DOCUMENT RESUME

ED 208 656 FL 012 569

AUTHOR Scholz, George E.; Scholz, Celeste M.
TITLE Multiple-Choice Cloze Tests of EST Discourse: An Exploration.
PUB DATE 81
NOTE 46p.; Revised version of paper presented at the Annual TESOL Convention (15th, Detroit, MI, 1981).

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Chinese; *Cloze Procedure; *English for Special Purposes; Language Proficiency; *Language Tests; Listening Comprehension; *Multiple Choice Tests; Postsecondary Education; *Reading Comprehension; Scoring; Second Language Learning; Social Influences; *Test Construction; Test Reliability; Test Validity; Vocabulary Skills; Writing (Composition)
IDENTIFIERS Algeria; *English for Science and Technology; Learning Strategies; Psychological Influences

ABSTRACT

This study examines the relationship between an open-ended cloze test and its multiple-choice versions generated from three sources: (1) interlingual learner-generated distractors--distractors selected from one language group and administered to a different language group, (2) intralingual learner-generated distractors--distractors selected from one language group and administered to a similar sample of the same language group, and (3) teacher-made distractors--distractors generated by a team of trained ESL/FL instructors and administered to their students. Open-ended cloze tests, multiple-choice cloze formats using intralingual learner-generated distractors, and cloze formats incorporating teacher-made distractors appear somewhat similar in terms of their relationship to general ESL proficiency. In selecting a cloze passage in terms of its discourse, teachers/testers need be aware of and sensitive to their learners' needs and goals. If a language program involves an English for science and technology focus, the appropriate discourse to be taught/tested includes science and technology as academic subjects. If a language program involves a professionally heterogeneous group of learners, a variety of authentic discourse types need to be tested/taught. Some cloze passages are more appropriate than others for language teaching, testing, and learning. Tests and distractors are included. (Author/JK)

Multiple-Choice Cloze Tests of EST Discourse: An Exploration*

George E. Scholz and Celeste M. Scholz

One alternative to facilitate correction of a cloze test is to incorporate a multiple-choice format in an nth word cloze deletion. Two cloze tests, one with the discourse of science as academic subject and the other containing the discourse of science as topic of popular interest were administered to 161 science students in Algeria. A set of three learner-generated distractors were selected from each group of wrong answers, based on the frequency of the incorrect response.

The cloze tests were then rebuilt in a multiple-choice format and administered to 96 scientists studying in the People's Republic of China. Also constructed and administered were multiple-choice cloze tests incorporating teacher-made distractors and distractors from the Chinese scientists' errors.

It was found that there was no significant difference in the sample variance attributable to the types of EST discourse used. However, a greater similarity between the science as academic subject discourse and ESL reading comprehension was indicated. In addition, it appeared that multiple-choice cloze format using teacher-made or learner-generated distractors may be similar to open-ended cloze tests.

*This article is a revised version of the paper presented at the Fifteenth Annual TESOL convention, Detroit, 1981, under the title "Learner-Generated Distractors: A Synthesis of Cloze Procedure and Multiple-Choice Format".

The authors would like to thank the teaching staff at INELEC and GELC for their help and support. Any errors found are solely our own.

George and Celeste Scholz are currently with the Guangzhou English Language Center, a UCLA China Exchange Program in the People's Republic of China.

A good test should be reliable, valid and practical. A reliable test has stable and consistent scores. The validity of a test examines what a test is in fact measuring and how well it does so. Practicality views a test in terms of how easy/difficult that test is to create, administer, score and interpret. Working within these parameters of a good test, a program can then begin to invest time and money into test development.

Apart from examining tests qua tests in the ESL/FL field, a language test may be used to look at learner characteristics as well as "hypotheses about psycholinguistic and sociolinguistic factors in the performance of language based tasks" (Oller 1979:4). If a language test can indicate and/or tap possible learner strategies and plausible hypotheses about the nature of language learning, then the test has immense instructional value. Cloze procedure as well as other pragmatic language tests/tasks (eg. dictation) appear to be a satisfactory language learning tool, as they challenge the hypothesized learner's "pragmatic expectancy grammar" (Oller 1979).

As a test, cloze procedure in an ESL/FL context has been shown to be reliable within a range from .61 to .95 (Brown 1980). When correlated with various discrete-point criterion measures of English language proficiency, cloze procedure has yielded validity coefficients of .63 to .89, a shared variance of 40 to 70% (Brown 1980). Apart from being a test of hypothesized overall language proficiency, cloze procedure may also be considered a reliable and valid test of reading comprehension, having a validity coefficient of .60 to .70 when correlated with multiple-choice reading tests as a criterion measure (Oller 1979: 357, Aitken 1977).

In terms of practicality, a cloze test is relatively easy to create and administer. A passage of appropriate content and length is selected. Allowing

for some context at the beginning and end of the passage, every nth word is systematically deleted (usually 5th to 7th word deletions are used). There should be between 25 to 50 deletions. The amount of time needed to administer a cloze test varies, but 30 minutes seems to be appropriate for a 50 item passage.

One major problem of practicality remains with cloze. In comparison to the multiple-choice format of standardized language tests, cloze is more difficult to score in terms of time and word acceptability by groups of evaluators. Word acceptability may possibly be eliminated as a problem by scoring only for exact words (the original word deleted from the passage), but scoring for acceptable learner responses is fairer to the learner given the nature of the task of completing a cloze test. Ten to fifteen minutes per 50 item passage are usually needed for correction.

One alternative to facilitate the correction of a cloze test is to incorporate a multiple-choice format in a cloze deletion instead of a blank to be completed by a learner. The problem then becomes one of creating distractors. Conceptual problems also arise as to whether or not a multiple-choice cloze would function similarly as compared to an nth deletion open-ended cloze. Would a multiple-choice cloze test be considered an integrative or discrete-point task? Would a multiple-choice cloze test require productive or receptive performance from the learner?

In distractor creation for a multiple-choice cloze test, two studies reviewed used learner-generated distractors selected on the basis of most frequent errors per deletion from an open-ended cloze (Brown 1980; Hinofotis and Snow 1980). In the Brown study, distractors were selected from an intermediate-advanced pretest ESL sample of 164. Distractors for the Hinofotis and Snow study were selected from a sample of 107 students entering an ESL program in the US at

various proficiency levels. In both studies an intermediate ESL text was used for a cloze passage. The two studies found relatively high correlations between an open-ended cloze and its multiple-choice version.

The present study attempted to examine the relationship between an open-ended cloze test and its multiple-choice versions generated from three sources:

1) Interlingual learner-generated distractors - distractors selected from one language group and administered to a totally different language group,
2) Intralingual learner-generated distractors - distractors selected from one language group and administered to a similar sample of the same language group,
3) Teacher-made distractors - distractors generated by a team of trained ESL/FL instructors and administered to their students.

In addition, an attempt was made to explore the relationship between two cloze passages (and their multiple-choice versions) of different but related scientific discourse: science as academic subject and science as topic of popular interest (Widdowson 1976).

Specifically, the hypotheses were:

1) There is no significant difference between the concurrent validity of a cloze passage, its multiple-choice formats, the criterion of ESL proficiency and its subtests.
2) There is no significant difference between cloze passages and their multiple-choice formats with regard to the different discourse of a) science as academic subject and b) science as topic of popular interest.
3) There is no significant difference between the concurrent validity of two cloze passages of different scientific discourse, their multiple-choice formats, the criterion of ESL proficiency and its subtests.

