

## **PSIKOLOGI KOGNITIF SEBAGAI DASAR PENGEMBANGAN TES KEMAMPUAN DASAR MEMBACA BIDANG SAINS**

*Ani Rusilowati*

Pendidikan Fisika FKIP-UNES

rusilowati@yahoo.com

### **Abstrak**

Penelitian ini bertujuan untuk: (1) menguji psikologi kognitif sebagai dasar pengembangan tes kemampuan dasar (TKD) membaca sains; (2) menentukan karakteristik TKD membaca sains yang dikembangkan; (3) menentukan kriteria penilaian dan cara menginterpretasikan hasil TKD. Penelitian ini merupakan penelitian pengembangan. Populasi penelitian adalah siswa SD kelas 5 di kota Ungaran, kabupaten Semarang. Sampel ditentukan secara *clustering random sampling*, dengan ukuran 549 siswa. Instrumen berupa tes kemampuan dasar membaca, bertema Sains, berbentuk tes objektif 3 pilihan, sebanyak tiga buku tes. Analisis dilakukan dengan bantuan *software microCat ITEMAN, BILOG, dan LISREL*. Hasil penelitian berupa: (1) tiga buku tes TKD-membaca yang telah diketahui karakteristik soalnya dan (2) TKD-meter sebagai alat interpretasi hasil tes kemampuan dasar membaca. Karakteristik soal TKD-membaca meliputi indeks tingkat kesukaran (berada pada kategori mudah sampai sukar) dan indeks daya beda (berada pada kategori baik). TKD-meter yang dihasilkan terdiri atas empat level. Level I merupakan level terendah, dan level IV merupakan level tertinggi. Siswa berada pada level I, setara dengan kemampuan siswa kelas 3 SD. Siswa berada pada level II, setara dengan kemampuan siswa kelas 4 SD.

Kata kunci: *tes kemampuan dasar membaca, sains, psikologi kognitif*

## COGNITIVE PSYCHOLOGY AS A BASIC TO DEVELOP BASIC COMPETENCY IN SCIENCE READING

*Ani Rusilowati*

Pendidikan Fisika FKIP-UNES

rusilowati@yahoo.com

### **Abstract**

The objectives of this study are (1) to develop a basic reading skill test, (2) to find out the characteristics of the developed basic reading skill test, and (3) to set the criteria of evaluation and interpretation the test results. The research population was the fifth-grade students in primary schools in Ungaran, Semarang Municipality. The cluster random sampling was used to select the schools, and the sample size is 549 students. A multiple-choice-item basic reading skill test with 3 options and Science-themed was used to collect the data. The data was analyzed employing ITEMAN, BILOG and LISREL software. The test item characteristics were analyzed by Item Response Theory (IRT). The results of this study were: (1) three basic reading skill tests, with their item characteristics, and (2) *TKD-meter* (basic reading skill-meter) for interpreting the test result. The test item characteristics were item difficulty (from easy to difficult) and item discrimination (good categories). *TKD-meter* used in the interpretation of student's ability had four levels. The first level was *low ability* and the fourth was *high ability*. If the student's ability was categorized at Level I, equal to the third-grade students in primary schools. If the student's ability was categorized at Level II, equal to the fourth-grade students in primary schools.

Key word: *basic reading skill tests, science, cognitive psychology*

## **Pendahuluan**

Pendidikan mempunyai peran yang amat menentukan bagi perkembangan dan perwujudan diri individu, terutama bagi pembangunan bangsa dan negara. Hal ini dapat dilihat dari kemajuan di berbagai sektor pembangunan yang semuanya merupakan dampak dari hasil pendidikan. Kemajuan suatu negara sangatlah bergantung pada cara negara itu mengenali, menghargai, dan memanfaatkan sumber daya manusia, serta perhatiannya terhadap kualitas pendidikan yang diberikan kepada masyarakat. Oleh sebab itu, perhatian terhadap peningkatan kualitas pendidikan perlu terus ditingkatkan, agar dapat bersaing di era global.

Pemerintah telah melakukan upaya peningkatan kualitas pendidikan melalui peningkatan kualitas sumber daya manusia, pemberlakuan sistem akreditasi, pengembangan kurikulum, hingga penetapan rambu-rambu evaluasi dan sistem penilaian hasil belajar. Perbaikan kualitas pendidikan pada umumnya, dan sistem evaluasi dan penilaian hasil belajar pada khususnya, akan memberikan dampak yang besar bagi peningkatan kualitas pendidikan. Hal ini dapat dimengerti karena dengan evaluasilah kualitas pendidikan dapat diperbandingkan, baik secara regional maupun nasional.

