

**THE VALIDITY OF CLOZE-TEST, TEST-C AND
SELF-ASSESSMENT AS A MEASUREMENT OF THE ABILITY
TO READ ENGLISH TEXTS**

Subarso

Abstract

This study aimed at measuring the validity of Cloze-test (CL) and test-C (C) as a measurement of the ability to read English texts using the Reading Comprehension (RC) test of the TOEFL and the validity of the Self-Assessment on the three tests as the references. Research population consists of students attending the Reading I, III and V enrolled in the first semester of 2006/2007 in the English Education department, Faculty of Languages and Arts, Yogyakarta State University. Sampling was done using the cluster random sampling (N = 121). Data about the reading comprehension skill were collected using the RC, CL and C-tests. Data on the students' self-assessment were collected using the three Likert-type questionnaires. Data were analyzed using the Pearson's product moment correlation technique. Findings suggest that there is a positive correlation between the RC and CL test scores, between the RC and C test scores and the PD score and the other three reading test scores.

Key words: *reading comprehension test, cloze test, test-C, self-assessment, validity*

**VALIDITAS TES *CLOZE*, TES-C, DAN PENILAIAN DIRI
SEBAGAI ALAT UKUR KEMAMPUAN MEMBACA
TEKS BAHASA INGGRIS**

Subarso

Abstrak

Penelitian ini bertujuan untuk mengetahui validitas tes *cloze* (CL) dan tes-C (C) sebagai alat ukur kemampuan membaca teks bahasa Inggris dengan tes *Reading Comprehension* (RC) dari TOEFL sebagai tolok ukur dan mengetahui validitas penilaian diri (PD) dengan mengacu pada ketiga tes tersebut. Populasi penelitian meliputi mahasiswa yang mengambil mata kuliah *Reading* I, III, dan V pada semester gasal tahun akademik 2006/2007 di Program Studi Pendidikan Bahasa Inggris, Fakultas Bahasa dan Seni, Universitas Negeri Yogyakarta. Sampel penelitian terdiri atas 121 mahasiswa yang dipilih dengan menggunakan teknik sampel acak gugus. Data kemampuan membaca dikumpulkan dengan tes RC, CL dan C. Data PD dikumpulkan dengan tiga kuesioner PD yang menggunakan model Likert. Data penelitian dianalisis dengan menggunakan teknik korelasi *Pearson product moment*. Hasil analisis data menunjukkan bahwa ada korelasi yang positif antara skor tes RC dan skor tes CL, antara skor tes RC dan tes C, dan antara skor PD dan ketiga skor tes membaca tersebut.

Kata kunci: *tes reading comprehension, tes cloze, tes-c, penilaian diri, validitas*

Pendahuluan

Penilaian hasil belajar yang adil, tepat, dapat menggambarkan kemampuan pembelajar yang sebenarnya, dan dapat diterima semua pihak adalah sesuatu yang didambakan dalam proses belajar mengajar. Namun, untuk mewujudkan dambaan tersebut tidak mudah. Di dalam perkuliahan, misalnya, kenyataan menunjukkan bahwa kadang-kadang mahasiswa tidak puas dengan nilai yang diberikan oleh dosen. Ketidakpuasan ini dapat disebabkan oleh berbagai faktor. Ada faktor yang berhubungan dengan alat ukurnya, dan ada faktor yang berhubungan dengan pihak-pihak yang dilibatkan dalam pengambilan keputusan mengenai nilai.

Penilaian yang tidak tepat dan tidak memuaskan dapat berdampak negatif pada proses belajar mengajar. Motivasi belajar mahasiswa dapat menurun atau mungkin juga terjadi tindak permusuhan kepada penilai atau dosen. Sebaliknya, penilaian yang tepat dan memuaskan merupakan pendorong timbulnya motivasi dan sikap positif, sehingga dapat memacu mahasiswa untuk lebih berprestasi.

Salah satu cara untuk meningkatkan ketepatan penilaian dan kepuasan mahasiswa terhadap nilai adalah dengan menggunakan alat ukur yang tepat dan dengan melibatkan mahasiswa dalam proses pengambilan keputusan mengenai nilai. Dengan alat ukur yang tepat, mahasiswa akan merasa puas ketika menerima keputusan yang berkaitan dengan kegiatan pembelajaran yang telah mereka ikuti. Lebih-lebih jika pengambilan keputusan tersebut mempertimbangkan pendapat atau pandangan mahasiswa. Untuk itu alat ukur yang tepat dan pelibatan mahasiswa diperlukan.

