

SLEP テストによる英語能力測定： 文京女子大学1年生の分析⁽¹⁾

Jeff Johnson*・アレン玉井光江・加須屋裕子**

Abstract

The Secondary Level English Proficiency (SLEP) test was given to 140 first-year women's university students. Reliability estimates ranged from .18 to .87 for the 10 subsections, with the dictation and comic subsections the most reliable, the map and second talk subsections the least reliable. The photograph, map, first talk, and cloze subsections had a high proportion of items that did not discriminate well between higher and lower level subjects. The second talk and reading passage subsections were considered too difficult, whereas the comic subsection was too easy. Inter-subsection correlations and factor analysis suggest that the listening section of the SLEP is more unidimensional than the reading section.

Key Words : English proficiency, item analysis, the SLEP test

はじめに

大学における外国語教育は過去10年を境に大きく変化してきている。よく言われるのが二極化現象である。外国語教育（ここでは主に英語教育をさす）を熱心に行う大学とカリキュラム

Measuring Proficiency with the SLEP Test : An Analysis with First-Year Women's Students

* Jeff Johnson ** Mitsue Allen-Tamai・Hiroko Kasuya (Department of Human Studies)

(1) 本研究は平成10年度文京女子大学共同研究費を使って行われた。また本研究は平成11年6月全国語学教育学会 (JALT) Teacher Education 大会にて口頭発表された。

Correspondence Address : Department of Business Administration, Bunkyo Women's University, 1196 Kamekubo, Oimachi, Iruma-gun, Saitama 356-8533, Japan.

Accepted October 18, 1999.

Published December 20, 1999.

から排除してしまう大学、外国語が大学の必修科目でなくなった以降多くの試みが行われているようだ。

本学では、英語は保育心理専攻の学生に対しては保育士資格を得るための必修科目の1つとなっているため毎年ほとんどの学生が受講する。しかし、福祉心理専攻の学生にとっては選択科目であるため受講者数は毎年在籍数の3分の1程度にとどまっている。そこで学部にあった英語教育を行うことを目的とし、本学人間学部1年生の英語能力と学習動機づけに関する調査を実施することにした。

研究目的

この研究調査の目的の1つは、将来の英語能力クラス分けテストとして使用可能な正確かつ信頼度の高い測定基準を作ることであった。The Secondary Level English Proficiency (2級英語能力テスト、これ以降 SLEP と呼ぶ) のフォーム3を測定に使用した。この研究は、古典的テスト理論による分析、信頼度予測、SLEP テストの聴解力と読解力セクションおよびそのサブセクションの dimensionality (特性) の分析結果を報告するものである。この研究の研究目的は、次の4つである。

1. SLEP テストの聴解力、読解力セクションが本学部の学生にとってどれほど信頼性があるものなのか。
2. SLEP テストの各項目は、本学部の学生にとって適切な困難度をもつか。
3. SLEP テストは、本学部の学生の英語能力の差を測定することができるか。
4. SLEP テストのサブセクションは聴解力と読解力の異なった特性を測定しているか。

方 法

被験者と手順

本研究の被験者は本学人間学部保育心理専攻と福祉心理専攻の学生である。そのサンプルは保育心理専攻の1年生が69% (106名)、保育心理専攻と福祉心理専攻の2年生が各4名ずつ、そして福祉心理専攻の1年生が22% (34名) であった。これらの学生は英書講読IIのクラスを受講していたが、そのクラスは、保育心理専攻の学生には必修科目であり、福祉心理専攻の学生には選択科目であった。SLEP テストは、本研究の研究者であり、これらのクラスの担当教員でもある2名の教師によって、1998年秋学期の第1週目に、聴解力テスト、2週目に読解力テストの順に行われた。テストは通常のクラス内で行われ、学生はこのテストが成績には一切関係がないことを告げられた。テストの各セクションの人数が同じではないのは、欠席者がでた

ことによる——143名が聴解力テストを、146名が読解力テストを受験し、したがって両方のテストを受けたのは141名であった。

材料

SLEP テストは、Educational Testing Service (ETS) により考案され、その本来の目的は英語を母語としない生徒がアメリカ中・高等学校（7年生から12年生）に入る時の英語のクラス分けの目安にすることにあった。我々が本学の学生にこのテストを実施しようとした理由は、次による。まず、このテストは他の英語能力テスト、例えば TOEFL や TOEIC に比べ易しいであろうと推察したこと。そして、このテストが授業時間内で実施可能であり、テストにかかる費用が割合に安価であったこと。さらに、他の大学でもこのテストをクラス分けやコース査定目的に使われていることなどの理由による。

SLEP テストは、聴解力、読解力セクション各75問の四択問題に分けられている。さらに、我々はこの聴解力セクションを Photograph (写真), Dictation (口述), Map (地図), Talk 1 (会話1), Talk 2 (会話2) の5部門に分けて考えた。Photograph の部では、文字は全くなく、25枚の白黒写真のコピーが印刷されていて、受験者は4つの文を聞き、その中から各写真の内容を描写している文を1つ選択することになる。SLEP 申込書(1998)によると、この Photograph の部は、「Minimal pair 対照, 接続, 強弱, 音のかたまり, 時制, 態, 前置詞, 語彙の正しい認識」(p.10) をテストするものである。Dictation の部では、18の文が読まれ、その各文ごとに似たような文が4つ示されている。受験者はその中からテープで聞いた文と同じものを1つ選ぶ。Map の部では、名前のついた通り、建物、また色々な方向に向かっている4台の車が示されている地図が提示されている。受験者は、車の中で交わされている2人の会話を聞き、車の位置や方向などを考慮に入れ、その会話が交わされているのがどの車であるのかを選ぶ。この Map の部の問題は12問ある。聴解力セクションの最後の2部門は、高校生と職員との会話文である。査定目的の為に、我々はこの2つの会話を異なったサブセクションとして扱うことにした。各会話について1, 2問の質問とそれに対する4つの選択肢が示されその中から受験者が回答を1つ選ぶという形式になっている。最初の会話は、14の質問文、2番目の会話は6問の質問文からなっている。問題の指示や各サブセクションの例題にかかる時間も含めて、聴解力セクションの総時間は約40分である。