Penghapusan EBTANAS di tingkat sekolah dasar menyebabkan tidak adanya instrumen pemantau mutu pendidikan untuk tingkat dasar. Pemerintah tidak dapat mengendalikan mutu pendidikan dasar yang sangat variatif. Oleh karena itu, diperlukan alat pemantau mutu pendidikan sekolah dasar (Lemlit UNY & Dinas Pendidikan DIY, 2004: 1). Salah satu upaya pemerintah untuk memantau mutu pendidikan sekolah dasar adalah menetapkan sistem penilaian hasil belajar yang sistematis dan berkelanjutan di sekolah dasar. Penilaian hasil belajar tersebut salah satunya adalah penilaian Tes Kemampuan Dasar, di samping penilaian kelas, ujian akhir sekolah, dan penilaian mutu pendidikan (Depdiknas, 2003d: 1). Untuk mengukur bahwa siswa SD menguasai kemampuan dasar perlu dilakukan tes kemampuan dasar. Tes untuk mengungkap kemampuan dasar ini disebut Tes Kemampuan Dasar (TKD). Depdiknas menetapkan kemampuan dasar untuk pendidikan tingkat dasar meliputi tiga keterampilan yaitu, membaca, menulis, dan berhitung yang disingkat

menjadi calistung. Pusat Penilaian Pendidikan, Balitbang Diknas telah menindaklanjuti permasalahan TKD ini dengan mengembangkan alat pemantau mutu pendidikan sekolah dasar berupa tes kemampuan dasar siswa kelas 3 SD. Tes ini masih bermanfaat jika suatu saat Ujian Nasional untuk SD diberlakukan kembali.

Di negara-negara maju seperti Amerika dan Australia, kemampuan dasar juga meliputi kemampuan membaca, menulis, dan berhitung. Salah satu tes kemampuan dasar yang dikembangkan di Amerika adalah *Iowa Tests of Basic Skills* (ITBS) dan di Australia *The Basic Skills Testing Program* (BSTP).

Maksud diberikan TKD adalah untuk mengetahui kemampuan membaca, menulis dan berhitung yang diperlukan dalam rangka memperbaiki program pembelajaran. Hasil tes ini diharapkan dapat memberi masukan tentang mutu pendidikan di sekolah dasar dan memberi manfaat untuk mengetahui kelemahan siswa dalam kemampuan dasar. Pemberian tes ini dilakukan sedini mungkin, agar kesalahan atau kelemahan siswa dapat segera diperbaiki.

Kenyataan saat ini menunjukkan bahwa ketersediaan TKD baku masih sangat terbatas. TKD yang dikembangkan oleh Balitbang Diknas masih dalam proses ujicoba, dan baru ditujukan untuk siswa kelas 3 SD. Ujicoba telah dilaksanakan secara nasional di 26 provinsi pada tahun 2003. Hasil ujicoba TKD, yang dilakukan di 36 SD di provinsi DIY, merekomendasikan beberapa hal berkaitan dengan perbaikan/pengembangan TKD. Rekomendasi yang berkaitan dengan TKD membaca adalah perlu mencermati atau bahkan merevisi butir soal yang berkaitan dengan kemampuan menentukan arti kata dalam kalimat, dan mendapatkan informasi yang tersirat dalam denah atau petunjuk (Lemlit UNY & Dinas Pendidikan DIY, 2004: 63). Di samping itu, kenyataan bahwa kemampuan dasar anak-anak Indonesia masih berada di bawah anak-anak dari negara ASEAN lain, juga mendorong dilakukannya penelitian ini. Rendahnya kemampuan dasar tersebut dapat dilihat dari hasil riset lembaga internasional, seperti IEA (*International Assosiation for the Evaluation of Educational Achievement*), PISA (*Programme for International Student Assessment*), dan PIRLS (*Progress in International Reading Literacy Study*). Hasil penelitian IEA pada tahun 1991, menunjukkan kemampuan membaca siswa SD

Indonesia berada pada posisi 26 dari 27 negara peserta studi. Pada tahun 1992, kemampuan membaca siswa SD Indonesia berada pada posisi 31 dari 32 negara peserta, dan pada tahun 2000 berada pada posisi 38 dari 39 negara peserta (Elley, 1992; Neni Utami, 2002: 8).

Hasil penelitian yang dilakukan oleh tim PISA berkerjasama dengan Balitbang Depdiknas menunjukkan bahwa kemahiran membaca anak usia 15 tahun di Indonesia masih memprihatinkan. Sekitar 37,6% hanya bisa membaca tanpa bisa menangkap maknanya, dan 24,8% hanya bisa mengaitkan teks yang dibaca dengan satu informasi pengetahuan (Kompas, 2 Juli 2003; Anis Suryani.) Hasil penelitian yang dilakukan oleh tim PIRLS tahun 2006, menunjukkan bahwa kemampuan membaca anak Indonesia berada jauh di bawah anak-anak dari negara lain. Kemampuan membaca anak Indonesia berada pada posisi 36 dari 40 negara peserta studi (IES *National Center for Educational Statistic*, 2007: 7). Perbedaan skor rata-rata antara jenis bacaan sastra dan informasi untuk anak Indonesia menempati posisi tinggi. Beda skornya sebesar 20. Beda skor ini lebih tinggi daripada beda skor anak Afrika Selatan, yang merupakan negara dengan posisi kemampuan membaca paling rendah. Bagi anak-anak dari negara Eropa, perbedaan skor antara bacaan sastra dengan informasi ini sangat kecil (antara 0 – 8) (*International Report IEA TIMSS & PIRLS*, 2007: 51). Rata-rata kemampuan membaca anak Indonesia ditinjau dari proses pemahaman bacaan, juga masih rendah. Skor rata-rata dalam menemukan informasi dan menarik kesimpulan dari informasi yang tertulis dalam bacaan sebesar 409. Skor rata-rata dalam menginterpretasikan, mengintegrasikan, dan mengevaluasi bacaan hanya sebesar 408. Skor tersebut masih di bawah *benchmark* yang sebesar 500 (*International Report IEA TIMSS & PIRLS*, 2007: 59).