Penelitian ini menitikberatkan pada penggunaan alat ukur alternatif untuk mengukur kemampuan membaca teks bahasa Inggris dan pada pelibatan mahasiswa dalam penilaian diri. Alat ukur alternatif yang diteliti adalah tes *cloze* dan tes-C, yang diacukan pada tes *Reading Comprehension* dari TOEFL. Penilaian diri dalam penelitian ini diacukan pada ketiga tes membaca tersebut.

Ada banyak definisi mengenai kemampuan membaca. Definisi yang paling umum adalah bahwa kemampuan membaca adalah kemampuan

seseorang untuk memahami teks tertulis. Namun, definisi yang tampak sederhana ini menyiratkan adanya kompleksitas karena kemampuan membaca itu sendiri merupakan gabungan keterampilan yang perlu dikuasai oleh seseorang ketika dia membaca. Grabe (2002: 51) mendefinisikan membaca sebagai keterampilan yang melibatkan “*rapid and automatic process, interacting process, flexible and strategic process, purposeful process, and linguistic process.*” Penjelasan yang senada dikemukakan oleh Brown (2001: 298-302). Membaca merupakan suatu proses interaktif yang terjadi antara pembaca dan teks, yang menghasilkan pemahaman. Gipe (1991: 5-7) mengidentifikasi empat dimensi dari proses membaca, yaitu proses kebahasaan, proses kognitif, proses psikologis atau afektif, dan proses fisiologis.

Luke (1995: 172) menyampaikan pergeseran pendekatan membaca. Di dalam kelas para pembelajar dianjurkan untuk menciptakan makna, tidak sekedar memahami makna. Dalam kegiatan membaca ada pergeseran dari “*text-driven*” ke “*reader-driven*”, dari “*bottom-up*” ke “*top-down*” ke “*transactional*”, dan dari “*outside-in*” ke “*inside-out*”. Pergeseran-pergeseran yang disebutkan di atas dilatarbelakangi oleh paham konstruktivisme (Tonjes, Wolpow, & Zintz, 1999: 101). Sejalan dengan salah satu ciri paham konstruktivisme, yaitu “*all knowledge is socially constructed*” (Richards & Schmidt, 2002: 114), Weaver (1994) menyatakan bahwa membaca bukan semata-mata proses psikolinguistik yang melibatkan otak pembaca dan bahasa yang digunakan untuk mengemas teks. Membaca juga merupakan proses sosio-psikolinguistik, karena transaksi antara pembaca dengan teks terjadi dalam konteks sosial dan situasional. Celce-Murcia (2001: 154) menyatakan bahwa membaca adalah proses interaktif dan sosiokognitif. Membaca melibatkan teks, pembaca, dan konteks sosial yang menjadi ajang berlangsungnya kegiatan membaca. Membaca dengan demikian merupakan kegiatan yang dilakukan oleh seseorang dalam merekonstruksi makna melalui transaksi dengan teks tulis yang dikemas dengan lambang yang berupa bahasa. Transaksi ini meliputi tindakan pembaca ketika menafsirkan teks, yang dipengaruhi oleh pengalaman, latar belakang kebahasaan, latar belakang budaya, dan tujuan membacanya.

Dari uraian tersebut dapat disimpulkan bahwa setidaknya-tidaknya ada dua pendekatan terhadap kegiatan membaca, yaitu pendekatan psikologis dan pendekatan sosiologis. Pendekatan psikologis menekankan latar belakang pembaca pada segi pengetahuan kognitif dan pengetahuan umum pembaca. Pendekatan sosiologis menekankan latar belakang pembaca pada segi latar belakang budaya, dan kebiasaan dan praktek berwacana dalam penggunaan bahasa di masyarakat.

Menurut Alderson (Phakiti, 2003: 691) kemampuan membaca dapat diukur dengan bermacam-macam cara. Suatu cara tertentu mungkin akan mengukur aspek membaca tertentu. Tidak ada satu cara yang paling baik untuk mengukur kemampuan membaca, karena masing-masing cara memiliki kelebihan dan kekurangannya. Oleh karena itu, masing-masing cara disukai oleh sejumlah orang dan tidak disukai oleh sejumlah orang yang lain (Farr & Carey, 1986: 33).