読解力セクション（これはまた、文法や語彙を測定する意味もある）も、5つのサブセクションに分けられる。それらは Comic (漫画), Illustration (イラスト), Cloze (穴埋め), Meaning (意味), そして Passage である。Comic の部には、10問の問題がある。受験者は、まず Bill Keane 作の Family Circus というアメリカの漫画を例題として見る。それぞれの質問文の答えとしてどの漫画の1シーンが一番合っているのかを選ぶ。フォーム3では、12の違った活動が1日の12時間に分けられ、時計の文字盤の上に重ねられて書いてある。Illustration のセクションは13問の問題があり、それぞれの文は4つの線描きイラストからなっていて、そのうち1つ

がその文を説明している。このセクションは、「前置詞，代名詞，副詞，数詞の使い方」(SLEP 申込書, p.15)をテストするものである。次の2つのサブセクションは Cloze と Meaning と名づけた。この2つのサブセクションにはそれぞれ穴埋め選択問題が組み込まれた3つの Passage があり，その後，読解力問題が続く。各 Cloze 問題は受験者が選択することになる単語1語，2語連続動詞，前置詞句などの4つのオプションをもつものである。最初の Passage は9つの Cloze 問題があり，3つの読解力問題がある。2番目の Passage には，12の Cloze 問題があり，4つの読解力問題がある。最終の Passage は9つの Cloze 問題と7つの読解力問題になっている。読み物セクションの最後のサブセクションは，Passage である。これは，Cloze サブセクションの Passage よりも長い読み物で，8つの選択問題が続く。受験者は，45分で読み物セクションを完成させる。

結 果

表1は，Descriptive Statistics と SLEP テストの10項目のサブセクションの信頼度係数，また聴解力セクションと読解力セクションの総数を表している。すべての被験者が各サブセクションの全項目を試みた訳ではないがそこには受験者数とそれぞれのサブテストの項目数があげられている（おのおの問題に対する欠落データ数が付録にある表の最後の欄に示されている）。セクションの難易度は central tendency の3つの測定値，mode，中間値，平均値の中に

Table 1 SLEP Test Descriptive Statistics and Reliability

	<i>N</i>	<i>k</i>	central tendency			dispersion			reliability	
			mode	medn	mean	<i>SD</i>	range	skew	kurt	<i>KR20</i>
<i>Listening Comprehension</i>										
Photograph	143	25	13	13	12.80	2.80	3-20	-.37	.96	.39
Dictation	143	18	13	12	11.27	3.70	2-18	.49	-.42	.77
Map	143	12	5	5	4.71	1.85	0-10	.26	.13	.25
Talk 1	143	14	4	4	4.23	1.81	0-9	.40	-.10	.18
Talk 2	143	6	1	1	1.45	1.01	0-6	.66	.40	.35
Total	143	75	35	35	34.45	6.66	14-57	.02	.32	.72
<i>Reading Comprehension</i>										
Comic	146	10	10	9	8.06	2.54	0-10	-1.33	.70	.87
Illustration	146	13	8	9	8.72	1.96	2-13	-.33	.39	.54
Cloze	146	30	12	11	10.93	3.56	1-19	.01	-.29	.55
Meaning	146	14	4	3.5	3.55	1.95	0-13	.35	-.07	.62
Passage	146	8	0	2	1.78	1.73	0-7	.51	-.77	.30
Total	146	75	35	33	33.04	7.65	16-58	.02	-.06	.82

Note. *N*=number of test-takers, *k*=number of items, medn=median score, *SD*=standard deviation, skew=skewness, kurt=kurtosis, *KR20*=Kuder-Richardson formula 20.

見られる、また、点数の広がりには標準偏差、幅、歪み、kurtosis 測定値で、示されている。表の平均値を見るといくつかのサブセクションは本学の学生にとってかなり難しかったことがわかる。聴解力セクションの中の地図、会話1と2がいずれも平均値50%をかなり下回っている。これと同じ結果が読解セクションの Cloze, Meaning, Passage のサブセクションにもでてくる。その一方で、Comic サブセクションは、中間値が10の満点であることから受験者にとっては大変易しすぎるものである。

信頼性

各サブテストと聴解力、読解力セクション全体の信頼度係数を示す、表1の一番右側の欄に、表されている。聴解力セクションは、Kuder-Richardson 公式20 (KR20) による測定で.72の信頼度である。読解力セクションの信頼度予測は.82である。サブセクションの信頼度は.18から.87の幅である。これらの数値の多くは、サブテストの長さを考慮に入れてもかなり低い。サブテストの中ではほんの少ししか満足できる信頼度に近い項目はない。——聴解力セクションの口述サブセクションが KR20信頼度係数.77で、読解力セクションの Comic と Meaning のサブセクションが、それぞれ係数.87と.62である。最も低い信頼度数のサブセクションはまた最も難しいセクションでもある——会話1 (KR20=.18, 平均値=4.23, 項目数14), 会話2 (KR20=.35, 平均値=1.45, 項目数6), Passage (KR20=.30, 平均値=1.78, 項目数8)。

サブセクションの dimensionality (特性)

SLEP テストの10のサブセクション間の関係を調べた。それは、これらのサブセクションが本学の学生の聴解力と読解力の違った特質を測定することができるかを調べるためであった。まず、サブテスト間の相関係数を測定し、さらに聴解力因子と読解力因子が現れるかを調べる為に10のサブセクションの因子分析を行った。これらの分析は10変数を比率に変換し SPSS (1995) コンピュータ統計プログラムを使って行われた。

