRT I PL) berdasarkan nilai INFIT MNSQ (0,77-1,30) baik pada uji coba terbatas maupun lebih luas. Tingkat kesukaran butir soal dilihat dari besarnya delta atau *threshold* pada rentang $\pm 2,00$. Indeks reliabilitas tes meningkat dari 0,97 pada uji coba terbatas menjadi 0,99 pada

uji coba lebih luas. Panduan penggunaan instrumen disusun agar guru dapat menggunakan instrumen secara mandiri.

B. Saran

1. Model instrumen yang dikembangkan dapat diujicobakan dan diimplementasikan dalam skala yang lebih luas tidak hanya regional Yogyakarta.
2. Diseminasi produk perlu dilakukan agar produk yang dikembangkan dapat lebih dikenal oleh praktisi pendidikan khususnya guru Bahasa Indonesia dan juga para pemerhati bahasa untuk pengembangan keilmuan.
3. Sebaiknya ada penelitian lebih lanjut terkait pengembangan bank soal dengan berdasarkan pada konstruk instrumen yang telah dikembangkan. Hal ini dapat dipahami karena soal-soal yang berkaitan dengan kompetensi pemahaman membaca sangat luas dan dapat menggunakan berbagai jenis tes.

DAFTAR PUSTAKA

- Abernathy, S. F. (2007). *No Child Left Behind and the public school*. Ann Arbor: The University of Michigan Press.
- Adams, R.J. & Kho, Seik-Tom. (1996). *Acer quest version 2.1*. Victoria: The Australian Council for Educational Research.
- Anderson, L. & Krathwohl, D.E. (2001). *A taxonomy for learning and assessing: A revision of Bloom's taxonomy of educational objectives* (Abridged). New York: Addison Wesley Longman, Inc.
- Bloom, B.S. (1956). *Taxonomy of educational objectives: Handbook 1 cognitive domain*. London: Longmans, green and co, Ltd.
- Bond, T.G & Fox, Ch.M. (2007). *Applying the rasch model: Fundamental measurement in the human science second edition*. London: Lawrence Erlbaum Associates Publisher.
- Borg, W & Gall, M. D. (1989). *Educational research*. New York & London : Longman.
- Cladwell, J. S. (2008). *Comprehension assessment: A classroom guide*. New York: The Guilford Press.
- Dalton, E. (2003). *The "new Bloom's taxonomy," objectives, and assessments questions*. Diambil pada tanggal 30 Agustus 2011 dari http://gaeacoop.org/dalton/publication/new_bloom.pdf.
- Dattalo, P. (2008). *Determining sample size: Balancing power, precision, and practicality*. New York: Oxford University Press.
- Dettmer, P. (2006). New Blooms in established fields: Four domains of learning and doing. *Roepers Review*, 28, 2 (70-78).
- Fasli Jalal & Nina Sardjunani. (2006). *Keaksaraan bagi kehidupan: Laporan pengawasan global PUS 2006*.

- Greaney, V & Kellaghan, T. (2008). *Assessing national achievement levels in education (Volume 1)*. Washington: The World Bank.
- Hambleton, R.K., Swaminathan, H., & Rogers, H.J. (1991). *Fundamentals of item responses theory*. Newbury Park: Sage Publications.
- Keeves, J.P & Master, G.N. (1999). *Advances in measurement in educational research and assessment*. Amsterdam: Pergamon, An imprint of Elsevier Science.
- Kementerian Pendidikan Nasional. (2010). *Rencana strategis kementerian pendidikan nasional 2010-2014*.
- Kintsch, W. & Kintsch, E. (2005). Comprehension. Dalam S.G. Paris & S.A. Stahl (Eds). *Children's Reading Comprehension and Assessment*. Mahwah, NJ: Erlbaum.
- Lawton, S. B. (2006). NCLB and diversity in school. Dalam F. Brown & R.C. Haunter. *No child left behind and other federal program for urban school district*. USA: Elsevier, Ltd.
- Olivert, D.P. (Ed). (2007). *No child left behind act: Text, interpretation and change*. New York: Nova Science Publisher, Inc.
- Parinas, N. (2009). Revised taxonomy: Reframing our understanding of knowledge and cognitive process. Dalam PEMEA, *The assessment Handbook: Continuing Education Program Vol 1*. Phillipines: PEMEA.
- Ostini, R. & Nering, M. (2006). *Polytomus item response theory models*. USA: SAGE Publication, Ltd.
- Pusat Penilaian Pendidikan & BSNP. (2010). *Laporan hasil ujian nasional tahun ajaran 2009-2010*.
- Singgih Santoso. (2010). *Statistik multivariat: Konsep dan aplikasi dengan SPSS*. Jakarta: PT Elex Media Komputindo.
- Unesco. (2007). *Education for all by 2015: Will we make it? EFA global monitoring report 2008*. UK: Oxford University Press.