METHOD

Subjects

The subjects were derived from two sources: the Institut National d'Electricite et Electronique (INELEC) in Boumerdes, Algeria, and the Guangzhou English Language Center (GELC), a UCLA-China Exchange Program in the People's Republic of China. The INELEC students were studying electrical technology and engineering through courses taught at the Institute in English by instructors from the US and England. Their INELEC English training which prepared them to attend these courses was given both prior to and at the same time as the technical courses. These technical courses dealt with material that was equivalent to the freshman year at a US university/technical institute in these disciplines. The students ranged in age from 17 to 24 and all had either Arabic or Arabic/Kabylie Berber as a first language and French as a second language. All of the students had received similar educational training in primary and secondary schools in Algeria including English language instruction. This instruction was usually taught by non-native speakers through a traditional grammar orientation. The students could talk about grammar but not use English for communicative purposes.

The students at the Guangzhou English Language Center were scientists, the majority of whom were in the fields of Engineering, Physics, Chemistry and Biology and preparing to study in the US, England and Canada. (Other fields included Medicine and various Agricultural Sciences.) The great majority of these scientists were involved in education as lecturers, researchers and professors in universities and technical institutes throughout China, while some worked in industry. For the most part, the students ranged in age from 37 to 45 and had been selected from many regions of China. This was reflected in their native dialects: Guilinese, Shanghaiese, Cantonese, Fujianese, Zhejianese, Wuhanese,

Shandongese, Sichuanese, Jianshuese and Mandarin. While most GELC students received formal classroom English language instruction in secondary school and/or university/technical institute between twenty and thirty years ago, a small percentage had never studied English formally before. Only one-fourth of all the students had received specialized English language training in the past six or seven years. Little, if any, of this English instruction had been given by native speakers. The remaining seventy-five per cent of the students had studied on their own using a variety of resources, including a limited number of textbooks and tapes available in China, radio programs and educational television. A very small percentage of the students could speak or read Russian, German or Japanese.

Materials

Two passages were selected, NEWTON, containing 415 words, and ENERGY, containing 440 words. Both passages were selected using the criterion that the passages must be taken from texts that a native speaker of English majoring in science or technology would probably encounter during the freshman year at a US university. An ESL instructor selected the science as topic of popular interest text, ENERGY, while a science teacher selected the science as academic subject text, NEWTON (see Appendix A). The latter passage was selected from a physics textbook and the former from a freshman composition textbook. Every 7th word in both passages was deleted, with one or two sentences at the beginning and end of every passage left intact to permit context, for a total of 50 items per passage.

In order to establish the concurrent validity of the NEWTON and ENERGY cloze passages and their multiple-choice formats, subtests from various standardized ESL batteries were selected. The criterion for choosing a particular subtest was the reliability of the subtest to the GELC sample. The five sub-

tests were: the Comprehensive English Language Tests of Structure and Listening (CELT: Structure, CELT: Listening), and the Test of English as a Foreign Language subtests of reading, vocabulary and writing (TOEFL: Reading, TOEFL: Vocabulary and TOEFL: Writing). The total score of these subtests was labelled GELC: Proficiency.

Procedure

Four multiple-choice formats were constructed for this study, -INTER, REV-INTER, -INTRA, and -TEACH (see Appendix B). 161 students at INELEC were administered both NEWTON and ENERGY cloze passages in July 1977. Based on raw scores on each cloze passage, the students were placed into three groups: high (top 25%), low (bottom 25%) and mid (26-74%). From this arrangement an attempt was made to select the most frequent error in each group. However, the majority of errors for an item/group was similar. The procedure was then revised to select the three most frequent wrong responses across all groups, a process similar to the Brown (1980) and Hinofotis and Snow (1980) studies. The three interlingual distractors and the exact word from the cloze passages were then put into a multiple-choice cloze format labelled NEWTON-INTER and ENERGY-INTER.

The original cloze passages, NEWTON and ENERGY, and the multiple-choice formats of NEWTON-INTER and ENERGY-INTER were administered to 91 and 96 GELC/Chinese students respectively in November 1980. This group of students was called Sample-1. In addition, the five subtests of GELC:Proficiency were administered. Selection of students for each group was random. Test administration procedure eliminated an ordering effect. NEWTON and ENERGY were scored on the acceptable word basis. NEWTON-INTER and ENERGY-INTER were scored on the basis of the exact word contained in the multiple-choice format.

The multiple-choice cloze passages, NEWTON-INTER and ENERGY-INTER, were item analyzed using the parameters of .15 through .85 for item facility and above .25

for item discrimination (Oller 1980). Items not meeting these parameters were eliminated. (While this procedure corrupted the 7th word deletion pattern, there was no alternative. As the distractors used in this study were learner-generated, only a very limited number were available.) Hence, the second multiple-choice cloze format, REV NEWTON-INTER and REV ENERGY-INTER, were derived from the interlingual distractors contained in NEWTON-INTER and ENERGY-INTER.

The errors generated from the original NEWTON and ENERGY cloze passages administered to the GELC/Chinese sample were selected as intralingual distractors on the basis of the three most frequent incorrect responses per item. In this fashion, the third multiple-choice format of NEWTON-INTRA and ENERGY-INTRA was developed.

The fourth multiple-choice cloze format was constructed by the GELC instructional team composed of Chinese and American ESL teachers. Working in groups, the teachers wrote what they felt to be the three most effective distractors for each passage. In this way, NEWTON-TEACH and ENERGY-TEACH were created.

The three multiple-choice formats of REV-INTER, -INTRA and -TEACH were then administered to three randomly selected groups at GELC in January 1981. This group of students was identified as Sample-2. The same subtests of GELC: Proficiency as administered to GELC Sample-1 in November 1980 were also given at that time. Test administration procedure eliminated an ordering effect.

Results

The summary test statistics of the criterion measures are given for the GELC/Chinese Sample-1 and Sample-2 in Table 1. In terms of reliability for the criterion measures, estimates ranged from .52 (CELT: Listening, GELC/Chinese Sample-2) to .90 (TOEFL: Vocabulary and GELC: Proficiency, GELC/Chinese Sample-2). The most reliable criterion measure across the GELC/Chinese samples was GELC: Proficiency, with a reliability estimate of .87 and .90 for GELC/Chinese

Sample-1 and Sample-2 respectively. The least reliable criterion measure appeared to be TOEFL: Listening, with a reliability estimate of .58 and .52 for the Chinese samples. TOEFL: Writing had similarly low reliabilities for the Chinese samples (.55 and .60). The reliability estimates may have been affected by the homogeneity of the samples with regard to a specific task and/or the number of items in each criterion measure. The high reliability of the GELC: Proficiency total may have been due in part to the large number of items involved (235).

Insert Table 1 about here.

Summary test statistics for NEWTON, ENERGY and related multiple-choice cloze formats are given in Table 2. The reliability estimates for the GELC/Chinese samples of the two cloze passages and their alternative multiple-choice formats revealed a fairly consistent reliability across formats which had 50 items (.52 to .68). The highest reliability for each cloze passage was with the INELEC/Algerian sample (.76 - NEWTON, .70 - ENERGY). This may have been due to more heterogeneity (hence higher reliability) in the INELEC/Algerian sample.