Table 2 Correlation Coefficients

	Photo	Dict	Map	Talk 1	Talk 2	Comic	Illust	Cloze	Meaning
Dictation	.26**								
Map	.37**	.24**							
Talk 1	.22**	.27**	.21*						
Talk 2	-.02	-.05	.07	.03					
Comic	.31**	.29**	.20*	-.04	.05				
Illustration	.28**	.06	.10	.09	-.09	.40**			
Cloze	.24**	.20*	.23**	.12	-.10	.32**	.47**		
Meaning	.28**	.25**	.40**	.22**	-.11	.23**	.36**	.46**	
Passage	.10	.08	.08	.03	.00	.05	-.01	.02	.16

* p<.10, ** p<.05

相関係数

それぞれのサブセクションのピアソン相関係数は表2に示されている。ボード体で書かれている上部の三角形にある係数は聴解力能力に関する相関関係を示す。それぞれのサブセクションが聴解力理解力の隠された特性を測っているのなら、この5つのサブセクションの点数はそれぞれ統計的に有意な相関を示すはずである。つまり収束的妥当性 (convergent validity) があるはずである。表2で示されている係数からは Photograph, Dictation, Map と Talk 1 はお互いに関連しているが、Talk 2 はどのサブセクションとも有意な相関をなしていないことがわかる。結果として Talk 2 は他の聴解力サブセクションで測られている力とは違うものを測定していると推察される。

同様に、ボード体で書かれている下部右にある三角形には読解力サブセクションの相関係数が書かれている。ここでも聴解力の時と同じように、Passage を除いた4つのサブセクションは有意に相関している。

このテストが2つの明らかに違う聴解力と読解力の特性 (trait) を測定しているのであれば、聴解力の5つのサブセクションのそれぞれの相関係数のほうが、聴解力が読解力の5つのサブセクションにもつ相関係数より高くなるはずである。つまり、divergent (分岐的妥当性)、もしくは discriminant validity (弁別的妥当性) を示すはずである。しかしながら表2で明らかなように、サブセクションの多くはこのような結果を示していない。聴解力サブセクションの Photograph と最も高い相関をもつのは聴解力のほかのサブセクション、Map ($r = .37$) であるが、2番目、3番目、4番目に高い相関係数をもつのは読解力サブセクションである。2番目は Comic で $r = .31$ 、3番目、4番目は Illustration $r = .28$ 、と Meaning 同じく $r = .28$ である。聴解力と読解力のサブセクションどうしの高い相関は Dictation と Comic の $r = .29$ や Map と Meaning の $r = .40$ に見られる。

相関係数分析において2つの興味深いことが発見された。1つは聴解力と読解力のサブセクションにおいて共通する何かが測定されたことである。これはこの2つの能力を測るサブセクションは、違いより共通点のほうが多いことを示唆している。もう1つの発見は聴解力サブセクションの Talk 2 と読解力サブセクションの Passage は他の8つのサブセクションとは違うものを測定していることが判明した。テストの Dimensionality をもっと深く研究するために次にこれら10のサブセクションに因子分析をかけることにした。

因子分析

理論的には聴解力の5つのサブセクションは1つの因子にそして読解力の5つのサブセクションはもう1つの違う因子に加せられるべきである。Principle axis factoring program と varimax 回転を使って2つの因子を抽出するようにプログラムを組んだ。表3に結果を報告している。最初の因子を「読解力能力因子」とよぶことができる。読解力の4つのサブセクションがこの因子に重い負荷を置いている。しかしながら、相関分析でも明らかなように、Passage

Table 3 *Principle Factor Analysis of the SLEP Subsections*

Subsections	Factor 1	Factor 2
<i>Listening Section</i>		
Photograph	.268	.501*
Dictation	.099	.496*
Map	.108	.579*
Talk 1	.009	.404*
Talk 2	-.124	.026
<i>Reading Section</i>		
Comic	.440*	.256
Illustration	.798*	.067
Cloze	.593*	.281
Meaning	.419*	.473*
Passage	.000	.177

サブセクションはやはり違った特性 (trait) を測定しているようで、どちらの因子にも加していない。4つの聴解力サブセクションが負荷を置いている2番目の因子を「聴解力能力因子」と名づける。同様に、Talk2はどちらの因子にも加していない。反対に読解力サブセクションのMeaningはどちらの因子にも強く関連している。ここで我々の研究の注意点を指摘しておかなければいけない。因子分析はテスト項目をそれぞれの隠された特性を表す因子に分けるには有効な統計方法であるが、本研究で取り扱った被験者の数が十分でない。因子分析には最低1変数に対し30名の被験者が必要であるが、我々は142人の被験者で10変数を扱っている。したがって統計結果は残念ながら注意して解釈しなければいけない。

項目分析

項目分析の結果は Appendix に載せている。それぞれの項目について困難度とそれぞれの項目がテスト得点の高い者と低い者をどの程度見分けるかを知らせる指標である弁別度を測った。項目の困難度は Item Facility (IF) と呼ばれ、それぞれの項目に正解した被験者の割合を指す。例えば、聴解力サブセクションの Photograph で143人中92名正解していたので Item Facility は .64となる。

弁別度を表す指標として Item discrimination と Point-biserial 相関係数を測定した。Item discrimination (ID) はまず、それぞれのサブセクションで総合得点を計算しその得点によって被験者を上位から3つのグループに分ける。それぞれグループには全体の約3分の1の被験者がいることになる。しかしながら本研究ではサブセクションごとの項目数が少ないことや難易度の差によって被験者を3分の1ずつきれいに分けることはできなかった。Appendix AとBの表にはサブセクションごとに3つのグループの幅と数を書き込んでいる。例えば、Photographでは上位グループ39名、中間グループ67名、下位グループ37名でそれぞれ総合得点15-20点、12-14点、3-11点の被験者が含まれている。また Dictation では上位グループ14-18点の被験者42