Alat ukur yang biasa digunakan untuk mengukur kemampuan membaca teks bahasa Inggris adalah tes *Reading Comprehension* dari *Test of English as a Foreign Language* (TOEFL) yang dikembangkan oleh *Educational Testing Service* (ETS). Alat ukur alternatif yang dapat digunakan adalah tes *cloze* dan tes-C. Tes *cloze* adalah jenis tes yang di dalamnya setiap kata ke *n* (bilangan yang dijadikan dasar untuk menghilangkan kata dalam bacaan) dihilangkan dan pembaca diminta untuk mengisi kata yang dihilangkan tersebut (Brown, 2004: 201; Shohamy, 1983: 231). Pada dasarnya tes-C adalah sejenis tes *cloze*. Tes-C juga meminta pembaca untuk melengkapi kata yang tidak sempurna dalam suatu teks. Pada tes *cloze*, pembaca diminta untuk mengisi kata yang dihilangkan, sedangkan pada tes-C pembaca diminta untuk melengkapi kata yang belum sempurna, penghilangan hurufnya mengikuti kaidah 2. Artinya, setiap kata kedua dan berkelipatan dua dibuat tidak lengkap dengan menghilangkan huruf paruh kedua pada setiap kata dan jika jumlah huruf dalam suatu kata ganjil, huruf yang dihilangkan adalah paruh kedua ditambah satu huruf (Brown, 2004: 203).

Penilaian diri merupakan cara untuk memotivasi pembelajar. Dengan menilai usaha mereka sendiri, para pembelajar dapat memperoleh pemahaman terhadap masalah atau pengalaman yang mereka hadapi. Cara ini dapat digunakan untuk mengatasi ketidakpuasan mereka terhadap

penilaian yang dilakukan hanya oleh pengajar, karena persepsi pembelajar terhadap usahanya sendiri tidak selalu sejalan dengan persepsi pengajar terhadap usaha pembelajar. Karena penilaian diri menuntut kejujuran pembelajar, hubungan saling percaya antara pengajar dan pembelajar perlu dibangun sejak awal proses belajar mengajar (Griffin & Nix, 1991: 64). Hal ini sejalan dengan Oskarsson (Rolfe, 1990: 164) yang menyatakan "... *the education of the learner to make reliable and valid autonomous judgements on the effectiveness of his communication is a necessary part of the learning process*".

Untuk mengetahui apakah alat ukur kemampuan membaca alternatif, yaitu tes *cloze* dan tes-C, memiliki validitas, alat ukur kemampuan membaca yang baku dapat dijadikan tolok ukur. Alat ukur baku yang biasa digunakan sebagai tolok ukur salah satunya adalah tes *Reading Comprehension* dari TOEFL. Untuk mengetahui validitas penilaian diri, tolok ukur yang digunakan adalah ketiga tes kemampuan membaca, yaitu tes *cloze*, tes-C, dan tes *Reading Comprehension* dari TOEFL. Dengan mengkorelasikan skor yang diperoleh melalui alat ukur alternatif kemampuan membaca dengan skor yang diperoleh melalui alat ukur baku kemampuan membaca, dan dengan mengkorelasikan skor penilaian diri dengan skor ketiga tes kemampuan membaca, koefisien-koefisien validitas tolok ukur atau *criterion-related validity* (Bachman, 1990: 248; Allen & Yen, 1979: 87) dapat diperoleh.

Validitas tolok ukur adalah jenis validitas yang didasarkan pada sejauh mana suatu alat ukur dapat dibandingkan atau dikorelasikan dengan alat ukur lain yang dijadikan tolok ukur. Untuk konsep ini, Gall, Gall, & Borg (2003: 194) menggunakan istilah *test-criterion evidence*. Mereka lebih lanjut menjelaskan bahwa jika skor suatu tes yang dicapai oleh sekelompok anggota sampel berkorelasi positif dan signifikan dengan skor tes lain yang diasumsikan mengukur kemampuan yang sama, tes-tes tersebut memiliki *convergent evidence*. Validitas tolok ukur dalam penelitian adalah *concurrent validity* (Bachman, 1990: 248), yaitu validitas yang didasarkan pada sejauh mana suatu alat ukur berkorelasi dengan alat ukur lain yang digunakan untuk mengukur kemampuan yang sama atau sejenis. Kedua alat ukur digunakan pada waktu yang bersamaan, atau pada waktu berlainan yang berselang tidak terlalu lama.

Penelitian mengenai tes *cloze* telah banyak dilakukan. Alderson (1983: 205-213) mengadakan penelitian tes *cloze* dalam kaitannya dengan kemahiran bahasa Inggris sebagai bahasa asing. Shohamy (1983: 229-236) meneliti reliabilitas *interrater* dan *intrarater* dari wawancara lisan dalam hubungannya dengan tes *cloze*. Brown (1983: 238) meneliti reliabilitas tes *cloze* dengan berbagai macam pendekatan. Penelitian mengenai tes-C antara lain dilakukan oleh Katona & Dornyei (1993: 35), yang mencoba mengkorelasikannya dengan beberapa tes baku yang biasa digunakan. Ikeguchi (1998: 3-8) meneliti beberapa jenis tes-C dan mengaitkannya dengan tingkat kemahiran bahasa Inggris sebagai bahasa kedua. Penelitian mengenai penilaian diri dilakukan antara lain oleh Rolfe (1990: 163-186), yang mengadakan penelitian tentang penilaian diri dan penilaian sejawat oleh pembelajar terhadap kemampuan berbahasa lisan. Penilaian diri dan penilaian sejawat ini diacukan pada penilaian pengajar. Lewis (1990: 202-203) mengembangkan dan merintis format penilaian diri untuk memantau kemajuan pembelajar secara berkesinambungan. Roever & Powers (2005: 1-18) mengadakan penelitian mengenai hubungan antara penilaian diri dengan skor TOEFL.