Berdasarkan kenyataan-kenyataan tersebut, peneliti memandang perlu untuk mengembangkan TKD-membaca. Pengembangan soal TKD harus terus dilakukan, baik untuk merevisi soal yang sudah ada maupun mengembangkan soal baru. Pengembangan TKD untuk sekolah dasar perlu diperluas setidaknya sampai pada siswa kelas 5, mengingat manfaat dari TKD pada mutu pendidikan sangat besar. Oleh sebab itu, peneliti

bermaksud mengembangkan TKD untuk kelas 5, agar dapat memenuhi kebutuhan akan alat pemantau pendidikan di tingkat dasar.

Alasan lain yang mendorong peneliti menitikberatkan pada TKD-membaca adalah, kenyataan bahwa kemampuan membaca tidak hanya mengukur tentang kemampuan kebahasaan, tetapi juga kepada pemahaman terhadap muatan yang terkandung di dalam bacaan. Hal ini berimplikasi kepada pengembangan TKD-membaca. Mengingat kemampuan Sains anak Indonesia yang masih rendah, maka muatan isi yang dikembangkan dalam bacaan TKD-membaca adalah Sains. Setiap pertanyaan dalam TKD mengacu kepada suatu bacaan bertema Sains. Penetapan soal dan bacaan dalam TKD dilakukan dengan memperhatikan proses, strategi dan struktur pengetahuan kognitif siswa.

Dengan adanya TKD-membaca, kemampuan membaca anak dapat diidentifikasi sejak dini. Jika diketahui kemampuan membaca siswa masih rendah, maka sedini mungkin dapat diperbaiki. Penggunaan materi Sains sebagai muatan isi suatu bacaan dalam TKD-membaca diharapkan berdampak positif bagi penguasaan konsep Sains. Hal yang paling mendasar dilakukannya penelitian ini adalah tersedianya TKD-membaca yang relatif standar, yang dapat digunakan untuk mengukur kemampuan membaca siswa SD kelas 5.

Ketersediaan tes ini dapat juga dimanfaatkan sebagai alat diagnosis kelemahan dan kekuatan siswa dalam memahami bacaan. Agar tes yang dikembangkan dapat berfungsi secara efektif, maka ketika menyusun perlu memperhatikan tingkat perkembangan psikologis siswa, khususnya psikologi kognitif. Oleh sebab itu, melalui penelitian ini pengembangan tes dengan menerapkan teori kognitif mulai dilakukan. Strategi kognitif pengembangan tes yang digunakan mengacu pada strategi kognitif yang dikembangkan oleh Embretson (1995, 1998, 2001, 2002). Strategi ini menggunakan dua kerangka kerja, yaitu kerangka konseptual dan kerangka prosedural. Kerangka konseptual menitikberatkan pada validitas konstruk, sedangkan kerangka prosedural lebih kepada langkah-langkah pengembangan tes berdasarkan desain kognitif.

Mengingat kondisi serta fasilitas sekolah di berbagai sekolah masih beragam, maka diperlukan pedoman penyelenggaraan tes serta penggunaan TKD. Pada penelitian ini juga akan dihasilkan buku pedoman penggunaan dan pelaksanaan TKD bagi guru. Buku ini diharapkan dapat menjadi acuan

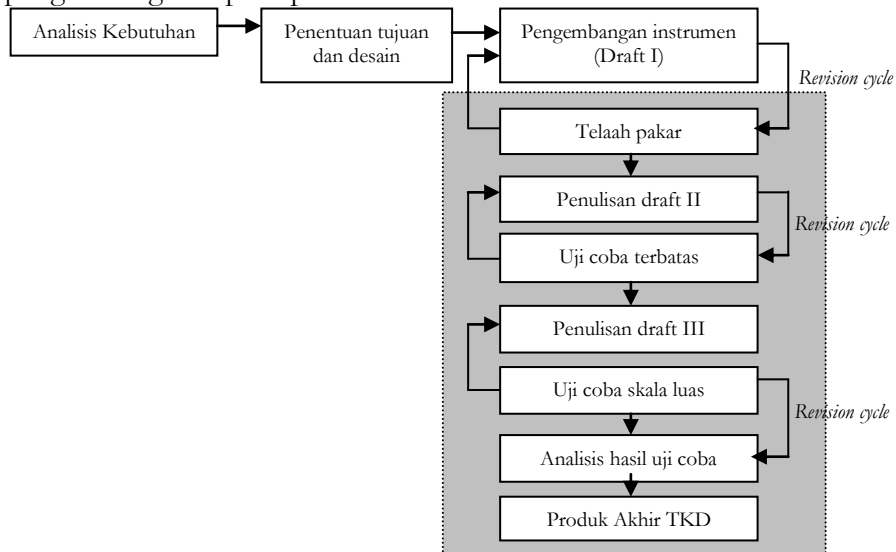
bagi penyelenggara pendidikan, sehingga pelaksanaannya terstandar untuk semua sekolah. Pada buku ini juga ditunjukkan cara menginterpretasikan hasil TKD, yaitu dengan TKD-meter. Pembuatan TKD-meter mengacu kepada penetapan level kemampuan dari *Profiles of Learning* yang dikembangkan oleh Masters, *et al.* (1990), *Benchmark Standar* dari Jackson & D'Andrea (2000), dan PIRLS (2007), dengan modifikasi.

