

# Effects of PILT (Program of Integrated Listening Training) on Listening Skills of Advanced Adult Learners

Masako SASAKI (Akita University College of Allied Medical Science)

Naoko FUJITA (Omagari Agricultural Senior High School)

Mitsu TANAKA (Noshiro Minami Junior High School)

## 1. Introduction

Listening skills are described as crucial for survival as well as for access to wider and richer experiences in the foreign environment (Littlewood 1981:65). Whether or not a learner stays in English-speaking countries, almost all serious learners of English in Japan would surely want to improve their listening skills in order to expand their experiences and views through English.

Advanced learners are no exception to this. They are never satisfied with their level of listening skills and would like to develop the skills to a level close to that of native speakers. It is often heard that even advanced learners have difficulties in understanding communication among native speakers. For instance, when advanced learners join a heated discussion among native speakers, they sometimes are at a loss for words and may even feel inferior in the situation because of an unsatisfactory level of listening skills. Lack of listening skills closes the door to full participation in discussions with native speakers. Such experiences may further prompt their aspirations to better understand natural speech, not simplified or contrived speech.

With the aim of enabling advanced learners to better listen to natural speech, a listening development program was designed and named "Program of Integrated Listening Training (PILT)". To let the participants in this program listen to natural speech, the materials were taken from recorded interviews between native speakers or television news programs. The word 'integrated' in the name of the program indicates another characteristic of the PILT in light of approaches and skills. It combined a top-down approach with two bottom-up approaches for the development of listening. Not only listening skills but also writing and speaking skills were used by the participants in the training. Writing a summary and having oral discussion on a related topic or writing a comment about a spoken text were incorporated into the program. Littlewood (1981:75) points out that listening activities are able to provide the input for further language activities.

## 2. Present Study

## 2.1 Purpose

The purpose of this study was to investigate whether the program will work effectively for the development of listening skills of Japanese advanced learners. To measure the effects of PILT, three kinds of listening tests were administered to the participants before and after the program: TOEFL, TOEIC and STEP 1st Grade test. The test materials used were mock tests published by Goken, ALC, and Goken, respectively. In addition, the effectiveness of the PILT was evaluated by the participants.

The following hypotheses were tested:

H1: The PILT will increase the participants' scores on these mock tests: 1) TOEFL, 2) TOEIC, and 3) STEP 1st-grade test

H2: The PILT will be approved by the participants.

## 2.2 Methodology

### 2.2.1 Participants

A total of 6 advanced adult learners (1 male and 5 females) with an average age of 34 participated in this study. 'Advanced level' in this study means the learners scored above 750 on TOEIC or passed the pre-1st grade of STEP. Two participants are in their 40's, three in their 30's and the other in her 20's. Participants are teachers of English except for one, who used to teach English and now is an office worker using English at times at her workplace.

### 2.2.2 Method of Testing Hypotheses

#### A. Materials

##### a. Pre- and Post-Tests

The participants were given three different listening tests: the listening section of a mock TOEFL (a previous version, not a renewal), the mock TOEIC, and the STEP 1st Grade mock test (a previous version, not a renewal).

##### b. PILT Materials

The listening materials used for the program were 22 authentic ones. They included 13 interviews and 9 items of news in either American English (15 materials) or British English (7 materials). The length of the 22 materials ranged from around 1.5 to 3 minutes. The order in which the participants worked on the materials was decided giving consideration to the following multiple factors: 1) topic, 2) speed, 3) length of time, 4) the number of exchanges, and 5) the number of technical terms.

## B. Procedure

### *Pre-Tests*

Before the participants started PILT training, they took the listening section of a mock TOEFL, a mock TOEIC, and a STEP 1st Grade mock test. The three tests were taken in order: TOEFL, TOEIC, and STEP, with a 20-minute intermission between the tests.