名, 中間グループ10-13点60名, 下位グループ2-9点41名となっている。3列目に書いている Item Discrimination は上位グループの Item Facility から下位グループの Item Facility をひいたものである。したがって Item Discrimination 1.0は上位グループの被験者が全員正解したのに対し下位グループの被験者は全員不正解であることを示す。反対に上位グループが全員不正解で, 下位グループの被験者が全員正解すると値は-1.0になる。Item Discrimination .00は上位, 下位のグループの解答に違いがないことを表す。

4列目にはもう1つの項目を分別する指標である Point-biserial 相関係数が示されている。この係数はテスト分析ではよく使われ1つのテスト項目と総合得点を比較するものであり, Item Discrimination と同様, いかによりよい項目がよい被験者とできない被験者を区別するかを表している。

それぞれの項目をその困難度と弁別度から「よい項目」「境界線上にある項目」「不出来な項目」に分類した。「よい項目」は, Item Facility の値が.30から.70の間にあり, Item Discrimination と Point-biserial 係数が.30以上のものとした。これらの項目は Appendix にボード体で記している。「境界線上にある項目」はイタリック体で示しているが, Item Facility の値を.79まで広げ, Item Discrimination と Point-biserial どちらかが.30の値をもつものとした。「不出来な項目」は, Item Facility の値が.80以上と簡単すぎるか, .30と難しすぎるものであり, Item Discrimination また Point-biserial とともに.30以下の分別能力の低いものである。表4はそれぞれのサブセクションの項目がどの分類に属するか, その数と割合を示している。これからもわかるように, それぞれのサブセクションで「不出来な項目」の数が「よい項目」の数を上回っている。全体的に見て聴解力の Dictation と読解力の Meaning が我々の被験者には一番適切

Table 4 Good and Borderline Items

Subsections	Total Items	Good Items	Percent	Good, Border-	
				line Items	Percent
<i>Listening Section</i>					
Photograph	25	6	24%	13	52%
Dictation	18	11	61%	17	94%
Map	12	6	50%	8	67%
Talk 1	14	5	36%	6	43%
Talk 2	6	2	33%	2	33%
Total	75	30	40%	46	61%
<i>Reading Section</i>					
Comic	10	1	10%	4	40%
Illustration	13	3	23%	4	31%
Cloze	30	9	30%	12	40%
Meaning	14	9	64%	10	71%
Passage	8	2	25%	2	25%
Total	75	24	32%	32	43%

なテスト項目を含むサブセクションであり、残りのサブセクションは彼女たちを能力別に分けるには役に立たないことがわかった。

結 論

最初の研究目的は SLEP テストの信頼性に関するものであった。読解力理解問題は全体で KR-20 係数 .82 というかなり信頼度の高いものであったが、聴解力問題の係数は .72 と少し下がった。それぞれのサブセクションの信頼係数に関しては聴解力 Dictation が .77, 読解力 Comic が .87 は高い係数を示したが、他のサブセクションはすべて低い信頼係数にとどまった。Comic でこれほど高い信頼係数をだしたのは、このサブセクションが被験者にとって全体的に簡単であったためだと考えられる。

SLEP テストの困難度を調べるのが 2 番目の研究目的であった。Comic サブセクションのみが被験者にとって簡単すぎ、彼女たちの能力より難易度が低いものであった。Talk1, Talk2, Cloze, Meaning, そして Passage はそれぞれあまりにも難しいものであった。それに反して、Photograph, Dictation, Map, Illustration は単独のテストとみなすと適度な困難度をもつテストであった。

3 番目の研究目的はいかに SLEP テストが能力の高い学習者と低い学習者を区別するのかというものであった。Item Discrimination, Point-biserial 相関係数で測定した結果、聴解力テスト 75 問中 44 問がどちらの係数も .30 以上の値を示し、弁別度が高いものであることを示した。読解力では 75 問中 39 問が弁別度の高いものであった。それぞれのサブセクションの弁別度の高い項目が全体に示す割合は次のようなものである。Photograph では 25 問中 9 問, Dictation では 18 問中 17 問, Map では 12 問中 6 問, Talk 1 では 14 問中 6 問, Talk 2 では 6 問中 6 問, Comic では 10 問中 8 問, Illustration では 13 問中 5 問, Cloze では 30 問中 10 問, Meaning では 14 問中 10 問, Passage では 10 問中 8 問となっている。高い弁別度をもつ項目のいくつかは高い困難度をもつものであったり (Talk2 と Passage の項目など), 反対にあまりに簡単なものもあった (Comic の項目)。したがって適正な困難度と弁別度をもつ項目は聴解力で 30 問, 読解力で 24 問であった。Dictation と Meaning は「よい項目」が高い割合であったのに反し, Photograph, Talk2, Comic, Illustration, Cloze, Passage においてはその割合がかなり低いものとなった。

我々の最後の研究目的は SLEP の 2 つのサブセクションに関する構成概念妥当性 (Construct Validity) についての研究であった。相関関係研究から読解力, 聴解力を測るそれぞれのセクションは 1 つの特性を測っていると予想した。ピアソン相関係数で測定した結果, それぞれの特性を測るサブセクション内では高い相互関係が見つかったが, Talk 2 と Passage はどちらの特性を測るサブセクション群にも相関を示さなかった。このことからこれら 2 つのサ

ブセクションは他とかなり違う能力を測っているのではないかと考えられた。また、聴解力のサブセクションと読解力のサブセクション間の相関も発見され、それはそれぞれのセクション内の相関より高い係数を示した。これらからサブセクションの中には聴解力能力、もしくは読解力能力だけを特別に測定していないものがあることが判明した。因子分析によって SLEP のサブセクションは 2 つの異なる特性を測っていることが明らかになった。Photograph, Dictation, Map, Talk 1 は聴解力因子に過重に負荷し、Comic, Illustration, Cloze は読解力因子に負荷した。しかし、Meaning は両方の因子に負荷し、Talk 2 と Passage はどちらの因子にも関連しなかった。被験者数が少ないためにこの因子分析の結果を解釈するのは注意が必要で、さらに研究を進める必要がある。

結論として、我々の学生である被験者を対象とした英語のクラス分けテストを行うとしたら聴解力のサブセクションである Dictation, そして読解力のサブセクションである Meaning と Cloze を使用することが以上の分析より最適であることが判明した。

文 献

Educational Testing Service. 1998. *Secondary Level English Proficiency Test Order Form*. Princeton NJ: Educational Testing Service.