The lowest reliability estimate for each cloze passage and formats was .52 for NEWTON-INTER and .48 for REV ENERGY-INTER. The greatest disparity in reliability between cloze passages of similar format occurred with the REV NEWTON-INTER (.67) and REV ENERGY-INTER (.48). While the reliability estimates were lower, in general, than previous cloze studies, sample homogeneity may have been responsible.

Insert Table 2 about here.

To test if there were any significant differences between the INELEC/Algerian sample, GELC/Chinese samples on the original cloze passages and the criterion measures, F-ratios were used. These results may be found in Table 3. It appeared that in terms of the cloze passages, the Chinese samples scored significantly higher than the Algerian sample (Table 3A).

Table 3B reveals that there was no significant difference between the GELC/Chinese samples, with the exception of vocabulary. This difference, however, did not affect the F-ratio of GELC:Proficiency scores at any statistically significant level. Thus, in terms of overall proficiency, its composite measures of structure, listening, reading comprehension and writing, the GELC/Chinese samples were similar.

Insert Table 3 about here.

To examine if there were any significant differences between cloze passages of similar format (eg. NEWTON-INTRA and ENERGY-INTRA) t-ratios for testing the difference between correlated variances were calculated. The t-tests computed indicated that there was no significant difference between scores on NEWTON and ENERGY, NEWTON-INTER and ENERGY-INTER, REV NEWTON-INTER and REV ENERGY-INTER, NEWTON-INTRA and ENERGY-INTRA and NEWTON-TEACH and ENERGY-TEACH. This seems to indicate that there was no significant difference in sample variance between a text sele ted as a sample of science as academic subject and a text selected as a sample of science as topic of popular interest. This observation was supported across two distinct language samples and various cloze test formats.

The validity (correlation) coefficients of the open-ended cloze and four multiple-choice cloze formats are given in Table 4. To determine if there were

Table 1

Summary Test Statistics
for Criterion Measures

A) GELC/Chinese Sample-1

Test	N	Items	Range	$\bar{x}$	SD	Reliability (KR_{21})	SEM
CELT: Structure	187	75	53-74	66.08	4.84	.67	2.77
CELT: Listening	187	50	16-41	29.01	5.29	.58	3.44
TOEFL: Reading	187	30	6-28	17.48	4.76	.70	2.60
TOEFL: Vocabulary	187	40	9-37	24.83	6.69	.81	2.92
TOEFL: Writing	187	40	12-34	25.62	4.45	.55	2.99
GELC: Proficiency	187	235	109-210	163.03	19.50	.87	6.97

B) GELC/Chinese Sample-2

CELT: Structure	167	75	46-75	66.82	5.63	.78	2.64
CELT: Listening	167	50	15-46	28.55	4.98	.52	3.46
TOEFL: Reading	167	30	5-28	18.18	4.97	.73	2.56
TOEFL: Vocabulary	167	40	12-39	30.11	7.66	.90	2.48
TOEFL: Writing	167	40	14-35	25.59	4.70	.60	2.98
GELC: Proficiency	167	235	112-205	168.98	21.29	.90	6.76

Table 2.

Summary Test Statistics for Cloze Tests

Cloze Test Type	Sample	Cloze Test Format	N	Items	Range	x	SD	Reliability (KR_{21})	SEM
NEWTON	INELEC/Algerian	open-ended 7th word deletion	161	50	5-45	28.34	6.98	.76	3.42
ENERGY	INELEC/Algerian	open-ended 7th word deletion	161	50	10-40	25.21	6.29	.70	3.46
NEWTON	GELC/Chinese Sample-1	open-ended 7th word deletion	91	50	21-45	33.47	5.80	.68	3.26
ENERGY	GELC/Chinese Sample-1	open-ended 7th word deletion	91	50	20-42	31.75	5.20	.58	3.36
NEWTON-INTER	GELC/Chinese Sample-1	multiple-choice Interlingual distractors	96	50	23-43	35.04	4.70	.54	3.20
ENERGY-INTER	GELC/Chinese Sample-1	multiple-choice Interlingual distractors	96	50	24-46	35.47	4.63	.53	3.18
REV NEWTON-INTER	GELC/Chinese Sample-2	multiple-choice Interlingual distractors revised through item analysis	57	20	5-20	12.86	3.54	.67	1.18
REV ENERGY-INTER	GELC/Chinese Sample-2	multiple-choice Interlingual distractors revised through item analysis	57	20	6-17	11.42	2.99	.48	2.16
NEWTON-INTRA	GELC/Chinese Sample-2	multiple-choice Intralingual distractors	56	50	19-41	30.96	5.56	.63	3.38
ENERGY-INTRA	GELC/Chinese Sample-2	multiple-choice Intralingual distractors	56	50	19-42	30.14	5.60	.63	3.40
NEWTON-TEACH	GELC/Chinese Sample-2	multiple-choice teacher-made distractors	54	50	24-42	33.63	4.76	.52	3.30
ENERGY-TEACH	GELC/Chinese Sample-2	multiple-choice teacher-made distractors	54	50	26-45	36.02	4.67	.55	3.13

Table 3

F-Ratio to Test for Significant Differences in Population Variance

A)

Test	Sample	df	SS	x	F	p*
NEWTON	INELEC/Algerian	160	137348	28.34	1.36	(.05)
	GELC/Chinese Sample-1	90	104990	33.47		
ENERGY	INELEC/Algerian	160	108671	25.21	1.55	(.01)
	GELC/Chinese Sample-1	90	94731	31.75		

B)

CELT: Structure	GELC/Chinese Sample-1	186	820907	66.08	1.02	----
	GELC/Chinese Sample-2	166	750911	66.82		
CELT: Listening	GELC/Chinese Sample-1	186	162587	29.01	1.03	----
	GELC/Chinese Sample-2	166	140252	28.33		
TOEFL: Reading	GELC/Chinese Sample-1	186	61320	17.48	1.07	----
	GELC/Chinese Sample-2	166	58673	18.08		
TOEFL: Vocabulary	GELC/Chinese Sample-1	186	123658	24.83	1.46	(.01)
	GELC/Chinese Sample-2	166	161118	30.11		
TOEFL: Writing	GELC/Chinese Sample-1	186	126433	25.62	1.00	----
	GELC/Chinese	166	113005	25.59		
GELC: Proficiency	GELC/Chinese Sample-1	186	5141065	163.03	1.08	----
	GELC/Chinese Sample-2	166	4843611	168.98		

*p > .05 unless otherwise indicated.

any differences between the validity coefficients of similar formats of NEWTON and ENERGY, Hotelling t values were calculated. Reviewing the concurrent validity coefficients in Table 4 shows moderately strong correlations of both cloze passages in an open-ended format with GELC: Proficiency (.67 - NEWTON, .64 - ENERGY) and reading (.73 - NEWTON, .61 - ENERGY). As well, there were similar validity coefficients of both cloze passages in a teacher-made multiple-choice format with overall GELC: Proficiency (.62 - NEWTON-TEACH, .74 - ENERGY-TEACH). NEWTON (open-ended) correlated significantly higher with the criterion measure of reading comprehension than ENERGY (open-ended), the difference in shared variance being 16%. In general, however, the -INTER, -INTRA and -TEACH multiple-choice formats of ENERGY correlated significantly higher with the criterion measures of reading comprehension and/or GELC: Proficiency than NEWTON passages of the same formats. In addition, the cloze passages revised through item analysis generally had the lowest validity coefficients of all cloze formats involved in this study.

