

KESEJAJARAN BENTUK TES PILIHAN GANDA DENGAN TES CLOZE DALAM TES MEMBACA PEMAHAMAN

Oleh: Burhan Nurgiyantoro

Abstrak

Ada banyak cara yang distandarkan untuk mengukur kemampuan membaca, antara lain mempergunakan bentuk betul-salah, melengkapi kalimat, pilihan ganda, pembuatan ringkasan atau rangkuman, cloze test, C-test. Teknik yang populer dipakai adalah format bentuk tes pilihan ganda. Namun, format tersebut sering dikritik karena jawaban benar dapat diperoleh lewat lebih dari satu cara, misalnya dengan cara menebak di samping adanya keraguan kemampuan siswa memahami sungguh-sungguh wacana yang ditekankan karena tanpa adanya penilaian dalam pemilihan jawaban yang benar. Proses pemilihan jawaban yang benar belum tentu mencerminkan proses yang melibatkan sebagaimana dalam konteks membaca yang sebenarnya. Untuk mengatasi keadaan itu dapat dipergunakan tes bentuk lain, yang dalam hal ini adalah tes cloze.

Sebagaimana terlihat dalam dukungan teori dan kemudian diperkuat lewat penelitian pengembangan alat tes tes membaca pemahaman, terdapat kesejajaran antara bentuk tes pilihan ganda dengan tes cloze. Hal itu menyuratkan suatu pengertian bahwa kedua bentuk tes tersebut sama-sama dapat dipakai untuk menyadap kemampuan membaca pemahaman. Keduanya dapat dipakai secara bersama, saling melengkapi dan saling menggantikan, dan akan memberikan hasil yang kurang lebih sama.

A. Pendahuluan

Tes kemampuan membaca adalah sebuah tes keterampilan berbahasa yang biasa dilakukan dalam pengajaran bahasa, baik dalam pengajaran bahasa pertama maupun bahasa kedua (asing). Kemampuan membaca merupakan salah satu dari keempat keterampilan berbahasa yang diajarkan —dan karenanya juga berkonsekuensi ditekankan— kepada pembelajar bahasa. Bersama dengan kemampuan menyimak kemampuan membaca tergolong kemampuan aktif reseptif, tetapi berbeda media penyampaiannya. Kemampuan menyimak dipergunakan untuk mengukur kemampuan memahami bahasa lisan, sedang kemampuan membaca untuk bahasa tulis. Ada banyak cara yang distandarkan

untuk mengukur kemampuan membaca. Sejumlah teknik pengukuran kemampuan membaca yang sering dipergunakan antara lain adalah dengan mempergunakan bentuk betul-salah, melengkapi kalimat, pilihan ganda, pembuatan ringkasan atau rangkuman, *cloze test*, *C-test*, dan lain-lain (Anderson, 1991:42). Teknik yang paling umum dipakai adalah format bentuk tes pilihan ganda. Namun demikian, format tersebut sering dikritik karena jawaban benar dapat diperoleh lewat lebih dari satu cara, misalnya dengan cara menebak. Di samping itu, juga diragukan kemampuan siswa memahami dengan sungguh-sungguh wacana yang diteskan karena tanpa adanya penilaian dalam pemilihan jawaban yang benar (Klein-Braley lewat Anderson, 1991:42). Dengan demikian, proses pemilihan jawaban yang benar belum tentu mencerminkan proses yang melibatkan sebagaimana dalam konteks membaca yang sebenarnya (Anderson, 1991:42).

Untuk mengatasi kritik tersebut, usaha pengukuran kemampuan membaca dapat ditempuh dengan mempergunakan lebih dari satu teknik. Misalnya, di samping dipergunakan bentuk pilihan ganda juga dipakai bentuk lain sebagai pendamping seperti teknik *cloze*. Teknik *cloze* juga cukup populer dan banyak dipergunakan untuk mengukur kemampuan membaca (Oller, 1979; Jonz, 1991; Brown, 1993), khususnya dalam pengajaran bahasa kedua. Namun, teknik *cloze* juga sering dipergunakan untuk mengukur kemampuan membaca oleh penutur asli. Jon Jonz (1991) misalnya, mengukur kemampuan membaca bahasa Inggris oleh mahasiswa Amerika Serikat dan pembelajar (mahasiswa) asing yang bukan penutur asli. Skor hasil pengukuran kedua kelompok pembelajar tersebut kemudian dibandingkan untuk mengetahui signifikansi perbedaannya. Penelitian itu menunjukkan adanya perbedaan skor yang diperoleh oleh mahasiswa Amerika yang penutur asli bahasa Inggris dengan mahasiswa asing yang bukan penutur asli (Jonz, 1991:10).

