

ABSTRACT

Fiharsono, Albertus. 2004. *An Analysis of the Use of Multiple-Choice Items of Reading Test in UAN for Students of SMP*. Yogyakarta: The Faculty of Teachers Training and Education, The Department of Language and Arts Education, The English Language Study Program. Sanata Dharma University.

Ujian Akhir Nasional (henceforth *UAN*) or National Leaving Exam is an academic evaluation that is conducted nationally at the end of school year in the last class of certain stage in education. As an evaluation test, *UAN* should be able to provide an accurate result of the students' achievement. Therefore, the items and the choice of type test should be designed well to describe the students' real ability.

In this research, the writer worked on reading items in *UAN* of 2002 to 2004 for students of *Sekolah Menengah Pertama* (henceforth *SMP*) or junior high school. Two research questions were proposed in this study. Firstly, to what extent did multiple-choice of reading items in *UAN* tests for students of *SMP* measure the objectives of reading stated in 1994 English curriculum? Secondly, to what extent was the use of multiple-choice type test of reading items in *UAN* for students of *SMP* appropriate to measure the objectives of reading stated in 1994 English curriculum?

The first problem was answered by classifying the reading items based on the objectives stated in 1994 English reading curriculum. The second problem was answered by applying Bloom's taxonomy of cognitive domain. The reading items measured on certain objectives were, then, classified based on the six levels of Bloom's taxonomy.

In answering the first problem, the writer found that not all objectives were tested in *UAN*. There were seven, out of eight, objectives measured in *UAN*. The seven objectives were *to find out particular information; to find out the general description of the text; to find out the explicit main idea; to find out the implicit main idea; to find out all the explicitly detailed information; to find out the implicitly stated information; and to interpret the meaning of words, phrases, and sentences*. The objective that was not measured was *to enjoy reading*. The absence of this objective is tolerable since *to enjoy reading* is immeasurable and is achieved through teaching learning activities in class. However, the writer also found that out of seven objectives measured in *UAN*, only six objectives were tested regularly each year. The objective *to find out the explicit main idea* was not tested in *UAN* of 2003 and 2004. The absences of this objective are considered as intolerable. It means that *UAN* did not measure all students' abilities that should be measured.

The writer also got some other findings. Firstly, the objectives *to find out particular information* (scanning) and *to find out the general description of the text* (skimming) are not appropriate to be tested in *UAN* since *UAN* did not present a limitation of time to test these two objectives. Testing these two objectives requires a limitation of time in which the students' need to do fast reading. Secondly, some items of the objective *to interpret the meaning of words, phrases, and sentences*

based on context did not successfully measure the student's real ability in interpreting the meaning of words, phrases, and sentences based on context since the items merely tested the students' vocabulary mastery. Testing this objective requires the alternatives that are simpler or, ideally, that the students know the meaning of all of them so that they need to interpret or deduce the meaning of words, phrases, and sentences based on the context.

In answering the second problem, the writer found that there were three, out of six, levels applied in *UAN*. They are the first level (knowledge), the second level (comprehension), and the fourth level (analysis). As Bloom states, items testing in knowledge, comprehension, and analysis, may apply multiple-choice type test. In short, the use of multiple-choice type test in *UAN* is appropriate.

However, it was also apparent that the items lied mostly on the first two levels, knowledge and comprehension, and it proves that *UAN* items were tested in relatively simple levels. The items were mostly tested on theoretical levels instead of application and production of unique responses as to elicit the students' creativity. It should be remembered also that multiple-choice also limits the students' creativity. In addition, it was also found out that good distractors were not always available in *UAN*, some stems were not well constructed, and some texts were not completely new to the students as to measure the students' real ability.

ABSTRAK

Fiharsono, Albertus. 2004. *An Analysis of the Use of Multiple-Choice Items of Reading Test in UAN for Students of SMP*. Yogyakarta: Program Studi Pendidikan Bahasa Inggris, Fakultas Keguruan dan Ilmu Pendidikan. Universitas Sanata Dharma.

Ujian Akhir Nasional (UAN) merupakan evaluasi belajar tingkat akhir yang diselenggarakan secara nasional untuk seluruh sekolah menengah di Indonesia. Sebagai alat evaluasi atas keseluruhan proses belajar selama tiga tahun di sekolah menengah, UAN harus dapat menunjukkan hasil yang akurat atas kemampuan dan pencapaian belajar siswa. Oleh karena itu, isi soal dan pilihan bentuk soal harus direncanakan dan disusun dengan baik agar dapat menggambarkan kemampuan siswa yang sesungguhnya.

Dalam riset ini, penulis menganalisis soal membaca yang terdapat di dalam UAN untuk siswa SMP yang dipakai pada tahun 2002, 2003, dan 2004. Ada dua permasalahan yang dirumuskan sebagai acuan penelitian ini. Pertama, seberapa jauh soal-soal membaca yang terdapat dalam UAN mengukur tujuan pembelajaran membaca yang tercantum di dalam kurikulum 1994? Kedua, sejauh mana pilihan model soal pilihan ganda layak dan cocok digunakan dalam UAN untuk mengukur tujuan-tujuan pembelajaran membaca tersebut.