Insert Table 4 about here.

To examine if there were any significant differences between a cloze passage, its alternative multiple-choice formats and the criterion measures, Fisher's z transformation was used. The results of those measures in which significant differences were found are contained in Table 5.

There appeared to be a significant difference in the concurrent validity

Table 4

Concurrent Validity Coefficients

A) GELC/Chinese Sample-1 (N=91)

	NEWTON		ENERGY				
	r	Variance (r^2)	r	Variance (r^2)	Hotelling t	p*	% difference in variance
ENERGY	.60	36%					
CELT: Structure	.39	15%	.40	16%	.12	---	---
CELT: Listening	.51	26%	.43	18%	1.49	---	---
TOEFL: Reading	.73	53%	.61	37%	1.94	(.03)	16%
TOEFL: Vocabulary	.42	18%	.42	18%	0	---	---
TOEFL: Writing	.51	26%	.52	27%	.13	---	---
GELC: Proficiency	.67	45%	.64	41%	.48	---	---

B) GELC/Chinese Sample-1 (N=96)

	NEWTON-INTER		ENERGY-INTER				
	r	Variance (r^2)	r	Variance (r^2)	Hotelling t	p*	% difference in variance
ENERGY -INTER	.50	25%					
CELT: Structure	.39	15%	.39	15%	0	---	---
CELT: Listening	.31	10%	.35	12%	.42	---	---
TOEFL: Reading	.42	18%	.57	32%	1.79	(.04)	14%
TOEFL: Vocabulary	.21	4%	.40	16%	2.00	(.02)	12%
TOEFL: Writing	.45	20%	.53	28%	.94	---	---
GELC: Proficiency	.45	20%	.58	34%	1.58	---	---

*$p > .05$ unless otherwise indicated.

Table 4 (continued)

C) GELC/Chinese Sample-2 (N=57)

	REV NEWTON-INTER		REV ENERGY-INTER				
	r	Variance (r^2)	r	Variance (r^2)	Hotelling t	p*	% difference in variance
REV ENERGY-INTER	.39	15%					
CELT: Structure	.29	8%	.28	8%	7	---	---
CELT: Listening	.29	8%	.08	6%	1.46	---	---
TOEFL: Reading	.56	31%	.39	15%	1.40	---	---
TOEFL: Vocabulary	.16	3%	.31	10%	1.05	---	---
TOEFL: Writing	.36	13%	.32	10%	.29	---	---
GELC: Proficiency	.42	18%	.38	14%	.30	---	---

D) GELC/Chinese Sample-2 (N=54)

	NEWTON-INTRA		ENERGY-INTRA				
	r	Variance (r^2)	r	Variance (r^2)	Hotelling t	p*	% difference in variance
ENERGY-INTRA	.63	40%					
CELT: Structure	.33	11%	.43	18%	.94	---	---
CELT: Listening	.41	17%	.49	24%	.79	---	---
TOEFL: Reading	.53	28%	.64	41%	1.24	---	---
TOEFL: Vocabulary	.35	12%	.42	18%	.66	---	---
TOEFL: Writing	.42	18%	.55	30%	1.33	---	---
GELC: Proficiency	.49	24%	.66	44%	1.93	(.03)	20%

*$p > .05$ unless otherwise indicated.

Table 4 (continued)

E) GELC/Chinese Sample-2 (N=54)

	NEWTON-TEACH		ENERGY-TEACH				
	r	Variance (r^2)	r	Variance (r^2)	Hotelling t	p*	% difference in variance
ENERGY-TEACH	.71	50%					
CELT: Structure	.35	12%	.48	23%	1.39	---	---
CELT: Listening	.44	19%	.48	23%	.43	---	---
TOEFL: Reading	.55	30%	.69	48%	1.83	(.03)	18%
TOEFL: Vocabulary	.53	28%	.66	44%	1.63	---	---
TOEFL: Writing	.54	29%	.59	35%	.59	---	---
GELC: Proficiency	.62	38%	.74	55%	1.71	(.04)	17%

*$p > .05$ unless otherwise indicated.

coefficients of science as academic subject cloze passage, NEWTON, and its multiple-choice formats with respect to reading comprehension. In terms of possible overall ESL proficiency as measured by GELC: Proficiency, the picture is less clear, as it appears that the open-ended cloze format, its multiple-choice interlingual and its multiple-choice teacher-made distractor formats had similar concurrent validity (r^2 = 24 to 45 %).

Focussing on ENERGY, there appeared to be no significant difference between the validity coefficients of this science as topic of popular interest cloze and its multiple-choice formats with GELC: Proficiency, provided that items were not discarded due to item analysis. In terms of the various multiple-choice formats, intralingual (-INTRA) and teacher-made (-TEACH) distractors appeared to be more similar than interlingual distractors. The difference between validity coefficients and the criterion of reading comprehension was negligible, with the ubiquitous exception of REV ENERGY-INTER. One interesting anomaly was that the teacher-made multiple-choice distractor format of the science as topic of popular interest cloze, ENERGY-TEACH, correlated significantly higher (r = .66) with the vocabulary criterion than any other format of the same passage. As well, there was a significant difference of GELC/Chinese Sample-2 in vocabulary as compared to GELC/Chinese Sample-1 (see Table 3). Perhaps this particular cloze passage and distractor type served more as a measure of a receptive vocabulary task than as an integrated task of reading or overall ESL proficiency.

Insert Table 5 about here.

Item facility and discrimination indices are reported in Table 6. The item analysis for each cloze passage and respective multiple-choice formats in-

Table 5

Fisher's z Transformation of Cloze Test

Validity Coefficients with Criterion Measures

(z values above diagonal; p* below diagonal)

NEWTON - Science as Academic Subject

A. 1) TOEFL: Reading

	NEWTON	NEWTON-INTER	REV NEWTON-INTER	NEWTON-INTRA	NEWTON-TEACH
r (r^2)	.73 (53%)	.42 (18%)	.56 (31%)	.53 (28%)	.55 (30%)
NEWTON		3.23	1.71	1.95	1.72
NEWTON-INTER	(.0001)		1.08	.83	1.02
REV NEWTON-INTER	(.04)	---		.22	.04
NEWTON-INTRA	(.03)	---	---		.18
NEWTON-TEACH	(.04)	---	---	---	

A. 2) TOEFL: Vocabulary

	NEWTON	NEWTON-INTER	REV NEWTON-INTER	NEWTON-INTRA	NEWTON-TEACH
r (r^2)	.42 (18%)	.21 (4%)	.16 (3%)	35 (12%)	.53 (28%)
NEWTON		1.58	1.66	.48	.81
NEWTON-INTER	---		1.30	.88	2.16
REV NEWTON-INTER	(.05)	---		1.06	2.20
NEWTON-INTRA	---	---	---		1.15
NEWTON-TEACH	---	(.02)	(.01)	---	

*p > .05 unless otherwise indicated.