Berdasarkan uraian di atas, masalah dalam penelitian ini dirumuskan sebagai berikut:

1. Bagaimanakah validitas tes *cloze* sebagai alat ukur kemampuan membaca jika diacukan pada tes baku kemampuan membaca, yaitu tes *Reading Comprehension* dari TOEFL?
2. Bagaimanakah validitas tes-C sebagai alat ukur kemampuan membaca jika diacukan pada tes baku kemampuan membaca?
3. Bagaimanakah validitas penilaian diri oleh mahasiswa jika diacukan pada tes baku kemampuan membaca?
4. Bagaimanakah validitas penilaian diri oleh mahasiswa jika diacukan pada tes *cloze*?
5. Bagaimanakah validitas penilaian diri oleh mahasiswa jika diacukan pada tes-C?

Penelitian ini bertujuan untuk mengetahui validitas tolok ukur (*criterion-related validity*) tes *cloze*, dan tes-C, sebagai alat ukur kemampuan membaca dengan diacukan pada tolok ukur tes baku kemampuan

membaca, yaitu tes *Reading Comprehension* dari TOEFL. Di samping itu, penelitian ini juga bertujuan mengetahui validitas tolok ukur penilaian diri dengan acuan ketiga tes tersebut.

Metode Penelitian

Penelitian ini termasuk jenis penelitian pengembangan yang bersifat penjangjagan. Penelitian ini dirancang dengan menggunakan model kuantitatif yang menggunakan teknik korelasional dengan cara menghubungkan skor tes *cloze*, tes-C, dengan skor tes baku kemampuan membaca yang akan dijadikan tolok ukur, yaitu tes *Reading Comprehension* dari TOEFL. Di samping itu, skor masing-masing dari ketiga tes tersebut dikorelasikan dengan skor penilaian diri untuk masing-masing tes.

Populasi penelitian ini adalah mahasiswa Program Studi Pendidikan Bahasa Inggris yang mengambil mata kuliah Reading I, Reading III, dan Reading V pada semester gasal tahun akademik 2006/2007. Sebanyak 121 mahasiswa dipilih sebagai sampel dengan menggunakan teknik sampel gugus.

Penelitian ini menggunakan enam macam instrumen yaitu tes *cloze* dan tes-C, yang masing-masing dikembangkan oleh peneliti, dan tes *Reading Comprehension* dari TOEFL, serta tiga lembar penilaian diri yang dikembangkan oleh peneliti, yang masing-masing untuk tes *cloze*, tes-C, dan tes *Reading Comprehension* dari TOEFL.

Instrumen yang sudah diujicobakan kemudian diberikan kepada anggota sampel penelitian untuk dikerjakan dalam dua periode waktu yang terpisah. Pada periode yang pertama tes *Reading Comprehension* dari TOEFL diberikan kepada mahasiswa. Setelah selesai mengerjakan tes tersebut, mahasiswa diminta untuk memberikan penilaian terhadap kemampuan mereka sendiri dalam mengerjakan tes yang baru saja mereka selesaikan. Penilaian tersebut mereka tuangkan dalam lembar penilaian diri.

Pada periode yang kedua, yaitu seminggu setelah tes *Reading Comprehension* dari TOEFL, tes *cloze* dan tes-C diberikan kepada mahasiswa. Setelah mengerjakan tes *cloze*, mahasiswa diminta untuk memberikan penilaian terhadap kemampuan mereka sendiri dalam mengerjakan tes

tersebut. Penilaian tersebut mereka tuangkan dalam lembar penilaian diri. Kemudian mahasiswa mengerjakan tes-C. Setelah selesai mengerjakannya, mahasiswa diminta untuk menilai kemampuan mereka sendiri dalam mengerjakan tes tersebut. Penilaian tersebut mereka tuangkan dalam lembar penilaian diri.

Data penelitian dianalisis secara deskriptif untuk mengetahui skor yang diperoleh mahasiswa untuk masing-masing tes dan penilaian diri terhadap masing-masing tes tersebut. Statistik deskriptif yang digunakan adalah rerata dan simpang baku. Untuk mengetahui validitas tolok ukur alat-alat ukur tersebut, teknik korelasi Pearson *product moment* digunakan, dengan cara mengkorelasikan skor tes *cloze* dan tes-C dengan skor tes *Reading Comprehension* dari TOEFL. Skor masing-masing penilaian diri juga dikorelasikan dengan skor masing-masing tes dengan teknik korelasi Pearson *product moment*. Statistik deskriptif dan penghitungan koefisien korelasi menggunakan program SPSS 12.