Terkait dengan permasalahan di atas, penelitian ini bertujuan untuk: (1) mengetahui aplikasi psikologi kognitif pada model pengembangan soal TKD-membaca, (2) mengetahui karakteristik soal TKD-membaca yang dikembangkan, dan (3) menentukan kriteria dan cara menginterpretasikan hasil TKD-membaca.

## Metode Penelitian

### Rancangan Penelitian

Penelitian ini merupakan penelitian pengembangan (R & D). Model pengembangan seperti pada Gambar 1.



Gambar 1. Model Pengembangan TKD

## Ujicoba Produk

Ujicoba dimaksudkan untuk mengetahui karakteristik soal yang memenuhi syarat untuk TKD, serta mengetahui kecocokan model struktural yang menunjukkan validitas konstruk TKD. Di samping itu, ujicoba diperlukan untuk menentukan banyaknya waktu yang diperlukan untuk menyelesaikan setiap tes, untuk memperkirakan jumlah soal dan jumlah waktu yang sesuai.

### 1. Desain Ujicoba

Ujicoba produk pengembangan (TKD) dilakukan melalui dua tahap, yaitu validasi pakar dan uji lapangan. Validasi pakar pertama dilakukan oleh pakar (materi, pengukuran, psikologi) dalam hal ini adalah tim promotor dan guru kelas. Kedua, dikenakan pada beberapa siswa kelas 5 SD. Uji lapangan dilaksanakan dua kali, pertama ujicoba terbatas, kedua, pada saat implementasi.

### 2. Subjek Ujicoba

Validasi pakar dilakukan oleh para pakar pengukuran, pakar IPA, dan praktisi pendidikan, sedangkan subjek ujicoba lapangan, adalah siswa kelas 5 SD. Teknik pengambilan subjek ujicoba lapangan dilakukan secara *clustering random sampling*. Sekolah dipilih secara random, dari setiap kluster negeri dan swasta. Sekolah yang terpilih sebagai subjek ujicoba terbatas adalah SDN 01, 02, 04, SDIT Assalamah dan SDIT Istiqomah. Jumlah kelas ujicoba terbatas adalah 7 kelas, jumlah subjek ujicoba 275 siswa. Sekolah yang digunakan sebagai ujicoba implementasi, adalah SDN 01, 02, 03, 04, 06, SDIT Assalamah dan SDIT Istiqomah. Jumlah kelas implementasi ada 17 kelas dan jumlah subjeknya 549 orang.

### 3. Jenis Data

Jenis data yang diperoleh dari ujicoba produk ada dua macam, yaitu data kualitatif dan kuantitatif. Data kualitatif berupa masukan-masukan dari para ahli, baik yang diperoleh secara lisan/wawancara maupun pengisian kuesioner. Data kuantitatif berupa skor jawaban siswa terhadap TKD yang diujikan kepadanya.



#### 4. Instrumen Pengumpulan Data

Instrumen yang digunakan dalam ujicoba ini berupa kuesioner dan tes kemampuan dasar (TKD) membaca. Kuesioner ditujukan kepada guru dan sekelompok kecil siswa kelas 5 SD. TKD diujikan kepada siswa kelas 5 SD sebagai subjek ujicoba terbatas dan implementasi. TKD ini dikemas dalam bentuk Buku Tes. Pengembangan tes dilakukan dengan menggunakan strategi kognitif. Tes yang dikembangkan berbentuk pilihan ganda, dengan 3 pilihan jawaban. Tes ini lebih ditekankan pada tes pemahaman membaca.

#### 5. Teknik Analisis Data

##### a. Validasi konstruk TKD membaca

Kecocokan model struktural yang direncanakan dianalisis dengan teknik analisis faktor konfirmatori *second-order*. Analisis ini juga untuk menguji validitas konstruk dari TKD membaca. Untuk keperluan menguji validitas konstruk ini digunakan program *software* LISREL 8.54, didukung perangkat lunak lain seperti *Microsoft Office Excel* 2003 dan SPSS 10.0.

##### b. Pengujian terhadap Model Kognitif yang Dikembangkan

Ada dua model kognitif yang perlu diuji, yaitu kecocokan model pengukuran dan kalibrasi karakteristik soal TKD-membaca. Model pengukuran yang direncanakan untuk mengestimasi probabilitas respons siswa terhadap soal TKD adalah model dua parameter logistik (2-PL). Uji kecocokan model menggunakan program BILOG, dengan melihat keluaran pada *phase* 2. Soal dikatakan cocok dengan model apabila harga CHISQ hitung lebih rendah dari harga tabel pada derajat kebebasan dan taraf signifikansi tertentu.