Table 5 (continued)

NEWTON - Science as Academic Subject

A. 3) CELC: Proficiency

	NEWTON	NEWTON-INTER	REV NEWTON-INTER	NEWTON-INTRA	NEWTON-TEACH
r (r^2)	.67 (45%)	.45 (20%)	.42 (18%)	.49 (24%)	.62 (38%)
NEWTON		2.19	2.10	1.58	.49
NEWTON-INTER	(.01)		.22	1.60	1.38
REV NEWTON-INTER	(.02)	---		.46	1.42
NEWTON-INTRA	---	---	---		.96
NEWTON-TEACH	---	---	---	---	

ENERGY - Science as Topic of Popular Interest

B. 1) TOEFL: Reading

	ENERGY	ENERGY-INTER	REV ENERGY-INTER	ENERGY-INTRA	ENERGY-TEACH
r (r^2)	.61 (37%)	.57 (32%)	.39 (15%)	.64 (41%)	.69 (48%)
ENERGY		.41	1.72	.28	.79
ENERGY-INTER	(.08)		1.38	.64	1.15
REV ENERGY-INTER	(.04)	---		1.79	2.23
ENERGY-INTRA	---	---	(.04)		.46
ENERGY-TEACH	---	---	(.01)	---	

*$p > .05$ unless otherwise indicated.

Table 5 (continued)

B. 2) TOEFL: Vocabulary

	ENERGY	ENERGY-INTER	REV ENERGY-INTER	ENERGY-INTRA	ENERGY-TEACH
r (r^2)	.42 (18%)	.40 (16%)	.31 (10%)	.42 (18%)	.66 (44%)
ENERGY		.16	.73	0	1.96
ENERGY-INTER	---		.60	.14	2.12
REV ENERGY-INTER	---	---		.66	2.42
ENERGY-INTRA	---	---	---		1.76
ENERGY-TEACH	(.03)	(.02)	(.01)	(.04)	

B. 3) GELC: Proficiency

	ENERGY	ENERGY-INTER	REV ENERGY-INTER	ENERGY-INTRA	ENERGY-TEACH
r (r^2)	.64 (41%)	.58 (34%)	.38 (14%)	.66 (44%)	.74 (55%)
ENERGY		.65	2.07	.20	1.09
ENERGY-INTER	---		1.53	2.28	1.65
REV ENERGY-INTER	(.02)	---		2.03	2.81
ENERGY-INTRA	---	(.01)	(.02)		.80
ENERGY-TEACH	---	(.05)	(.003)	---	

*$p > .05$ unless otherwise indicated.

dictates that both of the original cloze passages were easier for the GELC/Chinese sample than the INELEC/Algerian sample. That these differences were significant are supported by the F-ratios in Table 3. In general, the mean item facility figures indicate that both cloze passages and formats were relatively easy for the Chinese sample, with a range of .61 to .72. The mean item discrimination figures for all samples, cloze passages and formats are low (.21 to .41), indicating that the cloze passages were unable to finely discriminate among the subjects in each sample. This would probably indicate a high homogeneity of each sample with respect to language ability. The fairly low reliability estimates in Table 2 tend to support this.

As an item analysis is a tool used to improve a multiple-choice testing format by eliminating or revising items of extremely high/low facility and low discrimination, an exploration of effective items in both cloze passages and multiple-choice formats was conducted. Due to the high interdependence of cloze items in an open-ended format, however, this aspect of the study can only be speculative, as the actual application of item analysis may not be conceptually sound.

As indicated in Table 6, computing the revised mean item facility and revised mean item discrimination yielded a movement towards the ideal item facility level of .5 and an increase of the discrimination. The number of effective items (as determined by the parameters of item analysis) in the open-ended cloze passages were 25 for NEWTON and 30 for ENERGY, which constituted 50% and 60% of the original 50 items per respective passage. High commonality of effective items was found between the open-ended cloze passages, INELEC/Algerian and GELC/Chinese samples. For example, with 25 of the original items determined to be effective in the GELC/Chinese sample of NEWTON, 22 of the effective items in the INELEC/Algerian sample were similar for a commonality of 88% (22/25).

With the exception of the intralingual distractors (-INTRA) and NEWTON-TEACH, the multiple-choice formats had a much lower effective item commonality with their original cloze passages (33 to 47%). It seems possible that multiple-choice cloze formats were using somewhat different items in establishing their reliability and concurrent validity with criterion measures. However, surface difference in item similarity and obvious difference in testing format appeared to be statistically insignificant when the correlation coefficients with the criterion measure of GELC: Proficiency were reviewed in Table 5. One anomaly in the validity coefficients was the revised interlingual multiple-choice format (REV -INTER) of both cloze passages. Not only was the item commonality with the original cloze passage the lowest of all formats, but the effect of contextual constraints by mechanical deletions was totally eroded.

Insert Table 6 about here

Discussion

While there was no significant difference between the variance of two different cloze passages having similar format, there appeared to be significant differences between cloze passages in relation to the criterion of reading comprehension. Whereas NEWTON (open-ended), involving the discourse of science as academic subject correlated higher with reading than the science as topic of popular interest passage (ENERGY), the exact opposite was true when a multiple-choice format was used. One explanation may be that the science as academic subject cloze passage (NEWTON) correlated with the reading criterion because the discourse of the particular passage was similar to the discourse of the passages involved in the reading criterion measure. That is to say, the discourse of science as academic subject may be more similar to the discourse of ESL as academic subject than the discourse of science as topic of popular interest.

Table 6

Item Facility and Discrimination

for NEWTON and ENERGY Cloze Passages

(see Appendix A & B)

A) NEWTON

Sample:	NEWTON GELC/Ch	NEWTON INELEC/Alg	NEWTON-INTER GELC/Ch	REV NEWTON-INTER GELC/Ch	NEWTON-INTRA GELC/Ch	NEWTON-TEACH GELC/Ch
Mean Item Facility:	.66	.55	.70	.64	.61	.66
Mean Item Discrimination:	.27	.41	.21	.42	.28	.23
Revised Mean Item Facility:	.59	.50	.65	.60	.53	.60
Revised Mean Item Discrimination:	.41	.47	.35	.50	.38	.40
Items Remaining:	25	29	20	14	29	25
Item Commonality of Effective Items with NEWTON:*		22 (88%)	11 (44%)	8 (32%)	16 (64%)	16 (64%)

B) ENERGY

Sample:	ENERGY GELC/Ch	ENERGY INELEC/Alg	ENERGY-INTER GELC/Ch	REV ENERGY-INTER GELC/Ch	ENERGY-INTRA GELC/Ch	ENERGY-TEACH GELC/Ch
Mean Item Facility:	.64	.47	.70	.56	.60	.72
Mean Item Discrimination:	.27	.32	.22	.34	.28	.22
Revised Mean Item Facility:	.58	.51	.62	.61	.57	.62
Revised Mean Item Discrimination:	.37	.40	.37	.42	.38	.39
Items Remaining:	30	32	20	15	32	19
Item Commonality of Effective Items with ENERGY:*		22 (73%)	14 (47%)	10 (33%)	22 (73%)	19 (40%)

*as determined by item analysis.

Science and ESL as academic subjects may share the same discourse characteristics of being "typically metalinguistic, didactic, explicit" (Widdowson 1976:266). The discourse of science as topic of popular interest on the other hand varies "considerably...depending on the readership aimed at" (Widdowson 1976:266). Moreover, a popular science text may be too heavily culture-bound as it translates information from science that relies on the target language culture, "making appeal, and making concession, to social beliefs, attitudes, views of the world" (Widdowson 1979:52).