Appendix

Item Analysis Tables

Table A1

Listening Section Item Statistics

item	IF	ID	Rphi	Group	Options				
					A	B	C	D	blank
<i>Photograph Subsection</i> (Group ranges, n sizes : high 15-20, 39 ; mid 12-14, 67 ; low 3-11, 37)									
P01	.64	.48	.40	high	1	0	2	34*	0
				mid	8	12	4	41*	2
				low	6	11	5	17*	0
P02	.28	-.04	-.01	high	11	14	5	7*	0
				mid	13	24	5	24*	1
				low	12	11	7	9*	0
P03	.47	.26	.19	high	2	14	20*	1	0
				mid	10	18	36*	2	1
				low	8	16	11*	4	0
P04	.83	.15	.20	high	0	2	33*	2	0
				mid	8	2	56*	1	0
				low	7	3	29*	0	0
P05	.51	.18	.23	high	9	5	1	22*	0
				mid	16	11	5	35*	0
				low	5	9	8	16*	1
P06	.38	.31	.21	high	21*	5	10	1	0
				mid	23*	9	30	5	0
				low	10*	7	15	7	0
P07	.54	.26	.24	high	8	24*	5	0	0
				mid	16	38*	9	4	0
				low	10	15*	8	6	0
P08	.79	.23	.34	high	0	33*	3	1	0
				mid	7	54*	4	2	0
				low	5	26*	5	3	0
P09	.76	.12	.19	high	3	6	0	28*	0
				mid	8	9	3	47*	0
				low	5	7	2	25*	0
P10	.72	.43	.44	high	0	33*	3	1	0
				mid	1	52*	11	3	0
				low	2	18*	15	4	0
P11	.46	.42	.30	high	27*	1	3	6	0
				mid	27*	0	24	16	0
				low	12*	4	8	15	0

item	IF	ID	Rphi	Group	Options				
					A	B	C	D	blank
P12	.73	.38	.32	high	34*	2	0	1	0
				mid	50*	8	2	6	1
				low	21*	7	5	6	0
P13	.48	.21	.18	high	3	23*	6	5	0
				mid	10	29*	12	16	0
				low	4	16*	5	14	0
P14	.57	.53	.37	high	1	3	1	32*	0
				mid	11	14	5	37*	0
				low	8	12	6	13*	0
P15	.44	.39	.29	high	3	4	23*	7	0
				mid	4	16	31*	16	0
				low	5	13	9*	12	0
P16	.41	.39	.34	high	1	22*	8	6	0
				mid	8	28*	10	21	0
				low	4	8*	14	13	0
P17	.57	.42	.28	high	1	5	29*	2	0
				mid	5	12	39*	11	0
				low	10	10	14*	5	0
P18	.26	.07	.08	high	1	12*	16	8	0
				mid	1	15*	36	15	0
				low	3	10*	14	11	1
P19	.61	.48	.39	high	31*	2	1	3	0
				mid	42*	16	4	5	0
				low	14*	17	6	2	0
P20	.30	-.06	.04	high	2	22	2	10*	1
				mid	0	39	8	20*	0
				low	4	16	6	13*	0
P21	.50	.10	.16	high	19*	3	7	8	0
				mid	37*	9	11	10	0
				low	16*	4	12	7	0
P22	.72	.59	.53	high	0	1	36*	0	0
				mid	2	8	52*	5	0
				low	2	5	15*	17	0
P23	.24	.01	.10	high	8	16	4	9*	0
				mid	8	23	20	16*	0
				low	4	14	12	9*	0
P24	.58	.46	.37	high	0	3	1	33*	0
				mid	8	16	10	33*	0
				low	4	8	10	17*	0
P25	.08	.03	.09	high	4*	31	1	1	0
				mid	4*	56	4	3	0
				low	3*	30	3	3	0

item	IF	ID	Rphi	Group	Options				
					A	B	C	D	blank
<i>Dictation Subsection</i> (Group ranges, n sizes : high 14-18, 42 ; mid 10-13, 60 ; low 2-9, 41)									
D26	.57	.49	.42	high	0	4	3	35*	0
				mid	3	17	5	33*	2
				low	5	12	9	14*	1
<i>D27</i>	<i>.78</i>	<i>.51</i>	<i>.51</i>	high	0	41*	1	0	0
				mid	1	51*	8	0	0
				low	10	19*	8	0	4
D28	.42	.47	.38	high	0	9	2	31*	0
				mid	1	26	13	18*	2
				low	3	11	16	11*	0
D29	.61	.71	.56	high	0	2	40*	0	0
				mid	3	11	37*	8	1
				low	6	16	10*	7	2
D30	.65	.61	.53	high	1	0	1	40*	0
				mid	13	3	5	39*	0
				low	5	9	12	14*	1
<i>D31</i>	<i>.78</i>	<i>.44</i>	<i>.52</i>	high	0	2	40*	0	0
				mid	2	3	50*	3	2
				low	4	6	21*	7	3
D32	.56	.52	.43	high	0	7	0	35*	0
				mid	7	15	6	32*	0
				low	11	12	4	13*	1
D33	.57	.61	.56	high	35*	5	0	2	0
				mid	38*	13	2	5	2
				low	9*	16	9	7	0
D34	.46	.40	.33	high	0	13	29*	0	0
				mid	1	32	25*	0	2
				low	4	19	12*	6	0
D35	.56	.54	.39	high	0	35*	4	3	0
				mid	0	33*	23	4	0
				low	6	12*	13	7	3
D36	.66	.52	.50	high	5	0	3	34*	0
				mid	8	3	1	48*	0
				low	10	9	9	12*	1
D37	.55	.54	.46	high	30*	7	3	2	0
				mid	42*	13	5	0	0
				low	7*	13	11	8	2
<i>D38</i>	<i>.72</i>	<i>.54</i>	<i>.53</i>	high	0	0	0	42*	0
				mid	7	9	2	42*	0
				low	5	6	10	19*	1
<i>D39</i>	<i>.77</i>	<i>.34</i>	<i>.37</i>	high	0	38*	3	1	0
				mid	2	49*	4	5	0
				low	2	23*	10	5	1