Kalibrasi alat ukur berarti menentukan karakteristik soal yang meliputi tingkat kesukaran dan daya pembeda soal. Metode yang digunakan adalah IRT (*Item Respons Theory*). Program yang digunakan untuk keperluan ini adalah program BILOG. Estimasi probabilitas respons soal menggunakan model dua parameter logistik (2-PL). Model ini dipilih karena peneliti berasumsi bahwa siswa ketika mengerjakan tes tidak melakukan *guessing*.

c. Penyetaraan (*Equating*) Antarbuku Tes

TKD yang dikembangkan terdiri atas 3 Buku Tes. Masing-masing soal pada Buku Tes memiliki tingkat kesulitan yang berbeda-beda. Agar ketiga Buku Tes memiliki skala yang sama, maka dilakukan *equating* (penyetaraan) secara horisontal. Prosedur penyetaraan meliputi tiga tahapan proses, yaitu memilih desain pengumpulan data, menempatkan estimasi parameter pada skala yang sama, dan melakukan penyetaraan (Cook & Eignor, 1991; Kolen & Brennan, 1995). Desain pengumpulan data yang digunakan dalam penelitian ini adalah *single group desain*. Metode penyetaraan yang dipilih adalah *mean & sigma*.

d. Transformasi Skala

Indeks kesukaran soal hasil analisis dengan BILOG masih dalam bentuk skala logit. Oleh sebab itu, perlu ditransformasikan ke bentuk yang mudah dipahami. Transformasi skala yang digunakan adalah model WITs. Penetapan kategori setiap soal mengikuti ketentuan Hambleton, yaitu skala sebesar  $-2$  merupakan soal mudah, dan  $+2$  soal sukar. Ketentuan ini selanjutnya disesuaikan dengan skala transformasi dengan model WITs. Rentang indeks kesukaran soal ditetapkan dalam 3 kategori, yaitu mudah ( $b < 470$ ), sedang ( $470 \leq b \leq 530$ ), dan sukar ( $b > 530$ ).

e. Penginterpretasian Hasil TKD-membaca

Hasil uji dengan menggunakan TKD perlu diinterpretasikan agar dapat digunakan untuk melaporkan kemampuan membaca siswa. Pelaporan dapat dilakukan baik secara individual (untuk siswa dan orang tua), maupun klasikal bahkan regional. Oleh sebab itu, perlu dibuat alat penerjemah hasil. Alat ini berupa level-level kemampuan yang selanjutnya disebut TKD-meter. Pada penginterpretasian hasil TKD-membaca ini akan dibahas tentang pembuatan TKD-meter dan cara menggunakan TKD-meter tersebut.

TKD-meter dibuat untuk menginterpretasikan hasil TKD-membaca. TKD-meter ini dibuat dalam bentuk tingkatan-tingkatan atau level-level kemampuan membaca atau *benchmark*. Setiap level memiliki aspek kemampuan dan batas ambang (*cut-off value*) yang berbeda-beda. Penentuan *benchmark* mengikuti langkah-langkah sebagai berikut: (1) mengidentifikasi

kemampuan-kemampuan yang akan dicapai, (2) mengidentifikasi langkah-langkah untuk menentukan kemampuan, dan (3) menggunakan berbagai pendekatan untuk menetapkan batas standar (Jackson & D'Andrea, 2000: 89). Berdasarkan langkah-langkah tersebut dapat dibuat model proses penetapan *benchmark* dan batas ambang (*cut-off value*). Sedikitnya ada tiga cara penetapan *benchmark*, yaitu melalui identifikasi: tingkat kesukaran soal (Jordan, & Withnall, 1997), skor kemampuan dilanjutkan dengan tingkat kesukaran (IEA & PIRLS, 2007: 173-193), dan tingkat kesukaran dilanjutkan skor kemampuan (yang dilakukan dalam penelitian ini). Penggolongan level kemampuan mengacu pada Masters, *et al.* (1990), Darmiyati, Z., dkk. (2002), dan hasil analisis terhadap deskripsi kemampuan membaca.

Setelah TKD-meter dibuat, selanjutnya dapat digunakan untuk menentukan keberhasilan dan posisi kemampuan membaca siswa. Dengan membandingkan perolehan skor atau kemampuan ( $\theta$ ) siswa dengan rentang kemampuan dalam TKD-meter dapat diketahui keberhasilan dan posisi kemampuan siswa berada pada level I, II, III, atau IV. Apabila kemampuan siswa rendah, maka ia akan berada pada level I. Siswa dengan kemampuan tinggi akan berada pada level IV.

## **Hasil Penelitian dan Pembahasan**

Hasil penelitian ini berupa model pengembangan tes berdasarkan psikologi kognitif, karakteristik soal TKD-membaca, dan TKD-membaca yang karakteristik soalnya telah distandarisasi. Produk akhir berupa TKD-membaca, yang dikemas dalam tiga buah Buku Tes, dan manual penggunaannya. Aplikasi psikologi kognitif pada pengembangan soal TKD-membaca dilakukan sejak penetapan indikator kemampuan membaca. Pada penelitian ini, langkah pengembangan TKD-membaca meliputi penetapan indikator kemampuan membaca, pendefinisian konstruk kemampuan membaca, pemvalidasian konstruk pada kerangka konseptual, penerapan model kognitif dalam penulisan soal pada kerangka prosedural, dan pengevaluasian soal yang dikembangkan.