Hasil Penelitian dan Pembahasan

Untuk menjawab permasalahan bagaimanakah validitas tes *cloze* sebagai alat ukur kemampuan membaca jika diacukan pada tes baku kemampuan membaca, yaitu tes *Reading Comprehension* dari TOEFL, bagaimanakah validitas tes-C sebagai alat ukur kemampuan membaca jika diacukan pada tes baku kemampuan membaca, dan bagaimanakah validitas penilaian diri pembelajar jika diacukan pada ketiga tes tersebut, uji korelasi antar tes perlu dilakukan. Teknik korelasi yang sesuai dengan data penelitian adalah teknik korelasi Pearson *product moment*. Tabel 1 menunjukkan hasil analisis terhadap hubungan antar skor tes yang dikemukakan dalam pertanyaan dari data yang diambil dari semua mahasiswa yang menjadi anggota sampel penelitian ini tanpa mempertimbangkan semester.

Tabel 1. Hasil Uji Korelasi Pearson *Product Moment* untuk Skor Tes *Reading Comprehension* dari TOEFL, Tes *Cloze*, Tes-C, Penilaian Diri TOEFL, Penilaian Diri Tes *Cloze*, dan Penilaian Diri Tes-C Keseluruhan

Skor	<i>r</i>
TOEFL dan Tes <i>Cloze</i>	0,54
TOEFL dan Tes-C	0,54
TOEFL dan Penilaian Diri	0,65
Tes <i>Cloze</i> dan Penilaian Diri	0,76
Tes-C dan Penilaian Diri	0,65

Hasil analisis menunjukkan bahwa semua nilai *r* positif. Dengan demikian pertanyaan penelitian di atas dapat dijawab. Tes *cloze* dan tes-C dapat dijadikan sebagai instrumen alternatif untuk mengukur kemampuan membaca. Hasil analisis di atas diperoleh dari data semua mahasiswa yang menjadi anggota sampel penelitian. Dengan kata lain, semester tidak dipertimbangkan dalam analisis tersebut. Ini menyiratkan bahwa kedua tes tersebut dapat digunakan untuk pembelajar dari tingkat manapun.

Nilai *r* yang positif untuk ketiga penilaian diri di atas menunjukkan bahwa mahasiswa dapat menakar kemampuan dirinya. Dengan demikian, ketika pengajar hendak memutuskan nilai untuk pembelajar, pandangan pembelajar perlu dipertimbangkan.

Tabel 2 menunjukkan hasil analisis terhadap hubungan antar skor tes yang dikemukakan dalam pertanyaan dari data yang diambil dari mahasiswa semester 1 yang menjadi anggota sampel penelitian ini.

Tabel 2. Hasil Uji Korelasi Pearson *Product Moment* untuk Skor Tes *Reading Comprehension* dari TOEFL, Tes *Cloze*, Tes-C, Penilaian Diri TOEFL, Penilaian Diri Tes *Cloze*, dan Penilaian Diri Tes-C Semester 1

Skor	<i>r</i>
TOEFL dan Tes <i>Cloze</i>	0,37
TOEFL dan Tes-C	0,39
TOEFL dan Penilaian Diri	0,39
Tes <i>Cloze</i> dan Penilaian Diri	0,85
Tes-C dan Penilaian Diri	0,46

Hasil analisis menunjukkan bahwa semua nilai r positif. Tes *cloze* dan tes-C dapat dijadikan sebagai instrumen alternatif untuk mengukur kemampuan membaca. Hasil analisis di atas diperoleh dari data mahasiswa semester 1 yang menjadi anggota sampel penelitian. Semester dipertimbangkan dalam analisis tersebut. Ini menyiratkan bahwa kedua tes tersebut dapat digunakan untuk pembelajar tingkat dasar, dengan asumsi bahwa mahasiswa semester 1 baru memiliki kemampuan tingkat dasar.

Nilai r yang positif untuk ketiga penilaian diri di atas menunjukkan bahwa mahasiswa dapat menilai kemampuan dirinya.