item	IF	ID	Rphi	Group	Options				
					A	B	C	D	blank
D40	.74	.56	.50	high	41*	1	0	0	0
				mid	48*	6	0	4	2
				low	17*	9	9	4	2
D41	.85	.37	.53	high	40*	0	0	2	0
				mid	57*	1	0	2	0
				low	24*	10	2	5	0
D42	.62	.37	.32	high	0	9	33*	0	0
				mid	2	15	38*	5	0
				low	6	11	18*	6	0
D43	.40	.33	.24	high	1	16	1	24*	0
				mid	2	30	5	23*	0
				low	7	19	5	10*	0
<i>Map Subsection</i> (Group ranges, n sizes: high 6-10, 43; mid 4-5, 63; low 0-3, 37)									
M44	.22	.37	.29	high	6	3	15	17*	2
				mid	11	13	21	14*	4
				low	6	16	11	1*	3
M45	.42	.16	.15	high	3	21*	15	4	0
				mid	4	27*	21	7	4
				low	6	12*	12	5	2
M46	.35	.34	.28	high	24*	8	8	2	1
				mid	18*	22	16	3	4
				low	8*	12	12	5	0
M47	.62	.42	.32	high	4	33*	3	3	0
				mid	5	42*	11	4	1
				low	9	13*	5	10	0
M48	.25	.28	.25	high	6	16	3	18*	0
				mid	11	27	12	13*	0
				low	7	15	8	5*	2
M49	.37	.39	.36	high	8	25*	2	8	0
				mid	14	21*	13	15	0
				low	6	7*	10	12	2
M50	.43	.42	.36	high	2	12	26*	3	0
				mid	12	20	28*	3	0
				low	5	22	7*	3	0
M51	.41	.46	.48	high	28*	5	5	5	0
				mid	23*	13	16	10	1
				low	7*	4	18	6	2
M52	.48	.34	.27	high	0	5	25*	13	0
				mid	2	11	34*	15	1
				low	10	4	9*	12	2
M53	.45	.60	.51	high	34*	4	3	2	0
				mid	24*	10	17	12	0
				low	7*	11	13	5	1

item	IF	ID	Rphi	Group	Options				
					A	B	C	D	blank
M54	.47	.45	.38	high	4	8	30*	1	0
				mid	13	14	28*	8	0
				low	11	9	9*	7	1
M55	.25	.24	.27	high	2	16	9	16*	0
				mid	18	19	19	15*	1
				low	10	11	9	5*	2
<i>Talk 1 Subsection</i> (Group ranges, n sizes: high 5-9, 56; mid 4, 36; low 0-3, 51)									
T56	.22	.11	.17	high	14	16	18*	7	1
				mid	11	17	3*	4	1
				low	9	23	11*	7	1
T57	.56	.32	.30	high	7	40*	4	5	0
				mid	6	20*	5	5	0
				low	10	20*	12	9	0
T58	.17	.10	.11	high	24	9	14	9*	0
				mid	12	6	2	12*	4
				low	18	14	15	3*	1
T59	.15	.19	.36	high	9	14	16	15*	2
				mid	8	14	11	2*	1
				low	8	19	18	4*	2
T60	.22	.07	.14	high	7	15*	22	11	1
				mid	4	7*	20	4	1
				low	3	10*	28	8	2
T61	.39	.30	.29	high	5	9	12	30*	0
				mid	5	9	8	14*	0
				low	14	10	13	12*	2
T62	.20	.31	.45	high	22*	10	19	4	1
				mid	2*	13	10	9	2
				low	4*	20	19	7	1
T63	.31	.40	.37	high	3	28*	18	7	0
				mid	7	11*	9	9	0
				low	7	5*	19	17	3
T64	.12	.10	.09	high	3	38	9*	6	0
				mid	10	16	5*	2	3
				low	9	29	3*	9	1
T65	.44	.41	.36	high	35*	5	11	5	0
				mid	17*	6	8	4	1
				low	11*	8	23	8	1
T66	.42	.43	.42	high	4	10	34*	8	0
				mid	2	5	17*	12	0
				low	3	20	9*	18	1
T67	.24	.16	.22	high	11	18*	13	14	0
				mid	10	8*	11	7	0
				low	11	8*	20	12	0

item	<i>IF</i>	<i>ID</i>	<i>Rphi</i>	Group	Options				
					A	B	C	D	blank
T68	.27	.31	.28	high	16	23*	10	7	0
				mid	9	11*	13	3	0
				low	19	5*	19	8	0
T69	.52	.43	.44	high	1	7	43*	5	0
				mid	7	13	15*	1	0
				low	4	15	17*	15	0
<i>Talk 2 Subsection</i> (Group ranges, n sizes: high 3-6 44; mid 1-2, 77; low 0, 22)									
T70	.19	.45	.48	high	20*	6	12	6	0
				mid	7*	30	18	21	1
				low	0*	13	2	6	1
T71	.23	.34	.37	high	5	15*	9	14	1
				mid	27	18*	13	19	0
				low	11	0*	5	5	1
T72	.37	.73	.56	high	32*	1	6	5	0
				mid	21*	11	32	13	0
				low	0*	2	16	3	1
T73	.52	.82	.55	high	1	4	38*	1	0
				mid	15	19	36*	7	0
				low	8	9	0*	5	0
T74	.19	.43	.52	high	11	9	5	19*	0
				mid	21	35	13	8*	0
				low	5	10	6	0*	1
T75	.27	.55	.43	high	6	24*	11	3	0
				mid	16	14*	32	15	0
				low	6	0*	10	5	1