Permasalahan pertama dijawab dengan cara mengklasifikasikan soal-soal membaca yang terdapat di dalam UAN berdasarkan tujuan membaca yang tercantum di dalam kurikulum tahun 1994. Permasalahan kedua dijawab dengan cara mengklasifikasikan soal-soal membaca tersebut berdasarkan taksonomi Bloom. Untuk memberikan gambaran yang komprehensif, kedua klasifikasi di atas kemudian digabungkan menjadi satu; soal-soal membaca yang mengukur tujuan-tujuan kurikulum diklasifikasikan berdasarkan taksonomi Bloom.

Dalam menjawab permasalahan pertama, penulis menemukan bahwa dari delapan tujuan membaca yang terdapat di dalam kurikulum ada tujuh tujuan yang ditekankan di dalam UAN. Ketujuh tujuan tersebut adalah: *menemukan informasi tertentu; mendapatkan gambaran umum tentang isi bacaan; menemukan pikiran utama yang tersurat; menemukan pikiran utama yang tersirat; menemukan semua informasi yang tersurat; mendapatkan informasi yang tersirat; dan menafsirkan makna kata, frasa, dan kalimat berdasarkan konteks*. Tujuan yang tidak ditekankan adalah *membaca untuk mendapatkan kesenangan*. Tidak terukurnya kemampuan siswa dalam *membaca untuk mendapatkan kesenangan* ini masih dapat ditoleransi karena tujuan ini pada dasarnya memang tak dapat diukur secara jelas. Selain itu, pencapaian tujuan ini adalah melalui proses belajar mengajar di kelas. Namun demikian, penulis juga menemukan bahwa dari ketujuh tujuan yang terukur di dalam UAN, ternyata hanya enam tujuan yang selalu dimunculkan di dalam UAN setiap tahun. Tujuan *menemukan pikiran utama yang tersurat* tidak ditekankan dalam UAN tahun 2003 dan 2004. Tidak terukurnya kemampuan siswa dalam *menemukan pikiran utama yang tersurat* ini tentu saja tak dapat ditoleransi. Artinya, UAN tidak mengukur semua kemampuan siswa yang seharusnya diukur.

Selain itu, penulis juga menemukan beberapa temuan tambahan. Pertama, tujuan *menemukan informasi tertentu* (scanning) dan *menemukan gambaran umum tentang isi bacaan* (skimming) tidak dapat diteskan di dalam UAN secara efektif karena UAN tidak memberikan pembatasan waktu secara khusus untuk mengukur kemampuan siswa dalam kedua tujuan tersebut. Mengukur kemampuan siswa dalam *scanning* dan *skimming* membutuhkan adanya suatu pembatasan waktu tertentu sehingga kemampuan siswa benar-benar teruji dalam *scanning* dan *skimming*. Kedua, beberapa soal untuk tujuan *menafsirkan makna kata, frase, dan kalimat berdasarkan konteks* juga tidak efektif karena beberapa soal tersebut hanya menguji penguasaan kosa kata. Agar tidak tergelincir ke pertanyaan kosa kata, soal untuk tujuan ini seharusnya menyediakan alternatif pilihan ganda yang lebih sederhana, yang semuanya diketahui maknanya oleh siswa. Dengan demikian, siswa benar-benar menafsirkan makna berdasarkan konteks.

Menjawab pertanyaan kedua, penulis menemukan bahwa ada tiga level, dari enam level kognitif dalam taksonomi Bloom, yang terukur di dalam UAN. Ketiga level tersebut adalah level pertama (pengetahuan), level kedua (pemahaman) dan level keempat (analisis). Seperti yang dikatakan Bloom, soal dalam level pengetahuan, pemahaman, dan analisis, dapat menggunakan model pilihan ganda. Singkatnya, penggunaan model soal pilihan ganda dalam UAN dipandang tepat.

Namun demikian, sangatlah jelas bahwa sebagian besar soal membaca di dalam UAN diukur dalam level pertama (pengetahuan) dan kedua (pemahaman), dan fakta ini membuktikan bahwa soal-soal UAN berada pada level-level yang relatif mudah dan teoretis, bukan level-level yang dapat menggali kemampuan siswa untuk memberikan respon-respon yang unik, beragam, dan kreatif. Harus juga dipertimbangkan bahwa bentuk soal pilihan ganda membatasi kreatifitas siswa. Selain itu, ditemukan pula bahwa soal-soal pilihan ganda di dalam UAN tidak selalu bisa memberikan *distractors* yang baik, beberapa *stemnya* tidak tersusun dengan baik dan beberapa teksnya bukan merupakan teks yang baru.