Berdasarkan pengamatan selama ini di sekolah dan atau perguruan tinggi, misalnya dalam tes-tes sumatif akhir semester siswa sekolah menengah, Ebtanas, tes masuk perguruan tinggi, ataupun ujian-ujian semester mahasiswa jurusan Pendidikan Bahasa dan Sastra Indonesia dalam mata kuliah "Membaca", terlihat bahwa teknik *cloze* belum populer dipergunakan untuk mengukur kemampuan membaca siswa / mahasiswa penutur asli bahasa

Indonesia. Pengukuran kemampuan membaca bahasa Indonesia terlihat masih banyak mempergunakan teknik bentuk pilihan ganda. Oleh karena itu, bagaimana efektivitas teknik *cloze* untuk mengukur kemampuan membaca penutur asli bahasa Indonesia belum banyak diketahui.

Untuk itu, penggunaan tes *cloze* perlu segera dirintis bagi penutur asli bahasa Indonesia sebagai pendamping atau pelengkap bentuk tes kemampuan membaca yang lain yang sudah biasa dipakai khususnya pilihan ganda. Penelitian kecil ini —yang dapat dipandang sebagai suatu bentuk pengembangan atau uji coba bentuk atau format tes— dilakukan untuk menjajagi efektivitas penggunaan tes *cloze* sebagaimana dimaksudkan di atas. Untuk maksud itu, bentuk tes *cloze* tersebut sengaja dikembangkan bersama dengan bentuk tes pilihan ganda.

B. Aspek Dan Bentuk Tes Memebaca Pemahaman

Tes membaca pemahaman dimaksudkan untuk menyadap kemampuan peserta tes dalam memahami informasi yang terkandung dalam sebuah wacana. Dalam kegiatan sehari-hari orang pun sering berurusan dengan berbagai wacana yang harus dipahami informasinya untuk berbagai kebutuhan. Oleh karena itu, tes membaca pemahaman dapat dikategorikan sebagai tes pragmatik, tes komunikatif, tes integratif, dan paling tidak bukan sebagai tes diskrit.

Dalam tes membaca pemahaman masalah yang dipersoalkan adalah aspek apa saja yang perlu ditanyakan kepada peserta untuk dapat mengetahui seberapa banyak mereka dapat memahami informasi yang terkandung dalam wacana yang dibacanya. Untuk maksud pengukuran itu, ada sejumlah aspek yang biasa dilibatkan dalam tes membaca. Blaricom (1981:62) misalnya, menganggap bahwa paling tidak ada dua hal yang perlu ditanyakan, yaitu organisasi wacana dan ide pokok wacana. Organisasi wacana antara lain berupa hubungan antarkata, antarkalimat, antraparagraf. Secara per kata, per kalimat, atau per paragraf, mungkin saja kita mengetahui makna yang terkandung, namun bagaimana keterkaitan antarkata, kalimat, atau paragraf itu sehingga membentuk sebuah wacana yang lebih besar, belum tentu dapat dipahami dengan baik. Namun, kesulitan pemahaman itu mungkin saja disebabkan kurang baiknya pengorganisasian ide dalam wacana yang

bersangkutan. Dalam penelitiannya itu Blaricom (1981:71) menemukan bahwa organisasi wacana berpengaruh secara signifikan terhadap pembaca yang "kurang mampu", namun kurang berpengaruh terhadap pembaca yang lebih mampu.

Pemahaman terhadap organisasi wacana juga amat membantu untuk memahami ide-ide, termasuk ide pokok, yang terdapat dalam sebuah wacana. Pemahaman terhadap ide-ide (baca: informasi) yang terdapat dalam suatu wacana merupakan hal yang penting. Menurut Blaricom (1981:65) yang termasuk dalam aspek ini adalah pemahaman topik (tema?) keseluruhan wacana, penemuan judul (yang mencerminkan topik), dan pemahaman ide pokok, termasuk menemukan kalimat yang menyatakannya, dalam tiap paragraf.