### *PILT*

The participants had the PILT training for about 3 months (April 19 to July 11, 1998). The program was implemented in two ways: group training and individual training. The reason for setting up individual training was that the participants had been very busy and it was impossible for them to get together every time. In group training, the participants worked together on the training at a certain place, while in individual training they could work at their own pace at home.

Five group training sessions and seventeen individual training sessions (22 in total) were prepared (Appendix 1). The number of sessions actually implemented in their training, however, differed among the participants: Participant A had 21 sessions (Group (G)-4 + Individual (I)-17); participant B 19 sessions (G-5 + I-14); participant C 20 sessions (G-3 + I-17); participant D 21 sessions (G-5 + I-16); participant E 17 sessions (G-5 + I-12); and participant F 22 sessions (G-5 + I-17). It was impossible to force all the participants to complete all of the training materials during the 3-month training period because they had their personal circumstances.

Each session was made up of 6 steps: 1) 1st listening - listening without any particular task, 2) 2nd listening - writing what the material is about, 3) 3rd listening - note taking following the table which shows each speaker's turn and then answering several questions, 4) 4th listening - summary writing, 5) interpretation into Japanese, and 6) spot dictation (Appendix 2).

**Table 1: Activities of PILT**

Approach		Activity	Times of Listening
Top-down	1.	Summary Writing	4
Bottom-up	2.	Interpretation into Japanese	1 or 2
	3.	Spot Dictation (Rational Cloze Test)	1
(Consolidation)	4.	Discussion[G] / Comment[I]	0

The first activity, summary writing, was adopted as a top-down approach. Beneficial effects of summary writing on listening skills are stated in Okumura (1998), Miura (1997) and Donley (1975). Riley and Lee (1996) remark that summary writing is a useful method for making the subjects read more globally and focus on main ideas more than recall. This research is on reading, but the findings might be applied to listening by analogy. Besides, summary writing

also is given attention as a method for examining how much and how well learners understand what they have listened to. According to Rost (1994), summary writing was effective in identifying the degree of listeners' comprehension of academic lectures.

Interpretation into Japanese followed summary writing as a bottom-up approach (Activity 2 ). Miura (1997) and Torikai (1997) suggest that interpretation training could have positive influence on improvement of listening skills.

As the third activity, spot dictation was incorporated into this program. Spot dictation in the program employed rational cloze, which deletes important parts (Sato 1988:101-2). Sato suggests that taking rational cloze tests that delete key content words would gradually stimulate learners to listen intensively to important parts in terms of understanding a text. He also assumes that regular use of the tests that remove function words would accustom learners to the structure of English sounds naturally and efficiently.

The last activity, Activity 4, was either having discussion among the participants in group training or writing a comment in individual training. It is said that output promotes input. It was assumed that discussion or comment writing (output) would motivate listening (input) that precedes output activities.

#### *Post-Tests and Questionnaires*

Post-tests were administered to the participants to measure the difference of their listening skills before and after the program. The tests employed were the same as the pre-tests. Questionnaires were also employed to examine feedback about the PILT from the participants.

### 2.3 Data Analysis Method

The number of correct answers of each section of each test was counted and regarded as its score, for the marking systems used in the real TOEFL, TOEIC and STEP test were not available for this research. T-tests (one-tailed) were administered to see whether there were some statistically significant differences in the score of each section and the total scores between the pre- and the post-tests. Wilcoxon signed-ranks test was also carried out on the same data because the sample was so small that there was a need for concern that the sample was not drawn from a normally distributed population. As for the questionnaires, binominal tests were administered to see whether or not a difference could be detected in the participants' responses.

## 3. Results and Discussion

### 3.1 Based on Tests

In this section, descriptive and inferential statistics from the three different test scores are shown (Table 2, 3 and 4). As Table 3 and 4 indicate, the total scores on the TOEIC and STEP 1st Grade test between the pre- and the post-tests showed significant differences (respectively  $p < 0.05$  and  $p < 0.1$  by t-test; both  $p < 0.05$  by Wilcoxon signed-ranks test). However, there was no significant difference in the total score on the TOEFL as shown in Table 2. (In TOEFL, Wilcoxon signed -ranks test could not be administered, for the test cannot be carried out on data of 5 or less. The pre-test score of Person C could not be collected, because she was not available at the time of the test.).