Tabel 3. Hasil Uji Korelasi Pearson *Product Moment* untuk Skor Tes *Reading Comprehension* dari TOEFL, Tes *Cloze*, Tes-C, penilaian Diri TOEFL, Penilaian Diri Tes *Cloze*, dan Penilaian Diri Tes-C Semester 3

Skor	r
TOEFL dan Tes <i>Cloze</i>	0,64
TOEFL dan Tes-C	0,43
TOEFL dan Penilaian Diri	0,65
Tes <i>Cloze</i> dan Penilaian Diri	0,55
Tes-C dan Penilaian Diri	0,55

Tabel 3 menunjukkan hasil analisis terhadap hubungan antar skor tes yang dikemukakan dalam pertanyaan dari data yang diambil dari mahasiswa semester 3 yang menjadi anggota sampel penelitian ini. Hasil analisis menunjukkan bahwa semua nilai r positif. Tes *cloze* dan tes-C dapat dijadikan sebagai instrumen alternatif untuk mengukur kemampuan membaca. Hasil analisis pada Tabel 3 diperoleh dari data mahasiswa semester 3 yang menjadi anggota sampel penelitian. Semester dipertimbangkan dalam analisis tersebut. Ini menyiratkan bahwa kedua tes tersebut dapat digunakan untuk pembelajar tingkat menengah, dengan asumsi bahwa mahasiswa semester 3 memiliki kemampuan tingkat menengah. Nilai r yang positif untuk ketiga penilaian diri di atas menunjukkan bahwa mahasiswa dapat menilai kemampuan dirinya.

Tabel 4 menunjukkan hasil analisis terhadap hubungan antar skor tes yang dikemukakan dalam pertanyaan dari data yang diambil dari mahasiswa semester 5 yang menjadi anggota sampel penelitian ini.

Tabel 4. Hasil Uji Korelasi Pearson *Product Moment* untuk Skor Tes *Reading Comprehension* dari TOEFL, Tes *Cloze*, Tes-C, penilaian Diri TOEFL, Penilaian Diri Tes *Cloze*, dan Penilaian Diri Tes-C Semester 5

Skor	<i>r</i>
TOEFL dan Tes <i>Cloze</i>	0,77
TOEFL dan Tes-C	0,74
TOEFL dan Penilaian Diri	0,80
Tes <i>Cloze</i> dan Penilaian Diri	0,81
Tes-C dan Penilaian Diri	0,73

Hasil analisis menunjukkan bahwa semua nilai *r* positif. Tes *cloze* dan tes-C dapat dijadikan sebagai instrumen alternatif untuk mengukur kemampuan membaca. Hasil analisis di atas diperoleh dari data mahasiswa semester 5 yang menjadi anggota sampel penelitian. Semester dipertimbangkan dalam analisis tersebut. Ini menyiratkan bahwa kedua tes tersebut dapat digunakan untuk pembelajar tingkat lanjut, dengan asumsi bahwa mahasiswa semester 5 memiliki kemampuan tingkat lanjut. Nilai *r* yang positif untuk ketiga penilaian diri di atas menunjukkan bahwa mahasiswa dapat menilai kemampuan dirinya.

Hasil analisis korelasi Pearson *product moment* untuk data keseluruhan, tanpa mempedulikan semester, menunjukkan bahwa semua koefisien korelasi positif. Koefisien tersebut menunjukkan bahwa tes *cloze* dan tes-C memiliki validitas tolok ukur dan dapat digunakan sebagai alat ukur alternatif kemampuan membaca pada tingkat manapun. Penilaian diri juga dapat dilakukan pada tingkat manapun. Hal penting yang perlu diperhatikan adalah ada kejelasan petunjuk dalam pelaksanaan penilaian diri tersebut.

Jika dibandingkan antar semester, koefisien korelasi tersebut makin tinggi pada semester yang lebih tinggi, kecuali korelasi antara skor tes *cloze* dengan skor penilaian diri terhadap tes tersebut. Korelasi tertinggi justru

terdapat pada kelompok mahasiswa semester 1 dengan koefisien 0,85, sementara pada semester 3 dengan koefisien 0,55, dan pada semester 5 dengan koefisien 0,81. Keadaan ini menggambarkan apa yang dinyatakan oleh Cohen (1994: 199) bahwa penilaian diri kadang-kadang subjektif dan tidak ajeg karena dipengaruhi oleh berbagai faktor seperti prestasi di masa lalu, keinginan, harapan, perbandingan dengan yang lain, dan sebagainya. Dengan kata lain, makin tinggi semester seorang mahasiswa tidak selalu berarti makin mampu dia menilai dirinya sendiri.