Produk dari tahap prapengembangan meliputi: kesesuaian teori kognitif dengan kemampuan membaca, indikator kemampuan membaca berdasarkan hipotesis, indikator kemampuan membaca berdasarkan hasil workshop, topik bacaan setiap Buku Tes, spesifikasi tes, dan kisi-kisi soal. Pada tahap ini juga dihasilkan draf manual pelaksanaan tes untuk panduan guru.

Produk pada tahap pengembangan adalah sekelompok soal yang dikemas dalam 3 Buku Tes. Buku Tes 1 terdiri atas 36 soal, Buku Tes 2 sebanyak 35 soal, dan Buku Tes 3 sebanyak 36 soal. Buku Tes-Buku Tes ini merupakan draf I dari TKD membaca yang akan dikembangkan. Tes yang dihasilkan selanjutnya diujicobakan secara terbatas kepada sekelompok siswa kelas 5 SD. Jumlah subjek ujicoba sebanyak 275 siswa. Hasil uji coba diperoleh besarnya reliabilitas buku tes 1, 2, dan 3 berturut-turut adalah 0,687; 0,733; dan 0,743. Daya beda soal pada buku tes 1 dan 2 ada yang negatif, sedangkan buku tes 3 semua positif. Beberapa soal pada buku tes 1, 2 dan 3 pengecohnya belum berfungsi. Dengan demikian soal-soal tersebut perlu direvisi.

Data yang digunakan untuk uji validitas konstruk adalah skor TKD yang diperoleh ketika siswa mengerjakan tes ujicoba. Analisis data menggunakan program LISREL 8.54. Sebelum analisis dilakukan, persyaratan uji normalitas dan multikolinearitas data harus dipenuhi. Hasil analisis menunjukkan bahwa muatan faktor ( $\lambda$ ) dari variabel laten kemampuan pemahaman kosakata ( $\eta_1$ ), pemahaman literal ( $\eta_2$ ), pemahaman inferensial ( $\eta_3$ ) dengan variabel-variabel indikatornya secara statistik signifikan. Hasil tersebut menunjukkan bahwa variabel laten pemahaman kosakata ( $\eta_1$ ) secara teoretis dapat diukur dengan indikator Y1-Y3, pemahaman literal ( $\eta_2$ ) secara teoretis dapat diukur dengan indikator Y4-Y8, pemahaman inferensial ( $\eta_3$ ) secara teoretis dapat diukur dengan indikator Y9-Y20, dan aplikasi ( $\eta_4$ ) secara teoretis dapat diukur dengan indikator Y21-Y22.

Menurut (Garson, 2006), dukungan terhadap kecocokan model yang dikembangkan oleh data empirik paling tidak dilihat dari tiga ukuran kecocokan yang mewakili tiga kategori uji kecocokan model yang berbeda. Kategori yang dipilih untuk menentukan kecocokan model konstruk pada

penelitian ini adalah RMSEA, CFI dan PGFI. Hasil analisis menunjukkan bahwa besarnya RMSEA adalah 0,0545. Harga ini lebih kecil dari *cut off value* dari RMSEA yang sebesar 0,08. Besarnya nilai CFI yang diperoleh dari hasil analisis adalah 0,915. Harga ini lebih besar dari harga *cut off value* dari CFI yang sebesar 0,90. Harga PGFI yang diperoleh dari hasil analisis hanya sebesar 0,860. Dua dari tiga kategori uji kecocokan adalah signifikan. Berarti konstruk kemampuan membaca cocok dengan model yang dikembangkan. Hal ini juga didukung oleh harga *factor loading* dari setiap indikator. Hasil ini menggambarkan bahwa tes kemampuan membaca dikonstruksi atas 4 variabel laten yang mempunyai indikator  $Y_i$ , dengan  $i = 1, 2, \dots, 22$ .

Soal TKD yang telah direvisi, dan dipilih sesuai dengan aspek kemampuan, diujikan kembali kepada 549 siswa kelas 5 SD. Jumlah soal untuk masing-masing Buku Tes sebanyak 22 soal, dipilih soal dengan daya beda positif dan cocok dengan model. Berdasarkan hasil analisis dengan program BILOG tidak ada soal yang memiliki daya beda negatif, baik pada Buku Tes 1, Buku Tes 2, maupun Buku Tes 3. Dengan demikian semua soal memenuhi syarat untuk dianalisis lebih lanjut dengan IRT. Hasil karakteristik soal setelah dianalisis dengan IRT, disetarakan, dan ditransformasikan adalah sebagai berikut: (1) Buku Tes 1 memiliki 1 soal sukar, 3 soal sedang, 18 soal mudah, (2) Buku Tes 2 memiliki 4 soal sedang, 18 soal mudah, dan (3) Buku Tes 3 memiliki 1 soal sukar, 2 soal sedang, 19 soal mudah.