Dalam penelitiannya untuk menguji validitas konstruk sebuah tes membaca pemahaman yang diambil dari *Educational Testing Service*, Anderson dkk (1991) mengkategorikan pertanyaan-pertanyaan bacaan ke dalam tiga aspek, yaitu pemahaman ide pokok, pembuatan inferensi, dan pemahaman pernyataan langsung. Pertanyaan terhadap ketiga aspek tersebut dikatakan dapat menjamin untuk mengukur keterampilan-keterampilan yang berbeda yang semestinya diukur dalam tes itu. Dalam penelitiannya itu Anderson dkk (1991:51) menemukan adanya hubungan yang signifikan antara jenis-jenis pertanyaan tersebut dengan strategi membaca dan menjawab pertanyaan oleh peserta tes.

Pearson dan Johnson (lewat Anderson dkk, 1991) membedakan hubungan antara pertanyaan dan jawaban tes membaca pemahaman ke dalam tiga kategori, yaitu pertanyaan yang jawabannya ada secara tersurat dalam wacana (*textually explicit*), yang hanya ada secara tersirat (*textually implicit*), dan yang tersirat di luar wacana (*scriptally implicit*). Penelitian Anderson dkk (1991:52) menunjukkan tidak adanya hubungan yang signifikan antara ketiga kategori tersebut dengan strategi membaca dan menjawab peserta tes, namun signifikan dengan ketiga aspek pertanyaan di atas.

Bentuk tes yang dipergunakan untuk mengukur membaca pemahaman yang berkaitan dengan pengukuran aspek-aspek di atas pada umumnya adalah bentuk pilihan ganda. Tes pilihan ganda merupakan suatu bentuk pertanyaan yang menghendaki subjek yang dites untuk memilih salah satu alternatif

jawaban yang disediakan. Pilihan jawaban (*option*) yang disediakan pada umumnya berkisar antara tiga sampai lima butir, sedang jawaban yang benar (biasanya dikatakan: paling tepat) hanya satu butir. Cara penyusunan tes bentuk itu sendiri ada beberapa variasi.

Tes bentuk pilihan ganda pada hakikatnya tidak berbeda dengan tes bentuk benar salah karena sama-sama menuntut subjek untuk memilih di antara jawaban yang benar dan salah yang sengaja disediakan. Yang membedakan adalah tes bentuk pilihan ganda lebih banyak menyediakan jawaban yang salah lebih dari satu. Selain itu, faktor untung-untungan dalam menjawab pertanyaan sama-sama ditemukan dalam kedua bentuk tes tersebut. Jika dalam tes benar salah peluang menjawab benar atau salah secara untung-untungan sama besar, dalam tes pilihan ganda peluang itu lebih kecil, yaitu dengan persentase sebesar satu dibagi jumlah *option* kali 100 (jika jumlah *option* = n , rumusnya adalah $1 : n \times 100$). Semakin banyak jumlah *option* yang disediakan, semakin kecil peluang jawaban benar yang dipilih hanya secara spekulatif. Adanya peluang untuk berspekulasi inilah antara lain, seperti dikemukakan di atas, yang menjadi keberatan orang untuk mempergunakan tes bentuk itu sebagai satu-satunya cara, sehingga perlu dilengkapi dengan bentuk dan atau cara lain yang dapat mengurangi kelemahan tersebut. Untuk tes kemampuan membaca pemahaman, bentuk tes yang dimaksud dalam tulisan ini adalah tes *cloze*.

Bentuk pilihan ganda memang jauh lebih populer daripada tes *cloze*, Walau demikian, hal itu tidak usah diartikan bahwa tes pilihan ganda lebih baik daripada tes *cloze*. Ada banyak faktor mengapa orang lebih memilih tes pilihan ganda. Sebagai akibatnya, tes *cloze* juga jarang dimanfaatkan untuk tes bahasa Indonesia untuk penutur asli sehingga informasi tentang tes itu tidak mudah diperoleh.

C. Tes Cloze

Tes *cloze* merupakan suatu tugas untuk melengkapi kembali kata-kata yang sengaja dihilangkan pada sebuah wacana. Penghilangan kata-kata itu biasanya bersifat sistematis, yaitu setiap kata yang ke- n (5, 6, atau 7). Oller (1979:345) memandang tes *cloze* sebagai sebuah tugas pragmatik —dan juga merupakan salah satu bentuk tes komunikatif— berhubung pengisian kata-kata

tersebut harus tidak semata-mata mempertimbangkan kelayakan kebahasaan, melainkan juga ketepatan makna konteks. Menurut Oller tes *cloze* mampu untuk mengukur berbagai kemampuan berbahasa, bahkan juga dapat dipergunakan untuk menilai tingkat kesulitan sebuah wacana (Oller, 1979:348).