Untuk mengetahui apakah koefisien korelasi tes *cloze* dan tes-C dengan tes *Reading Comprehension* dari TOEFL sebagai tolok ukur dan koefisien korelasi penilaian diri dengan ketiga tes tersebut secara statistik berbeda berdasarkan semester, *critical ratio* dari koefisien-koefisien korelasi tersebut dapat digunakan (Furlong, Lovelace, & Lovelace, 2000: 2006-2007). *Critical ratio* dapat dicari dengan mengubah koefisien korelasi menjadi nilai z. Hasil yang diperoleh dibandingkan dengan nilai kritis z, yaitu $\pm 1,96$. Jika *critical ratio* lebih besar atau sama dengan nilai kritis z, koefisien-koefisien korelasi tersebut secara statistik berbeda. Tabel 5 menunjukkan *critical ratio* korelasi antar tes kemampuan membaca dan korelasi penilaian diri dengan ketiga tes tersebut.

Tabel 5. *Critical Ratio* Korelasi antar Tes Kemampuan Membaca dan Korelasi Penilaian Diri dengan Ketiga Tes Kemampuan Membaca

Skor	Semester 1 dan 3	Semester 3 dan 5	Semester 1 dan 5
	CR	CR	CR
TOEFL dan Tes <i>Cloze</i>	1,50	1,06	2,60*
TOEFL dan Tes-C	0,19	1,99*	2,22*
TOEFL dan Penilaian Diri	1,47	1,31	2,83*
Tes <i>Cloze</i> dan Penilaian Diri	1,47	2,06*	3,59*
Tes-C dan Penilaian Diri	0,49	1,26	1,77

* CR > 1,96

Secara statistik, tidak ada perbedaan kemampuan penilaian diri antara mahasiswa semester 1 dengan mahasiswa semester 3. Ini ditunjukkan oleh

perbandingan *critical ratio* dari koefisien korelasi antara skor yang diperoleh oleh mahasiswa semester 1 dan 3. Tidak ada *critical ratio* yang lebih besar atau sama dengan $\pm 1,96$. Dengan kata lain, koefisien korelasi yang tidak sama secara statistik tidak berbeda.

Pada mahasiswa semester 3 dan 5, kemampuan penilaian diri yang berbeda adalah kemampuan dalam menilai tes *cloze*. *Critical ratio* untuk koefisien korelasi ini adalah 2,06, yang lebih besar daripada $\pm 1,96$. Ini menunjukkan bahwa mahasiswa semester 5 lebih memiliki kemampuan menilai dirinya sendiri dalam tes *cloze*.

Perbandingan antara mahasiswa semester 1 dan 5 menunjukkan bahwa kedua kelompok tersebut memiliki kemampuan yang berbeda dalam menilai diri dalam tes *Reading Comprehension* dan tes *cloze*. Ini ditunjukkan oleh *critical ratio* untuk koefisien korelasi skor penilaian diri dengan skor kedua tes tersebut. *Critical ratio* untuk kedua koefisien korelasi tersebut lebih besar daripada $\pm 1,96$. Dengan kata lain, peningkatan kemampuan penilaian diri terjadi secara bertahap dan tidak selalu signifikan dari segi statistik. Peningkatan akan signifikan setelah rentang waktu yang relatif lama, yaitu dari semester 1 ke semester 5, sementara dari semester 1 ke semester 3 dan dari semester 3 ke semester 5 tidak ada peningkatan yang signifikan kecuali kemampuan penilaian diri dalam tes *cloze*.

Simpulan

1. Tes *cloze* memiliki validitas tolok ukur karena skor tes tersebut memiliki korelasi yang positif dengan skor tes baku kemampuan membaca, yaitu tes *Reading Comprehension* dari TOEFL.
2. Tes-C memiliki validitas tolok ukur karena skor tes tersebut memiliki korelasi yang positif dengan skor tes baku kemampuan membaca.
3. Penilaian diri untuk tes baku kemampuan membaca memiliki validitas tolok ukur karena skor penilaian diri memiliki korelasi yang positif dengan skor tes baku kemampuan membaca.
4. Penilaian diri untuk tes *cloze* memiliki validitas tolok ukur karena skor penilaian diri memiliki korelasi yang positif dengan skor tes *cloze*.

5. Penilaian diri untuk tes-C memiliki validitas tolok ukur karena skor penilaian diri memiliki korelasi yang positif dengan skor tes-C.

Saran

Para pengajar membaca di kelas bahasa Inggris hendaknya berusaha menggali kemungkinan menggunakan alat ukur yang beragam. Alat ukur kemampuan membaca bukan hanya tes *Reading Comprehension* dari TOEFL atau yang sejenis. Tes seperti tes *cloze* dan tes-C beserta variasinya dapat digunakan, karena tes model tersebut mudah disusun, digunakan, dan diskor. Namun, dalam penyusunan tes perlu ada kecermatan dalam pemilihan teks. Teks yang dijadikan tes hendaknya memiliki tingkat kesulitan yang sedang.