When a multiple-choice format is introduced, the culture bound and stylistic aspect of the science as topic of popular interest is reduced, making the passage more manageable for the learner, though not necessarily as a reading processing task but perhaps a vocabulary task. As well, the explicit and didactic features of a science as academic subject are increased to the point where the test itself loses its strength as a task of reading comprehension.

Reviewing distractors types, intralingual distractors may be related more closely to teacher-made distractors than to an open-ended format. This relationship appears to be especially true when both teachers and students are in the same context, ie. language program. Similarities between multiple-choice formats, discrete-point in nature, are greater than between integrative and discrete-point formats. Hence, multiple-choice formats are somewhat different than integrative open-ended cloze formats.

These differences, however, do not necessarily obviate the intralingual or teacher-made multiple-choice cloze formats' membership to integrative tasks. The differences between a multiple-choice and open-ended cloze format may be due to a receptive/productive demand. A multiple-choice cloze task may be receptive in the sense that a learner need only identify the correct answers, while an open-ended cloze task requires written production of the correct

answer or a semantically acceptable equivalent. Moreover, what claims could a multiple-choice cloze format make to being discrete-point in nature when the items are selected on the basis of nth word deletion rather than an a priori item selection based on syntax or lexis?

Conclusion

Responding to the initial hypotheses in this paper, the results of this study indicate that open-ended cloze tests, multiple-choice cloze formats using intralingual learner-generated distractors and cloze formats incorporating teacher-made distractors appear somewhat similar in terms of their relationship to general ESL proficiency. Based on a sample of science students and scientists, it appears that an open-ended cloze containing the discourse of science as academic subject remains a more valid measure of ESL reading comprehension than its multiple-choice counterparts. On the other hand, a cloze passage representing the discourse of science as a topic of popular interest appears to be of the same validity in terms of ESL reading comprehension as its multiple-choice formats employing intralingual or teacher-made distractors.

While there is no significant difference of population variance between the discourse of science as academic subject and science as topic of popular interest passages, the former seems more closely related to ESL reading comprehension than the latter. A multiple-choice format proceeds to make opposite claims--closely relating a science as topic of popular interest discourse to ESL proficiency, reading and vocabulary. It appears that across discourse types, cloze procedure and its multiple-choice varieties make competing claims. However, in the testing of a given discourse, an open-ended cloze procedure might also be accomodated by both intralingual learner-generated and teacher-made distractors. This may serve to alleviate the problem of scoring cloze. Yet, instructionally, open-ended cloze remains stronger due to its ability of ex-

ploiting doubt and choice, two crucial elements in the development of communicative competence.

In selecting a proper cloze passage in terms of its discourse, teachers/testers need to be aware of and professionally sensitive to their learners' needs and goals. If a language program involved an English for Science and Technology (EST) focus, the appropriate discourse to be taught/tested would include science and technology as academic subjects. If a language program is involved with a professionally heterogeneous group of learners, then perhaps a variety of authentic discourse types would have to be tested/taught. Some cloze passages may be more appropriate than others with respect to language teaching, testing and learning.

REFERENCES

Aitken, K. G. 1977. Using Cloze Procedure as an Overall Language Proficiency Test. TESOL Quarterly 11, 1:59-67.

Ary, D., Jacobs, L. C., Razavieh, A. 1972. Introduction to Research in Education. New York: Holt, Reinhart and Winston.

Brown, J. D. 1980. Relative Merits of Four Methods for Scoring Cloze Tests. Modern Language Journal 64, 3:311-317.

Guilford, J. P., Fruchter, B. 1973. Fundamental Statistics in Psychology and Education, New York: Mc-Graw Hill.

Harris, D. P. 1969. Testing English as a Second Language. New York: Mc-Graw Hill.

Hinofotis, F., Snow, B. G. 1980. An Alternative Cloze Testing Procedure: Multiple-Choice Format. Research in Language Testing. J. Oller and K. Perkins (Eds.). Rowley, Mass.: Newbury House Publishers, Inc.

Minium, E. W. 1970. Statistical Reasoning in Psychology and Education. New York: John Wiley.

Morray, M. 1981. INELEC: Teamwork in an EST Program. English Teaching Forum 19, 1.

Oller, J. W. 1979. Language Tests at School. London: Longman Group Limited.

Oller, J. W., Perkins, K (Eds.). 1980. Research in Language Testing. Rowley, Mass.: Newbury House Publishers, Inc.

Widdowson, H. G. 1976. The Authenticity of Language Data. On TESOL '76. Washington, D.C.: TESOL.

Widdowson, H. G. 1979. Explorations in Applied Linguistics. London: Oxford University Press.

Appendix A

Cloze: Newton
(Science as Academic Subject)

Newton assumed that gravitational force depended on the masses of the attracting bodies as well as inversely on the square of their distance of separation. He then generalized his concept of gravitational attraction into a law of universal gravitation. He thought that all bodies, no (1) (matter)* where they were located, exerted forces (2) (of)*gravitational attraction upon one another. In (3) (order)* to discover the exact nature of (4) (this)*attractive force, he had to consider (5) (bodies)*of various different masses at significantly (6) (different) distances from one another. He could (7) (not)*change the distance between the center (8) (of)* the earth and a body on (9) (the)*earth very appreciably, however. It was (10) (for) this reason that he first compared (11) (the) motion of the moon and a (12) (body)* on earth. The force between different (13) (macroscopic)*bodies on the earth was so (14) (small) that it was not detected in (15) (Newton's)*time. Newton apparently realized that this force (16) (was) small and easily masked by frictional (17) (or) other forces. Hence, he focused his (18)(attention) on the motion of the planets (19) (in)* an attempt to confirm his ideas.

(20) (The)*earliest scientific attempts to understand the (21) (solar)*system were made by the Greeks. (22)(A) detailed description of the conclusions of (23) (Greek) astronomers was given by Ptolemy. His (24) (system) is known as the Ptolemaic, or (25)(geocentric) theory. It assumes that the earth (26) (is)*stationary at the center of the (27) (universe), with the sun, moon, planets and (28) (stars) all revolving about the earth in (29) (complex) orbits. This theory was accepted for (30) (almost)*fifteen centuries and greatly influenced philosophy (31) (and) literature as well as science. However, (32) (the)*theory was quite complex and could (33) (not) quantitatively account for an increasing number (34) (of)*observations. In the sixteenth century Copernicus (35) (suggested)* that a simpler description of celestial (36) (motions)*could be given by assuming that (37) (the)* sun was at rest at the (38) (center)*of the universe. In the Copernican (39) (or) heliocentric theory, the earth was a (40) (planet)*rotating on its axis and revolving (41) (around)*the sun, and the other planets (42) (had) similar motions.

The growing controversy over (43) (the) two theories stimulated astronomers to obtain (44) (more) accurate observational data. Such data were (45) (compiled) by Tycho Brahe, who was the (46) (last) great astronomer to make observations without (47) (the) use of a telescope. His data (48) (on)* planetary motions were analyzed and interpreted (49) (for) about twenty years by Johannes Kepler, (50) (who) had been Brahe's assistant. Kepler found important regularities in the motion of the planets. These regularities are known as Kepler's three laws of planetary motion.