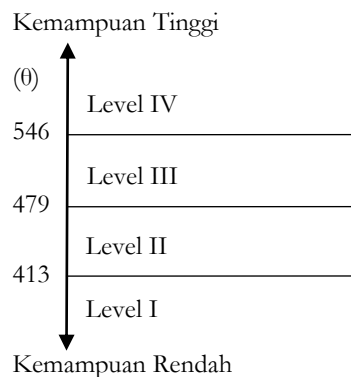
Hasil uji dengan menggunakan TKD perlu diinterpretasikan agar dapat digunakan untuk melaporkan kemampuan membaca siswa. Pelaporan dapat dilakukan baik secara individual (untuk siswa dan orang tua), maupun klasikal bahkan regional. Oleh sebab itu perlu dibuat alat penerjemah hasil. Alat ini berupa level-level kemampuan yang selanjutnya disebut TKD-meter. Pembuatan TKD-meter meliputi penentuan *benchmark* dan batas ambang tiap level.

Berdasarkan hasil analisis empiris dapat ditentukan deskripsi kemampuan membaca menurut level kemampuan dan urutan tingkat kesukaran. Kemampuan membuat analogi merupakan kemampuan paling sukar, berada pada level IV. Kemampuan paling mudah adalah

menemukan informasi dalam bacaan, dan berada pada level I. Berdasarkan kemampuan tertinggi, terendah, dan interval tiap level yang telah dihitung sebelumnya, serta setelah ditransformasikan ke skala WTTs, dapat ditentukan rentang kemampuan untuk tiap levelnya sebagai berikut.

Level IV:  $(\theta) \geq 546$ ; Level III:  $479 \leq (\theta) < 546$ ; Level II:  $413 \leq (\theta) < 479$ ; dan Level I:  $(\theta) < 413$ .

Rentang kemampuan ini dapat diletakkan pada TKD-meter sesuai dengan level kemampuan, dan TKD-meter dapat digunakan untuk menentukan posisi kemampuan membaca siswa. Berdasarkan penetapan rentang kemampuan berdasarkan  $(\theta)$  dapat ditentukan TKD-meter seperti pada Gambar 2.



Gambar 2. TKD-meter Berdasar Kemampuan  $(\theta)$  Siswa

TKD-meter dapat digunakan untuk menentukan posisi kemampuan membaca siswa. Dengan menghitung tingkat kemampuan  $(\theta)$  dan membandingkannya dengan batas-batas skor yang ada di TKD-meter, dapat ditentukan posisi kemampuan membaca seorang siswa. Selanjutnya dapat ditentukan letak kekuatan dan kelemahan siswa dalam kemampuan membaca untuk pemahaman.

Agar TKD dapat digunakan oleh semua orang, maka perlu dilengkapi dengan manual. Manual Penggunaan TKD dimaksudkan untuk memandu guru dalam mengujikan TKD-membaca kepada siswa. Panduan

ini berisi informasi tentang tujuan dan materi TKD-Membaca, persiapan testing, petunjuk pelaksanaan, petunjuk penskoran, penafsiran hasil tes, dan penggunaan hasil tes dalam pengembangan pengajaran. Hal yang perlu diperhatikan dalam mempersiapkan testing antara lain adalah: kualifikasi pelaksana tes, keseragaman, penyusunan fisik, penjadwalan tes, halaman praktik, dan sikap murid untuk testing. Dengan adanya manual penggunaan ini, diharapkan guru dapat menggunakan TKD-Membaca dengan mudah.

## **Simpulan dan Saran**

### **Simpulan**

1. Penetapan indikator kemampuan membaca telah dikaitkan dengan kemampuan kognitif, seperti operasi kombinasi/klasifikasi, asosiasi, reversibilitas, dan identitas. Di samping itu, kemampuan membaca pemahaman juga dikaitkan dengan aspek kognisi yang meliputi jenjang ingatan, pemahaman, aplikasi, analisis, sintesis, dan evaluasi.
2. Produk dari penelitian ini adalah TKD-membaca yang dikemas dalam bentuk Buku Tes. TKD-membaca ini dikembangkan untuk mengukur kemampuan membaca siswa kelas 5 SD dan yang sederajat, khususnya membaca pemahaman terhadap bacaan bidang Sains. Kemampuan membaca ini tersusun dari konstruk kemampuan pemahaman kosakata, pemahaman literal, pemahaman inferensial, dan aplikasi.
3. Indikator yang digunakan untuk mengungkap kemampuan membaca sebanyak 22, tersebar dalam 4 level kemampuan. Ketiga Buku Tes yang dikembangkan memiliki karakteristik soal dengan indeks daya beda (*slope*) berkisar 0,074 sampai dengan 1,174 pada kategori baik dan tingkat kesukaran (*threshold*) berkisar antara 354 sampai dengan 537 (skala WITs), rata-rata pada kategori mudah sampai sedang. Dari tiga Buku Tes yang dikembangkan hanya ada dua soal pada kategori sukar. Ketiga Buku Tes telah disetarakan, sehingga memiliki skala yang sama dan dapat diperbandingkan.
4. Interpretasi hasil TKD-membaca menggunakan TKD-meter, yang berbentuk garis vertikal yang terbagi dalam level-level kemampuan membaca dan memuat berbagai indikator yang mengukurnya. Level

kemampuan membaca dibagi dalam 4 level, Level IV merupakan level tertinggi dan Level I merupakan level terendah. Setiap level memiliki aspek kemampuan yang berbeda. Kriteria level kemampuan pada TKD-meter ditentukan berdasarkan tingkat kemampuan siswa ( $\theta$ ) sebagai berikut: Level IV :  $(\theta) \geq 546$  ; Level III:  $479 \leq (\theta) < 546$ ; Level II:  $413 \leq (\theta) < 479$ ; dan Level I :  $(\theta) < 413$ . Cara menginterpretasikan hasil TKD-membaca yaitu membandingkan kemampuan yang diperoleh siswa ketika mengerjakan tes dengan level kemampuan pada TKD-meter.