Menurut berbagai penelitian dalam *ESL* (Brown, 1993; juga Oller, 1979) tes *cloze* mempunyai korelasi yang signifikan dengan berbagai keterampilan berbahasa (menyimak, berbicara, membaca, dan menulis). Penelitian Djiwandoro (1990:95) tentang tes *cloze* bahasa Indonesia untuk penutur asli juga menunjukkan adanya korelasi antara tes *cloze* dengan mengarang dan tata bahasa, yaitu masing-masing sebesar 0,59 dan 0,37. Dalam berbagai penelitian tentang tes *cloze* tersebut berdasarkan pengamatan Brown (1993:94) tingkat reliabilitas dan validitas bervariasi antara penelitian yang satu dengan penelitian yang lain. Tingkat reliabilitas berkisar antara 0,31 sampai 0,96, sedang tingkat validitas yang dicari dengan mengorelasikannya dengan hasil tes TOEFL berkisar antara 0,19 sampai 0,83. Menurut Brown faktor-faktor yang mempengaruhi tingkat reliabilitas dan validitas tes *cloze* tersebut adalah (1) cara penyekoran, (2) banyaknya kata-kata yang dihilangkan, (3) frekuensi penghilangan atau jarak antara kata-kata yang dihilangkan, (4) tingkat keterbacaan wacana, (5) hasil tes untuk penutur asli dan bukan penutur asli, dan panjangnya tes atau jumlah keseluruhan butir tes.

Jonz (1991) membedakan jenis-jenis pertanyaan —yang berupa kata-kata yang dihilangkan— tes *cloze* ke dalam empat kategori, yaitu pertanyaan (1) dalam satu klausa yang dapat berwujud sintaktik dan kosa kata, (2) antarklausa dalam satu kalimat, (3) antarkalimat dalam teks, dan (4) ekstratekstual. Namun, dalam penelitiannya itu tidak ditemukan perbedaan yang signifikan antara nilai rata-rata untuk keempat jenis pertanyaan tersebut (Jonz, 1991:13).

D. Kesejajaran Bentuk Tes Pilihan Ganda dengan Tes Cloze dalam Tes Membaca Pemahaman

1. Uji Pengembangan Alat Tes

Kesejajaran antara kedua bentuk tes membaca pemahaman yang

disajikan berikut didasarkan pada hasil penelitian pengembangan terhadap kedua bentuk tersebut.

Peserta tes adalah mahasiswa STIE Malangkusewara Malang tingkat I (semester 1) yang masih dan sedang menempuh mata kuliah Bahasa Indonesia (MKDU). Jumlah mahasiswa peserta tes adalah 33 orang. Sebagai penutur asli dan telah bertahun-tahun diajar bahasa Indonesia sampai tingkat perguruan tinggi, tentunya mereka dapat dipandang sebagai penutur asli yang terpelajar sehingga kemampuan berbahasa Indonesianya diharapkan cukup baik. Bahan tes berupa wacana yang diambil dari tiga buah surat kabar, yaitu *Kompas*, *Republika*, dan *Jawa Pos*. Wacana yang diambil adalah berupa artikel ilmiah yang berisi masalah teknologi, teknologi-ekonomi, dan perekonomian makro yang ditulis tokoh-ahli dalam bidang yang bersangkutan. Pemilihan topik itu disebabkan subjek yang akan dites adalah mahasiswa ekonomi sehingga sedikit banyak hal itu diharapkan menarik minat mereka.

Alat tes yang dikembangkan terdiri atas dua bentuk, yaitu tes pilihan ganda dan *cloze*, keduanya terdiri dari 30 butir soal. Wacana tes pilihan ganda adalah 6 buah, dan sengaja diurutkan dari yang pendek ke yang lebih panjang agar secara psikologis mahasiswa tidak merasa langsung dihadapkan pada tugas yang berat. Wacana tes *cloze* diambil dari artikel yang sama dengan wacana tes pilihan ganda. Dengan demikian, secara teoretis tingkat keterbacaan keduanya seimbang. Wacana terdiri dari atas lima alinea pendek. Penghilangan kata berkisar antara kata yang ke-5 sampai ke-9 karena penghilangan itu ditujukan pada kata-kata tertentu yang secara makna penting di samping juga menghindari kata-kata tugas. Kalimat pertama dan terakhir sengaja dibiarkan utuh, sedang yang lain setiap kalimat paling tidak terdapat dua kata yang dihilangkan. Penyekoran dilakukan dengan dasar kelayakan konteks, maka untuk keperluan pengoreksian sengaja disiapkan daftar kata yang secara konteks dapat mengisi tempat kosong tersebut untuk tiap butir soal. Validitas alat tes pilihan ganda dicapai dengan validitas konstruk, yaitu sebuah validitas yang pencapaiannya dilakukan dengan mencocokkan butir-butir tes dengan teori yang melatarbelakanginya. Kemampuan peserta membuat inferensi-inferensi yang berkaitan dengan informasi wacana penting ditanyakan, yaitu yang berupa ide pokok dan informasi yang terkandung dalam wacana itu baik secara tersurat maupun tersirat (Anderson dkk, 1991). Dilihat dari segi isi tes,