Para pengajar hendaknya melibatkan pembelajar ketika hendak memutuskan nilai untuk mereka karena mereka sebenarnya dapat menilai diri mereka sendiri. Pengajar dapat memberikan pengantar sebelum tes dilaksanakan, bahwa setelah tes pembelajar diminta untuk menilai kemampuan dirinya dengan instrumen yang disediakan dan diminta mengisinya dengan sejujur-jujurnya. Jika pembelajar jujur dalam penilaian diri, skor penilaian diri akan bermanfaat dalam pengambilan keputusan mengenai nilai akhir.

Daftar Pustaka

- Alderson, J. C. (1983). The cloze procedure and proficiency in English as a foreign language. Dalam Oller, J. W., Jr. (Ed.) *Issues in language testing research*. Rowley: Newbury House Publishers, Inc.
- Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press
- Brown, H. D. (2001). *Teaching by principles (2nd ed.)*. New York: Addison Wesley Longman, Inc.
- Brown, H. D. (2004). *Language assessment*. New York: Addison Wesley Longman, Inc.

- Brown, J. D. (1983). A closer look at cloze: validity and reliability. Dalam Oller, J. W., Jr. (Ed.) *Issues in language testing research*. Rowley: Newbury House Publishers, Inc.
- Celce-Murcia, M. (2001). *Teaching English as a second or foreign language (3rd ed.)*. Boston: Heinle and Heinle Thomson Learning, Inc.
- Cohen, A. D. (1994). *Assessing language ability in the classroom (2nd ed.)*. Boston: Heinle & Heinle Publishers
- Farr, R. & Carey, R. F. (1986). *Reading: what can be measured? (2nd ed.)* Newark: International Reading Association, Inc.
- Furlong, N. E., Lovelace, E. A., & Lovelace, K. L. (2000). *Research methods and statistics*. Fort Worth: Harcourt College Publishers
- Gall, M. D., Gall, J. P., & Borg, W. R. (2003). *Educational research: an introduction (7th ed.)*. Boston: Pearson Education, Inc.
- Gipe, J. P. (1991). *Corrective reading techniques for the classroom teacher (2nd ed.)*. Scottsdale: Gorsuch Scarisbrick, Publishers
- Grabe, W. (2002). Reading in a second language. Dalam Kaplan, R. B. (Ed.) *Applied linguistics*. Oxford: Oxford University Press
- Griffin, P. & Nix, P. (1991). *Educational assessment and reporting*. Sydney: Harcourt Jovanovich, Publishers.
- Ikeguchi, C. B. (1998). Do different C-tests discriminate proficiency of EL2 learners? *JALT Testing & Evaluation SIG Newsletter* [Versi elektronik], Vol. 2 No. 1 Oktober, 3 – 8
- Katona, L. & Dornyei, Z. (1993). The C-test: a teacher friendly way to test language proficiency. *Forum* [Versi elektronik], Vol. 31 No. 2 Januari, 35
- Lewis, J. (1990). Self-assessment in the classroom: a case study. Dalam Brindley, G. (Ed.) *The second language curriculum in action*. Sydney: National Centre for Language Teaching and Research, Macquarie University

- Luke, A. (1995). The social practice of reading. *Proceedings of the 21st Australian Reading Association*, National Conference, Sydney, 167-187
- Phakiti, A. (2003). A closer look at gender and strategy use in L2 reading. *Language Learning*, 53:4, Desember 2003, 649 – 702
- Richards, J. C. & Schmidt, R. (2002). *Dictionary of language teaching & applied linguistics (3rd ed.)*. London: Longman Pearson Education
- Roever, C. & Powers, D. E. (2005). *Effects of language of administration on a self-assessment of language skills*. Princeton: ETS
- Rolfe, T. (1990). Self- and peer- assessment in the ESL curriculum. Dalam Brindley, G. (Ed.) *The second language curriculum in action*. Sydney: National Centre for Language Teaching and Research, Macquarie University
- Shohamy, E. (1983). Interrater and intrarater reliability of the oral interview and concurrent validity with cloze procedure in Hebrew. Dalam Oller, J. W., Jr. (Ed.) *Issues in language testing research*. Rowley: Newbury House Publishers, Inc.
- Tonjez, M., Wolpow, R., & Zintz, M. (1999). *Integrated content literacy*. Boston: McGraw-Hill College
- Weaver, C. (1994). *Reading process and practice (2nd ed.)*. Portsmouth: Heinemann

Biodata Penulis

Subarso. Lembaga asal, Universitas Negeri Yogyakarta. Pendidikan S1, Pendidikan Bahasa dan Sastra Inggris, IKIP Yogyakarta, tahun lulus 1983. Pendidikan S2, Penelitian dan Evaluasi Pendidikan, Universitas Negeri Yogyakarta, tahun lulus 2008.