*correct response, bold = well-working item, italic = borderline item

Table A2
Reading Section Item Statistics

item	IF	ID	Rpbi	Group**	Options				
					A	B	C	D	blank
<i>Comic Subsection</i> (Group ranges, n sizes : high 10, 62 ; mid 9, 29 ; low 0-8, 55)									
C01	.71	.71	.73	high	0	0	62*	0	0
				mid	0	0	25*	4	0
				low	6	4	16*	28	1
C02	.94	.16	.57	high	0	62*	0	0	0
				mid	0	29*	0	0	0
				low	2	46*	4	3	0
C03	.64	.73	.54	high	62*	0	0	0	0
				mid	16*	1	1	11	0
				low	15*	11	13	16	0
C04	.90	.25	.58	high	0	0	0	62*	0
				mid	0	1	0	28*	0
				low	6	4	4	41*	0
C05	.75	.62	.80	high	62*	0	0	0	0
				mid	27*	1	0	1	0
				low	21*	3	8	22	1
C06	.76	.55	.73	high	0	0	62*	0	0
				mid	3	0	24*	2	0
				low	11	2	25*	17	0
C07	.86	.38	.76	high	0	62*	0	0	0
				mid	0	29*	0	0	0
				low	5	34*	12	4	0
C08	.82	.45	.79	high	0	0	62*	0	0
				mid	1	0	27*	1	0
				low	6	7	30*	11	1
C09	.86	.35	.54	high	0	62*	0	0	0
				mid	0	27*	1	1	0
				low	7	36*	7	5	0
C10	.84	.42	.80	high	62*	0	0	0	0
				mid	29*	0	0	0	0
				low	32*	3	9	11	0
<i>Illustration Subsection</i> (Group ranges, n sizes : high 11-13, 42 ; mid 9-10, 35 ; low 2-8, 39)									
I11	.95	.15	.42	high	0	0	0	42*	0
				mid	1	0	0	34*	0
				low	3	1	2	33*	0
I12	.90	.23	.44	high	0	42*	0	0	0
				mid	1	34*	0	0	0
				low	5	30*	2	2	0
I13	.91	.21	.48	high	0	0	42*	1	0
				mid	0	1	34*	0	0
				low	1	6	31*	1	0

item	IF	ID	Rphi	Group	Options				
					A	B	C	D	blank
I14	.60	.65	.52	high	4	2	0	36*	0
				mid	9	0	0	26*	0
				low	19	10	2	8*	0
I15	.90	.23	.45	high	0	42*	0	0	0
				mid	2	30*	2	1	0
				low	3	30*	2	4	1
I16	.68	.57	.52	high	2	0	39*	1	0
				mid	3	0	29*	3	0
				low	3	4	14*	18	0
I17	.71	.42	.29	high	0	6	0	36*	0
				mid	0	8	0	27*	0
				low	2	19	1	17*	0
I18	.84	.36	.49	high	0	0	0	42*	0
				mid	1	1	0	33*	0
				low	4	2	8	25*	0
I19	.88	.28	.50	high	41*	1	0	0	0
				mid	32*	2	0	1	0
				low	27*	9	2	1	0
I20	.28	.40	.32	high	2	19*	11	10	0
				mid	1	14*	7	11	2
				low	3	2*	12	22	0
I21	.52	.43	.35	high	1	32*	8	1	0
				mid	3	18*	5	7	2
				low	6	13*	12	8	0
I22	.97	.05	.36	high	0	1	41*	0	0
				mid	0	0	35*	0	0
				low	1	2	36*	0	0
I23	.25	.39	.20	high	0	4	24*	14	0
				mid	2	5	4*	24	0
				low	2	5	7*	25	0
<i>Cloze Subsection</i> (Group ranges, n sizes: high 13-19, 56; mid 10-12, 44; low 1-9, 46)									
C24	.34	.45	.42	high	14	34*	5	3	0
				mid	24	9*	9	2	0
				low	24	7*	10	5	0
C25	.36	.47	.42	high	4	17	5	30*	0
				mid	1	20	4	19*	1
				low	12	25	5	3*	1
C26	.51	.15	.17	high	7	3	33*	13	0
				mid	10	6	21*	7	0
				low	8	7	20*	11	0
C27	.44	.48	.42	high	39*	4	5	8	0
				mid	15*	6	18	5	0
				low	10*	7	9	20	0

item	IF	ID	Rphi	Group	Options				
					A	B	C	D	blank
C28	.27	.18	.21	high	3	5	27	21*	0
				mid	12	3	18	10*	1
				low	10	8	19	9*	0
C29	.32	.15	.16	high	22*	20	6	8	0
				mid	14*	21	4	4	1
				low	11*	15	12	8	0
C30	.36	.17	.14	high	7	14	10	23*	2
				mid	7	10	8	18*	1
				low	14	14	6	11*	1
C31	.29	.22	.27	high	15	22*	5	12	2
				mid	8	13*	5	17	1
				low	8	8*	12	17	1
C32	.12	.00	-.02	high	7	5*	30	12	2
				mid	6	8*	18	11	1
				low	10	4*	19	11	2
C36	.79	.25	.36	high	2	2	48*	4	0
				mid	0	4	39*	1	0
				low	6	4	28*	6	2
C37	.58	.38	.31	high	5	5	4	41*	1
				mid	6	9	1	27*	1
				low	12	10	5	16*	3
C38	.25	.31	.27	high	3	7	24	22*	0
				mid	6	8	19	10*	1
				low	4	5	30	4*	3
C39	.52	.26	.28	high	20	34*	1	1	0
				mid	12	26*	5	1	0
				low	16	16*	4	7	3
C40	.49	.23	.23	high	35*	4	10	7	0
				mid	19*	4	13	8	0
				low	18*	2	8	16	2
C41	.32	.08	.12	high	18*	24	12	2	0
				mid	17*	20	3	3	1
				low	11*	22	9	3	1
C42	.21	.04	.09	high	5	12*	12	27	0
				mid	4	11*	10	18	1
				low	7	8*	14	16	1
C43	.49	.50	.39	high	3	40*	3	10	0
				mid	7	22*	5	9	1
				low	13	10*	8	14	1
C44	.29	.27	.35	high	26*	12	12	6	0
				mid	7*	20	8	8	1
				low	9*	14	12	10	1
C45	.69	.50	.52	high	3	1	0	51*	1