pertanyaan yang dikembangkan terdiri atas tiga kategori pertanyaan, yaitu (1) tema, (2) informasi yang terkandung, dan (3) hubungan antaralinea, sedang dilihat dari jenis pertanyaan inferentif yang harus dibuat dibedakan ke dalam pertanyaan inferensi tersurat dan inferensi tersirat, dan tidak menanyakan inferensi tersirat di luar wacana sebagaimana yang dilakukan Pearson dan Johson (via Anderson, 1991).

Tes *cloze* dimaksudkan untuk menyadap kemampuan peserta memahami wacana, maka penghilangan kata-kata sengaja ditujukan terhadap kata-kata kunci dan bukan kata tugas. Berbeda halnya dengan tes *cloze* dalam bahasa Inggris yang penuh dengan kendala bahasa, misalnya masalah kala dan jumlah, dalam bahasa Indonesia hal itu tidak dijumpai. Namun, hal itu tidak berarti tak ada kendala bahasa, atau paling tidak tempat kosong yang jawabannya harus mempertimbangkan kalimat dan bentuk kata yang lain. Misalnya, adanya tuntutan bentuk paralelistis, sebelum kata "oleh" adalah kata kerja bentuk pasif, sesudah kata kerja aktif transitif adalah kata benda, dan lain-lain. Dalam tes *cloze* yang dikembangkan ini terdapat 10 buah tempat kosong yang isian mengandung kendala bahasa, sedang 20 yang lain berupa isian kata-kata kunci yang mendukung gagasan wacana.

2. Prosedur Pengukuran

Pelaksanaan pengetesan kemampuan membaca pemahaman untuk tes bentuk pilihan ganda dan *cloze* dilakukan bersamaan. Kedua perangkat tes dijadikan satu kesatuan. Setelah peserta mengerjakan tes pilihan ganda kemudian diikuti pengerjaan tes *cloze* masing-masing selama 30 menit.

Hasil pengukuran dianalisis dengan teknik statistik deskriptif, t-tes, dan korelasi. Teknik analisis yang pertama dipergunakan untuk menghitung nilai rata-rata dan simpangan baku skor tes pilihan ganda dan *cloze*. Hasil penghitungan ini dipergunakan untuk menghitung indeks reliabilitas K-R 21 di samping juga untuk mempertimbangkan kemampuan membaca pemahaman mahasiswa baik secara individual maupun secara kelompok. Teknik t-tes dipergunakan untuk menghitung perbedaan jawaban pertanyaan inferentif yang tersurat dan tersirat. Teknik korelasi dipakai untuk mengetahui kesejajaran antara tes pilihan ganda dengan tes *cloze*, kesejajaran antara ketiga aspek isi

pertanyaan yang berupa tema, informasi, dan hubungan antaralinea, dan kesejajaran antara pertanyaan inferentif tersurat dan tersirat. Penghitungan indeks reliabilitas K-R 21 dihitung secara manual dengan kalkulator, sedang penghitungan statistik yang lain dilakukan dengan program SPSSPC.

Selain itu, juga dilakukan kerja analisis butir soal yang dimaksudkan untuk menghitung indeks tingkat kesulitan dan daya beda tiap butir soal yang diteskan. Keduanya dimaksudkan untuk menentukan kelayakan butir-butir tes, yaitu apakah dipertahankan, direvisi, atau diganti. Suatu butir tes dinyatakan layak jika indeks tingkat kesulitannya berkisar antara 0,15 sampai 0,85 dan daya bedanya paling tidak mencapai 0,25 (Oller, 1979: 247). Anderson dkk (1991:50) mengatakan bahwa indeks tingkat kesulitan yang tergolong mudah adalah $p > 0,67$, tingkat sedang $p = 0,33—0,66$; dan tingkat sulit $P < 0,33$. Jika digabungkan dengan kriteria Oller, kriteria itu menjadi: butir yang tergolong mudah berkisar antara 0,68—0,85, sedang: 0,33—0,67, dan sulit: 0,15—0,32. Rara-rata tingkat kesulitan butir tes *cloze* dihitung berdasarkan rata-rata hitungnya.