*items ineffective through item analysis (see Table 6)

Cloze: Energy
(Science as Topic of Popular Interest)

Just after World War II, uranium seemed to many to be just another fuel that could be exploded in anger or consumed more slowly in calculation - doing more damage than other fuels in the first case, but more good in the second.

Unfortunately, uranium is more than just another fuel - and worse.

What the general public didn't realize, (1) (at) first, was that the atomic bomb, (2) (when) exploded, did not do all its (3) (damage)* at the moment of explosion. When (4) (uranium)undergoes fission, the ash it produces (5) (is) atoms that are fiercely radioactive, and (6) (if) these spread out over the land, (7) (they) will carry a deadliness that would (8) (last) for decades.

Uranium undergoing controlled fission (9) (produced) useful energy for mankind - but the (10) (radioactive) ash is still produced; it is (11) (still) there. It must not be allowed (12) (to)* escape into the environment, either before (13) (disposal)* or after. This aspect of fission (14) (energy) makes it a most uneasy substitute (15) (for) the fossil fuels, coal and oil.

(16) (In) fact, as we look toward the (17) (future), we find that coal, oil and (18) (uranium) are alike in this: we can (19) (depend) on none of them for the (20) (21st)*century. The oil will be, for (21) (the)* most part, gone by then. The (22) (coal), in the quantities needed, will destroy (23) (the) land in the course of being (24) (dug) out, and will pollute the air (25) (in)* the course of being burned. As (26) (for)* uranium, that will have within it, (27) (always)* the fearsome danger of the radioactive (28) (contamination)* of the world.

Yet we must (29) (have)* energy if civilization is to survive. (30) (Where) will it come from, if oil, (31) (coal)* and uranium are eliminated as possible (32) (sources)?*

One thing to remember is that (33) (there)* are many sources of energy that (34) (we) do use, have used in the (35) (past), and can use more of in (36) (the)* future. There is still the energy (37) (of)* the wind, of running water, of (38) (the)* tides, of the Earth's internal heat.

(39) (Even) all of these together may not (40) (be) enough to supply all of mankind's (41) (energy) needs. Properly exploited, however, they can (42) (offer) man a large percentage of the (43) (energy)* he needs, and can relieve the (44) (weight) of our dependence on oil, coal (45) (and)* uranium, until the really suitable source (46) (is) developed.

Each of these sources, moreover, (47) (is) inexhaustible. Wind and running water will (48) (last) as long as the sun shines, (49) (the) tides will be there as long (50) (as)* the Earth turns, and the Earth will have internal heat as long as it exists. All that energy is there and is being expended; it is only necessary to use it, rather than letting it go to waste.

---

*items ineffective through item analysis (see Table 6)

Appendix B

Test - Cloze: Newton

Interlingual Distractors (Algerian/Chinese)	Intralingual Distractors (Chinese/Chinese)	Teacher-made Distractors (GELC team/ Chinese)
1. *(matter) place importance one	*(matter) place space regard	*(matter) place wonder need
3. *(order) fact experience way	*(order) addition time trying	*(order) fact time contrast
5. (bodies) all some many	(bodies) all some many	(bodies) forces experiments pieces
7. (not) also be then	*(not) also easily make	(not) only easily always
9. (the) which that on	*(the) which that on	*(the) an this another
11. +(the) a with between	+(the) a with between	*(the) each all some
13. *(macroscopic) mass two separated	(macroscopic) mass two distance	*(macroscopic) attractive moving gravitational

Interlingual Distractors (Algerian/Chinese)	Intralingual Distractors (Chinese/Chinese)	Teacher-made Distractors (GELC team/ Chinese)
2. *(of) by are on	(of) by as a	(of) by with from
4. (this) an body gravitational	(this) order earth gravitation	(this) an every its
6. *(different) same the of	*(different) same various far	*(different) long measurable great
8. *(of) on to and	*(of) on to and	*(of) at in near
10. +(for) by in not	*(for) just from of	+(for) by from with
12. (body) motion center gravitation	*(body) motion center gravitation	*(body) motion matter force
14. *(small) different difficult far	+(small) different close various	+(small) significant changeable strong

*ineffective items (Table 6)
+item commonality (Table 6)

	Interlingual Distractors (Algerian/Chinese)	Intralingual Distractors (Chinese/Chinese)	Teacher-made Distractors (GELC team/ Chinese)
15.	(Newton's)	(Newton's)	(Newton's)
	same	same	Galileo's
	this	that	modern
	the	any	enough
17.	*(or)	+(or)	+(or)
	with	with	with
	to	to	beside
	of	of	through
19.	(in	(in)	*(in)
	and	and	under
	of	with	with
	for	as	for
21.	*(solar)	(solar)	*(solar)
	gravitational	sun	gravitational
	old	universe	universal
	motion	planet	Newtonion
23.	+(Greek)	*(Greek)	+(Greek)
	the	the	European
	what	which	solar
	an	early	early
25.	*(geocentric)	+(geocentric)	+(geocentric)
	Ptolemy's	Ptolemy	solar
	his	Ptolemaic	gravity
	the	heliocentric	planetary
27.	*(universe)	*(universe)	+(universe)
	earth	earth	atmosphere
	space	space	world
	system	sun	system
29.	+(complex)	+(complex)	+(complex)
	the	the	regular
	its	its	changing
	same	planets	confusing

*ineffective items (Table 6)
+item commonality (Table 6)

	Interlingual Distractors (Algerian/Chinese)	Intralingual Distractors (Chinese/Chinese)	Teacher-made Distractors (GELC team/ Chinese)
16.	+(was)	+(was)	*(was)
	is	is	got
	are	too	looked
	of	very	became
18.	+(attention)	+(attention)	+(attention)
	experience	discover	instrument
	comparison	idea	telescope
	force	notice	concept
20.	*(The)	(The)	(The)
	An	An	Same
	Some	Some	Very
	In	In	All
22.	+(A)	+(A)	+(A)
	He	He	This
	They	They	Their
	It	More	The
24.	+(system)	*(system)	+(system)
	conclusion	conclusion	attempt
	experiment	idea	experiment
	name	name	exploration
26.	*(is)	(is)	*(is)
	was	was	was
	rotates	rotated	revolves
	has	located	hangs
28.	*(stars)	+(stars)	*(stars)
	are	are	galaxies
	they	they	meteors
	the	were	heavens
30.	(almost)	(almost)	(almost)
	the	the	those
	only	before	over
	all	at	all

	Interlingual Distractors (Algerian/Chinese)	Intralingual Distractors (Chinese/Chinese)	Teacher-made Distractors (GELC team/ Chinese)
31.	*(and) in of by	*(and) in religion on	*(and) through of with
33.	*(not) be have take	+(not) be explain make	+(not) always even also
35.	*(suggested) said is was	*(suggested) gave assumed considered	(suggested) asked denied wondered
37.	*(the) as at that	*(the) as at that	*(the) when if a
39.	+(or) the and of	+(or) the sun celestial	+(or) named and held
41.	*(around) on to of	*(around) on motion from	*(around) with above towards
43.	*(the) this in are	+(the) this of between	+(the) their Newton's Ptolemy's
45.	*(compiled) made given discovered	*(compiled) made gave found	+(compiled) invented named suggested