item	IF	ID	Rpbi	Group	Options				
					A	B	C	D	blank
C46	.40	.31	.30	mid	6	3	5	30*	0
				low	4	7	15	19*	1
				high	9	7	31*	8	1
C47	.36	.12	.23	mid	12	6	17*	8	1
				low	10	9	11*	14	2
				high	19	25*	3	7	2
C52	.53	.35	.28	mid	13	13*	4	13	
				low	7	15*	11	12	1
				high	7	39*	6	4	0
C53	.47	.50	.47	mid	5	23*	9	6	1
				low	9	16*	5	9	7
				high	2	40*	1	13	0
C54	.13	.17	.20	mid	2	18*	4	19	1
				low	8	10*	5	17	6
				high	11*	2	23	20	0
C55	.19	.02	-.07	mid	7*	4	14	17	2
				low	1*	8	12	16	9
				high	4	32	11*	9	0
C56	.36	.36	.36	mid	6	18	9*	9	2
				low	10	17	8*	3	8
				high	30*	8	11	5	2
C57	.11	.02	-.01	mid	15*	10	8	8	3
				low	8*	8	12	8	10
				high	5*	10	12	28	1
C58	.71	.39	.30	mid	8*	6	11	17	2
				low	3*	7	10	14	12
				high	1	51*	3	0	1
C59	.14	.11	.15	mid	7	29*	4	2	2
				low	5	24*	4	2	11
				high	19	10	15	10*	2
C60	.23	.32	.25	mid	12	10	12	8*	2
				low	14	6	10	3*	13
				high	26	4	23*	2	1
<i>Meaning Subsection</i>				mid	17	15	7*	3	2
				low	16	7	4*	6	13
				<i>(Group ranges, n sizes : high 7-13, 47 ; mid 5-6, 42 ; low 0-4, 57)</i>					
M33	.23	.25	.35	high	13	14	12	7*	1
				mid	8	12	8	11*	3
				low	14	15	14	13*	1
M34	.55	.31	.47	high	28*	9	3	7	0
				mid	24*	8	8	2	0
				low	29*	11	7	9	1
M35	.41	.52	.51	high	25*	16	3	3	0

item	IF	ID	Rppi	Group	Options				
					A	B	C	D	blank
M48	.44	.49	.50	mid	15*	13	8	4	2
				low	20*	18	8	9	2
				high	3	20	21*	2	1
M49	.57	.47	.45	mid	2	15	18*	7	0
				low	8	19	26*	2	2
				high	5	29*	8	5	0
M50	.37	.34	.45	mid	5	24*	11	2	0
				low	7	30*	14	3	3
				high	11	10	9	16*	1
M51	.43	.43	.41	mid	9	8	9	15*	1
				low	8	13	10	22*	4
				high	6	15	21*	5	0
M61	.48	.65	.55	mid	6	9	20*	6	1
				low	12	9	22*	11	3
				high	5	8	5	23*	6
M62	.41	.52	.43	mid	5	5	5	21*	6
				low	4	10	4	26*	13
				high	21*	11	6	4	5
M63	.46	.68	.50	mid	20*	9	5	4	4
				low	20*	11	8	2	16
				high	9	1	26*	6	5
M64	.32	.31	.23	mid	5	2	22*	7	6
				low	8	9	18*	7	15
				high	6	21*	8	7	5
M65	.17	.10	.05	mid	7	10*	13	5	7
				low	6	16*	11	7	17
				high	8	14	9*	9	7
M66	.23	.45	.38	mid	6	15	7*	8	6
				low	10	9	9*	12	17
				high	3	9	9	17*	9
M67	.26	.26	.23	mid	12	5	9	8*	8
				low	9	12	7	8*	21
				high	5	14*	17	2	9
Passage Subsection				mid	3	10*	12	9	8
				low	2	14*	14	9	18
				Group ranges, n sizes: high 4-7, 26; mid 2-3, 51; low 0-1, 69					
P68	.29	.67	.50	high	0	6	1	19*	0
				mid	5	16	11	19*	0
				low	8	8	4	4*	45
P69	.24	.64	.46	high	2	6	17*	1	0
				mid	8	20	17*	6	0
				low	7	8	1*	5	48
P70	.23	.66	.56	high	18*	1	6	0	1

item	<i>IF</i>	<i>ID</i>	<i>Rphi</i>	Group	Options				
					A	B	C	D	blank
P71	.12	.25	.25	mid	13*	15	14	8	1
				low	2*	6	9	2	50
				high	6	7*	7	5	1
P72	.33	.72	.50	mid	10	10*	22	6	3
				low	4	1*	8	4	52
				high	1	19*	2	3	1
P73	.12	.41	.38	mid	3	28*	9	9	2
				low	5	1*	6	6	51
				high	4	5	5	11*	1
P74	.36	.74	.33	mid	16	13	13	6*	3
				low	4	7	6	1*	51
				high	2	20*	3	1	0
P75	.14	.28	.23	mid	5	31*	7	5	3
				low	6	2*	8	2	51
				high	8*	7	5	3	3
				mid	11*	8	22	8	2
				low	2*	6	6	4	51

*correct response, bold=good item, italic=borderline item