3. Hasil Penelitian

Hasil pengukuran dapat dibedakan ke dalam dua kelompok, yaitu (1) hasil tes dan kaitannya dengan penampilan peserta tes, dan (2) hasil tes dan kaitannya dengan alat tes itu sendiri. Namun, sesuai dengan tujuan penulisan ini pembicaraan lebih ditujukan kepada masalah yang kedua, sedang pertama hanya terlihat secara tidak langsung..

Data hasil pengukuran menunjukkan bahwa kemampuan membaca pemahaman mahasiswa yang dijadikan sampel uji coba ini belum tinggi. Rata-rata hitung yang hanya sebesar 14,939 (untuk pilihan ganda) dan 16,00 (untuk *cloze*) dari kemungkinan skor tertinggi 30 menunjukkan betapa masih kurangnya kemampuan membaca mereka. Perbedaan yang relatif kecil antara nilai rata-rata tes bentuk pilihan ganda dengan tes *cloze* mengindikasikan bahwa keduanya tidak berbeda kadar kesulitannya. Kemungkinan hal itu juga disebabkan oleh kenyataan bahwa kedua tes itu diambil dari teks yang sama sehingga dapat diperkirakan berkadar keterbacaan yang setingkat pula. Untuk pengujian masalah ini lebih lanjut akan dilakukan pada pembicaraan tentang kesulitan butir soal.a. Kesejajaran Tes Pilihan Ganda dan Tes *Cloze*.

Tes *cloze* yang dicobakan dalam penelitian ini dimaksudkan sebagai pendamping atau pelengkap tes pilihan ganda di samping untuk memperoleh informasi tentang kesejajaran keduanya. Untuk mengetahui kesejajaran itu dilakukan uji statistik dengan teknik korelasi *product moment*. Hasil korelasi antara tes pilihan ganda dengan tes *cloze* untuk membaca pemahaman adalah 0,7047 dengan $p = 0,000$. Jadi, terdapat korelasi yang signifikan antara kedua bentuk tes itu.

Hal itu berarti dapat ditafsirkan bahwa terdapat kesejajaran antara tes membaca pemahaman bentuk pilihan ganda dengan tes *cloze*. Implikasi dari keadaan itu adalah kedua tes tersebut dapat dipakai untuk saling menggantikan untuk mengukur kemampuan membaca pemahaman peserta tes. Artinya, untuk tujuan menyadap kemampuan membaca pemahaman peserta tes dapat dipergunakan salah satu, atau keduanya, dari kedua bentuk tes tersebut. Hasilnya akan memberikan informasi yang kurang lebih sama.

4. Hasil Tes sebagai Pengembangan Alat Tes

a. Reliabilitas Alat Tes

Reliabilitas alat tes, baik yang berbentuk pilihan ganda maupun tes *cloze* dihitung dengan rumus K-R 21. Hasil penghitungan menunjukkan bahwa indeks reliabilitas tes pilihan ganda adalah sebesar 0,6033251, atau dibulatkan menjadi 0,60, sedang untuk tes *cloze* adalah sebesar 0,586774 atau dibulatkan menjadi 0,59. Tuckman (1975:256) mengemukakan bahwa untuk tes buatan guru dinyatakan reliabel jika indeks reliabilitas alat tes itu mencapai 0,60. Berdasarkan ketentuan itu dapat dikatakan bahwa tes membaca pemahaman bentuk pilihan ganda yang diujikan ini reliabel. Sebaliknya, indeks reliabilitas tes *cloze* tidak reliabel. Namun, berhubung besarnya indeks reliabilitas itu hanya sedikit di bawah 0,60 (hanya berbeda 0,02), paling tidak alat tes tersebut dapat dikatakan mendekati reliabel. Selain itu, sebagaimana terlihat di atas bahwa terdapat korelasi yang signifikan antara kedua alat tes itu yang mengindikasikan adanya kesejajaran, hal itu sekaligus juga dapat ditafsirkan bahwa reliabilitas keduanya tidak banyak berbeda.