*ineffective items (Table 6)
+item commonality (Table 6)

	Intenlingual Distractors (Algerian/Chinese)	Intralingual Distractors (Chinese/Chinese)	Teacher-made Distractors (GELC team/ Chinese)
32.	*(the) Ptolemy a all	*(the) Ptolemy Ptolemaic his	*(the) man's Newton's geocentric
34.	(of) for to in	(of) by under astronomic	(of) by with through
36.	*(motions) which it that	(motions) which data system	motions) bodies experiments theories
38.	*(center) earth rest end	*(center) system rest gravitation	*(center) beginning top middle
40.	*(planet) center circle system	*(planet) subject itself ceaseless	*(planet) star sun meteor
42.	+(had) have in are	+(had) have were did	*(had) with showing gave
44.	*(more) the an by	*(more) the an much	*(more) the some few
46.	*(last) most first best	+(last) most Greek famous	+(last) one first only

	Interlingual Distractors (Algerian/Chinese)	Intralingual Distractors (Chinese/Chinese)	Teacher-made Distractors (GELC team/ Chinese)
47.	+(the)	+(the)	+ (the)
	an	an	this
	to	to	his
	a	by	any
49.	*(for)	+(for)	* (for)
	after	later	after
	by	into	within
	in	before	over

Test - Cloze: Energy

1.	*(at)	*(at)	* (at)
	the	the	the
	that	that	for
	in	damage	in
3.	(damage)	(damage)	(damage)
	energy	energy	impact
	effect	good	effect
	effects	work	damages
5.	*(is)	+ (is)	+ (is)
	by	by	destroy
	the	from	destroys
	are	in	are
7.	+ (they)	+(they)	+ (they)
	it	it	it
	and	and	that
	which	which	theirs
9.	+ (produced)	+(produced)	* (produced)
	and	produce	for
	by	provide	producing
	is	as	results

	Interlingual Distractors (Algerian/Chinese)	Intralingual Distractors (Chinese/Chinese)	Teacher-made Distractors (GELC team/ Chinese)
48.	*(on)	(on)	* (on)
	of	of	of
	is	proved	about
	was	showed	around
50.	*(who)	* (who)	* (who)
	that	that	that
	he	he	he
	and	it	one
2.	+(when)	* (when)	+ (when)
	it	it	if
	which	while	while
	was	as	after
4.	+(uranium)	+ (uranium)	* (uranium)
	the	atom	energy
	an	explosion	plutonium
	it	bomb	reactors
6.	+(if)	+ (if)	* (if)
	all	then	therefore
	and	after	during
	make	makes	nonetheless
8.	*(last)	+ (last)	* (last)
	be	be	increase
	rest	damage	diminish
	use	keep	lessen
10.	*(radioactive)	*(radioactive)	+(radioactive)
	atoms	atom	energetic
	most	uranium	invisible
	energy	fission	harming

*ineffective items (Table 6)
+item commonality (Table 6)

	Interlingual Distractors (Algerian/Chinese)	Intralingual Distractors (Chinese/Chinese)	Teacher-made Distractors (GELC team/ Chinese)
11.	*(still) in not dangerous	+(still) remained problem over	*(still) in yet from
13.	*(disposal) explosion fission now	*(disposal) explosion fission use	*(disposal) explosion fission production
15.	*(for) of to from	+(for) of to like	+(for) of to with
17.	+(future) energy uranium world	+(future) energy fuel fuels	*(future) energy past production
19.	*(depend) find use say	+(depend) find use consume	*(depend) work rest substitute
21.	*(the) its example using	*(the) its example a	*(the) its their a
23.	*(the) in by a	+(the) in from earth	+(the) many its of
25.	(in) of and by	(in) when while during	*(in) for after on

*ineffective items (Table 6)
+item commonality (Table 6)

	Interlingual Distractors (Algerian/Chinese)	Intralingual Distractors (Chinese/Chinese)	Teacher-made Distractors (GELC team/ Chinese)
12.	*(to) the b an	(to) which it that	*(to) the for their
14.	+(energy) which that it	+(energy) which process uranium	*energy) control process force
16.	*(In) The A By	*(In) The At By	*(In) The A On
18.	*(uranium) fuels fuel they	+(uranium) fuels gas fossil	*(uranium) resources reactors energy
20.	*(21st) last 20th 19th	(21st) last 20th recent	(21st) passing future near
22.	*(coal) uranium energy oil	+(coal) uranium fuel oil	*(coal) power energy coals
24.	+(dug) burned spread it	+(dug) burned exploded digged	*(dug) brought found taken
26.	(for) the energy an	*(for) the radioactive to	(for) regarding with concerning

	Interlingual Distractors (Algerian/Chinese)	Intralingual Distractors (Chinese/Chinese)	Teacher-made Distractors (GELC team/ Chinese)
27.	*(always) ash is and	*(always) however has produces	*(always) because with for
29.	*(have) need find know	(have) need save gain	(have) need keep make
31.	*(coal) is fuel fuels	*(coal) is fuel fuels	*(coal) reactors energy resources
33.	*(there) they uranium oil	*(there) they uranium sun	*(there) they these those
35.	+(past) present world life	+(past) present period future	+(past) present world earth
37.	*(of) by in from	(of) source in as	*(of) by using with
39.	*(Even) In If At	+(Even) Put Although Almost	*(Even) In With Therefore
41.	*(energy) of he it	+(energy) living daily consume	+(energy) resourceful material social

	Interlingual Distractors (Algerian/Chinese)	Intralingual Distractors (Chinese/Chinese)	Teacher-made Distractors (GELC team/ Chinese)
28.	(contamination) and effects energy	(contamination) ash all pollution	(contamination) contaminating polluting pollutants
30.	*(Where) What We Energy	*(Where) What Which When	*(Where) What Which When
32.	(sources) as now then	(sources) as energy exhausted	(sources) contaminants substitutes substitutions
34.	+(we) can they man	+(we) can will may	*(we) cars they animals
36.	*(the) for a on	*(the) for few near	*(the) its a near
38.	(the) heating sun using	(the) running solar raising	(the) these water those
40.	*(be) sufficient have use	*(be) sufficient much rich	+(be) get obtain find
42.	+(offer) help be supply	*(offer) help make meet	*(offer) help aid produce

*ineffective items (Table 6)
+item commonality (Table 6)

	Interlingual Distractors (Algerian/Chinese)	Intralingual Distractors (Chinese/Chinese)	Teacher-made Distractors (GELC team/ Chinese)
43.	*(energy) life mankind's object	(energy) sources alternative amount	*(energy) sources uranium plutonium
45.	*(and) in of fuel	*(and) gas radioactive energy	*(and) or with but
47.	*(is) are was and	+(is) are was isn't	+(is) are was were
49.	+(the) and using but	*(the) and risen there	*(the) those no few

	Interlingual Distractors (Algerian/Chinese)	Intralingual Distractors (Chinese/Chinese)	Teacher-made Distractors (GELC team/ Chinese)
44.	*(weight) energy need most	*(weight) energy needs situation	+(weight) lack power force
46.	*(is) was are will	+(is) was are will	*(is) was are has
48.	+(last) be take have	+(last) be use run	+(last) be work decrease
50.	*(as) time that possible	*(as) time that possible	*(as) after if while

*ineffective items (Table 6)
+item commonality (Table 6)