### **Saran**

Berdasarkan hasil penelitian, dapat disarankan hal-hal sebagai berikut: (1) tema bacaan untuk mengukur kemampuan membaca sebaiknya divariasikan sesuai dengan kelompok-kelompok bidang studi, (2) perlu menambah bacaan yang berkaitan dengan sastra, (3) soal yang mengacu pada bacaan sebaiknya dianalisis dengan model *polytomous*.

### **Daftar Pustaka**

- Anderson, J. R. (1995). *Cognitive psychology and its implications (4<sup>th</sup> ed.)*. New York: W.H. Freeman Company.
- Cook, L. L., & Eignor, D. R. (1991). IRT equating methods. *Journal of Educational Measurement*, 37-45.
- Darmiyati Zuchdi, Sukamto, dan Suryanto. (2002). *Pengembangan alat ukur peringkat keterampilan membaca, menulis, dan matematika pada jenjang sekolah dasar*. Laporan Penelitian. Yogyakarta: UNY.
- Elley, W. B. (1992). *How in the world do student read*. Hamburg: Grindeldruck GMBH.
- Embretson, S. E. (1995). A measurement model for linking individual learning processes and knowledge: Application to mathematical reasoning. *Journal of Educational Measurement*, 32 (3), 277-294.



- \_\_\_\_\_. (1998). A cognitive design system approach to generating valid test: Application to abstract reasoning. *Psychological Method*, 3, 380-396.
- \_\_\_\_\_. (1999). Generating item during testing: Psychometric issues and models. *Psychometrika*, 64, 407-433.
- \_\_\_\_\_. (2002). Generating abstract reasoning item with cognitive theory. Dalam Irvine, S. H., & Kyllonen, P. C. *Item generation for test development* (pp 219-250). Mahwah, N. J.: LEA Publisher.
- Embretson, S. E., & Gorin, J. S. (2001). Improving construct validity with cognitive psychology principles. *Journal of Educational Measurement*, 38 (4), 343-368.
- Garson, D. (2006). *Structural equation modeling*. Diambil pada tanggal 20 Agustus 2006, dari <http://www2.chass.ncsu.edu/garson/pa765/structur.htm>.
- Gorin, J. S. (2006a). Test design with cognition in mind. *Educational Measurement: Issues and Practice*, 25 (4), 21-35.
- \_\_\_\_\_. (2006b). Manipulation of processing difficulty on reading comprehension test questions: The feasibility of verbal item generation. *Journal of Educational Measurement*, 42, 351-373.
- Gorin, J. S., & Embretson, S. E. (2006). Item difficulty modeling of paragraph comprehension item. *Applied Psychological Measurement*, 30 (5), 394-411.
- Hieronimus, A.N., Linnquist, E.F., & Hoover, H.D. (1982). *Manual for school administrators Iowa tests of basic skills*. Chicago: The Riverside Publishing Company.
- IEA TIMSS & PIRLS International Study Center. (2007a). *PIRLS 2006 International Report*. Diambil pada tanggal 7 April 2007, dari <http://pirls.bc.edu/pirls2006/intl-rpt.html>.

- IEA TIMSS & PIRLS International Study Center. (2007b). *PIRLS 2006 Technical Report*. Diambil pada tanggal 7 April 2007, dari <http://pirls.bc.edu/pirls2006/tech-rpt.html>.
- IES National Center for Educational Statistic. (2007). *The reading literacy of U. S. fourth-grade students in an international context*. Diambil pada tanggal 7 April 2007, dari <http://nces.ed.gov/survey/pirls/>
- Jackson, N., & D'Andrea V. (2000). Benchmarking the outcomes of learning. Dalam Jackson, N., & Lund, H. (Eds). *Benchmarking for higher education*. (pp 85-105). London: SHRE and Open University Press.
- Jordan, A., & Withnall. (1997). *The core attributes of biological sciences graduates*. London: Institute of Biology.
- Lemlit UNY & Dinas Pendidikan DIY. (2004). *Laporan kegiatan: Analisis tes kemampuan dasar siswa sekolah dasar di DIY*.
- Kolen, M. J., & Brennan, R. L. (1995). *Test equating, scaling, and linking methods and practices* (2nd ed.). New York: Spring Science & Business Media, LLC.
- Masters, G. et al. (1989). *Profiles of learning*. Hawton, Victoria: ACER.
- Neni Utami Adiningsih. (2002). Mencermati generasi gagap aksara. *Pikiran Rakyat Cyber Media* Kamis, 07 Nopember 2002.