Akan tetapi, jika pemertimbangan reliabilitas tersebut lewat konsultasi tabel nilai-nilai kritis *product moment* —semua penghitungan koefisien r pada hakikatnya dapat diinterpretasikan dengan nilai tabel *product moment*—

koefisien r yang diperoleh sangat signifikan, karena melewati batas signifikansi 1% yang "hanya" menuntut besarnya $r = 0,403$. Berdasarkan keadaan itu dapat dinyatakan bahwa reliabilitas kedua bentuk tes tersebut reliabel dengan kadar reliabilitas yang cukup tinggi.

Jika kadar reliabilitas alat tes dilihat dari kriteria tradisional sebagaimana yang dikemukakan Tuckman, secara keseluruhan reliabilitas kedua bentuk tersebut kurang tinggi. Paling tidak terdapat dua kemungkinan yang menyebabkannya. Pertama, jumlah butir tes instrumen tes itu sendiri kurang. Seandainya keduanya ditingkatkan jumlahnya, misalnya masing-masing menjadi 50 butir soal, reliabilitas keduanya pasti akan lebih tinggi. Kedua, keadaan mahasiswa peserta tes itu homogen. Artinya, kemampuan mereka tidak terlalu banyak berbeda sehingga rentangan skornya tidak terlalu besar dan simpangan bakunya pun relatif kecil. Indeks variabilitas kelas untuk tes pilihan ganda adalah 28,39 dan untuk tes *cloze* 25,66. Hal itu menunjukkan bahwa keadaan kelas yang dijadikan sampel penelitian ini homogen (indeks variabilitas 29 ke bawah: homogen; 30–39: normal; 40 ke atas heterogen) (Nurgiyantoro, 1995: 410).

b. Triangulasi Validitas Konstruk

Sebagaimana dikemukakan pada bagian depan validitas konstruk instrumen tes dipertanggungjawabkan secara teoretis. Untuk pengembangan tes membaca pemahaman ini, validitas tes dipertanggungjawabkan dari segi isi tes yang tertuang dalam isi pertanyaan.

Dengan memodifikasi model tes yang dikembangkan Anderson dkk (1991), isi tes membaca pemahaman dibedakan ke dalam tiga kategori, yaitu (1) tema yang mencakup gagasan pokok alinea, tema seluruh wacana, dan judul wacana, (2) informasi yang terkandung dalam wacana, dan (3) ide yang menghubungkan antaralinea. Korelasi antara ketiga kategori tersebut dapat dilihat pada tabel berikut.

Tabel Korelasi Antaraspek Isi Tes Membaca Pemahaman

No.	Aspek Isi Tes	Tema	Informasi	Hubungan
1.	Tema	1,00	0,7896 P=0.000	0,4345 P=0,012
2.	Informasi		1,000	0,492 P=0,004
3.	Hubungan			1,000

Korelasi antaraspek isi tes membaca pemahaman di atas kesemuanya signifikan. Hal itu memperkuat validitas konstruk tes membaca pemahaman yang secara teoretis sebagaimana dikemukakan sebelumnya lewat triangulasi dari sumber data hasil tes.

Sebagaimana dikemukakan di atas Pearson & Johnston (Anderson dkk, 1991) membedakan aspek tes membaca pemahaman ke dalam tiga kategori, yaitu pertanyaan yang jawabnya ada tersurat dalam teks, hanya tersirat, dan tersirat di luar teks. Dalam penelitian ini hanya diambil dua yang pertama. Hasil korelasi jawaban peserta tes terhadap kedua aspek tersebut adalah $r = 0,7100$, $P = 0,000$. Hal itu juga memperkuat validitas konstruk tes membaca pemahaman lewat triangulasi dari sumber data hasil tes.

Jadi, berdasarkan pertanggungjawaban teoretis dan triangulasi data hasil tes, tes bentuk pilihan ganda memiliki validitas konstruk yang dapat dipertanggungjawabkan. Oleh karena terdapat korelasi yang signifikan antara tes pilihan ganda dengan tes *cloze*, hal itu juga dapat diartikan bahwa tes *cloze* memiliki validitas konstruk yang baik pula.

c. Tingkat Kesulitan dan Daya Beda Butir Tes

Rata-rata indeks tingkat kesulitan dan daya beda tes pilihan ganda masing-masing adalah 0,48 dan 0,39. Berdasarkan kriteria pembagian Tuckman (Anderson dkk, 1991) sebagaimana dirujuk di atas, hal itu berarti bahwa tingkat kesulitan butir soal instrumen tes itu berada dalam kategori sedang ($0,33 > ITK < 0,67$). Indeks daya beda pun secara rata-rata tergolong baik karena mampu membedakan antara kelompok atas dan bawah. Rata-rata tingkat kesulitan tes *cloze* adalah 0,485. Dengan demikian, tingkat kesulitan

dua bentuk tes itu seimbang. Hal itu sebenarnya dapat diperkirakan dari besarnya nilai rata-rata yang tidak banyak berbeda di atas, di samping juga didukung oleh kenyataan bahwa bahan kedua bentuk tes itu diambil dari artikel wacana yang sama. Berdasarkan analisis tingkat kesulitan dan daya beda butir soal tersebut dapat pula “kekuatan” masing-masing butir tes sehingga dapat ditentukan kelayakan butir tes yang bersangkutan untuk dipertahankan, direvisi, atau diganti.

E. Penutup

Sebagaimana yang dikemukakan di atas lewat kajian teori dan kemudian diperkuat lewat penelitian pengembangan alat tes tes membaca pemahaman, terdapat kesejajaran antara bentuk tes pilihan ganda dengan tes *cloze*. Hal itu menyuratkan suatu pengertian bahwa kedua bentuk tes tersebut sama-sama dapat dipakai untuk menyadap kemampuan membaca pemahaman. Keduanya dapat dipakai secara bersama, saling melengkapi dan saling menggantikan, dan akan memberikan hasil yang kurang lebih sama.

Oleh karena itu, untuk tujuan penyadapan kemampuan membaca pemahaman siswa atau mahasiswa, disarankan penggunaan keduanya, baik yang bersifat melengkapi dalam arti keduanya dipakai sekaligus, maupun saling menggantikan dalam arti dipakai salah satu. Jika teknik *cloze* dinilai tidak layak pakai dalam tes-tes seperti Ebtanas, UMPT, atau bahkan ujian akhir semester –penggunaan bentuk tertentu melibatkan banyak faktor penyebab– paling tidak bentuk tes itu dapat dipakai di kelas dalam tes-tes formatif. Penggunaan alat ukur yang bervariasi akan memberikan dampak psikologis yang menguntungkan, misalnya untuk mengurangi kejenuhan baik bagi guru (dosen) maupun siswa (mahasiswa). Selain itu, penyusunan soal pilihan ganda yang membutuhkan waktu lama, juga sedikit dapat diatasi dengan teknik *cloze* yang relatif singkat pembuatannya. Atau paling tidak, adanya kesadaran bahwa terdapat alat ukur lain yang sama-sama baik yang dapat dipergunakan untuk menyadap kemampuan membaca pemahaman, sudah merupakan hal yang positif. Dalam tes bahasa Inggris untuk penutur asli dan asing, teknik *cloze* pun diketahui sejajar dengan keterampilan-keterampilan berbahasa yang lain. Untuk memastikan apakah hal itu juga berlaku untuk tes bahasa Indonesia bagi penutur asli, kiranya perlu juga dilakukan tes pengembangan yang sejenis.

DAFTAR PUSTAKA

- Anderson, Neil J. dkk. 1991. "An Exploratory Study into the Construct Validity of a Reading Comprehension Test: Triangulation of Data Sources", dalam *Language Testing*, 8, 1, hlm. 41--65.
- Blaricom, Ginger van. 1981. "The Effect of Passage Organization on Main Idea Comprehension at Three Response Levels", dalam Clifford Pennock (ed), *Reading Comprehension at Four Linguistics Levels*. Delaware: International Reading Association.
- Brown, James Dean. 1993. "What are the Characteristics of Natural Cloze Tests?", dalam *Language Testing*, 10, 2, hlm. 93-115.
- Djiwandoro, Soenardi. 1989/1990. *Pengembangan Tes Kemampuan berbahasa Indonesia*, Malang: Laporan Penelitian DPPM (Tidak terbit).
- Jonz, Jon. 1991. "Cloze Item Types and Second Language Comprehension", dalam *Language Testing*, 8, 1, hlm. 1-22.
- Nurgiyantoro, Burhan. 1995. *Penilaian dalam Pengajaran Bahasa dan Sastra*. Yogyakarta: BPFE.
- Oller, John W. 1979. *Language Test at School, A Pragmatic Approach*. London: Longman.
- Perkins, Kyle. 1992. "The Effect of Passage Topical Structure Types on ESL Reading Comprehension Difficulty", dalam *Language Testing*, 9, 2, hlm. 163-72.
- Tuckman, Bruce W. 1975. *Measuring Educational Outcomes, Fundamentals of Testing*, New York: Harcourt Brace Jovanovich.