**Table 2: Statistics of TOEFL Scores ( $N=5$ )**

	Part A[20] <sup>1</sup>		Part B[15]		Part C[15]		Total[50]	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Mean	15.6	16.0	11.6	12.6	10.0	10.6	37.2	39.2
SD	3.13	1.22	1.52	1.52	1.73	1.67	5.50	3.35
t-test( $t$ )	0.30		1.41		0.65		1.19	

1. [ ]=full marks

**Table 3: Statistics of TOEIC Scores ( $N=6$ )**

	Part I[20] <sup>1</sup>		Part II[30]		Part III[30]		Part IV[20]		Total[100]	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Mean	15.3	14.8	20.8	23.7	19.7	21.2	10.7	13.3	66.5	73.3
SD	0.82	0.98	5.04	4.32	4.18	3.87	3.01	3.01	10.33	9.58
t-test( $t$ )	-1.0		2.49**		2.24**		2.04**		8.2***	
wsrt <sup>2</sup> ( $p$ )	0.28		0.06*		0.08*		0.08*		0.03**	

1. [ ]=full marks 2. Wilcoxon signed-ranks test \*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.1$

**Table 4: Statistics of STEP Scores ( $N=6$ )**

	Preliminary Test[10] <sup>1</sup>						Final Test [30]		Total[40]	
	Part 1[5]		Part 2[5]		Total[10]		Pre	Post	Pre	Post
	Pre	Post	Pre	Post	Pre	Post				
Mean	2.5	3.5	3.8	4.2	6.3	7.7	11.8	15.5	18.2	23.2
SD	2.07	1.64	1.17	1.17	3.14	2.73	8.01	8.38	9.77	10.63
t-test( $t$ )	1.58*		1.58*		1.66*		2.10**		3.23**	
wsrt <sup>2</sup> ( $p$ )	0.10		0.16		0.11		0.07*		0.03**	

1. [ ]=full marks 2. Wilcoxon signed-ranks test \*\* $p < 0.05$ , \* $p < 0.1$

The score of each part of the TOEIC (except for Part I) and the STEP 1st Grade test showed significant differences between the pre- and the post-tests. (The judgement of the Wilcoxon signed-ranks test is more severe than that of the t-test.) However, no significant differences were seen in the score of any part of the TOEFL. The test items on the TOEFL Part B, the TOEIC Part III, and the STEP preliminary test Part 1 are similar. The questions in those parts are based on a conversation between two persons. TOEFL Part C, TOEIC Part IV, STEP preliminary test Part 2, and STEP final test also consist of similar test items: questions based on a passage. Only

the scores from the TOEFL in both cases, however, showed no significant differences in spite of similar test items. Item analysis is needed to clarify this point.

Part I was the only section that showed no significant difference in all the parts of the TOEIC. The test items in this part require participants to choose very quickly a sentence which corresponds with a picture. This kind of listening skill might be least likely fostered by the PILT.

Obviously, the number of participants was insufficient, though they were the total of advanced learners we could ask to participate at the time of this study. To produce more reliable results, more advanced learners should be called upon to be participants in our future research.

### 3.2 Based on Questionnaires

The following are the questions about each activity given to the participants and summary tables showing the responses from them:

#### Activity 1. Summary Writing

- Q1. Was the number of times of listening for summary writing adequate?  
 Q2. Was it useful for comprehension to listen to the other participants' answers in Step 2 (What is this interview/news about?) in group training?  
 Q3. Was the table of Step 3-1 useful for comprehension?  
 Q4. Was the Q & A of Step 3-2 useful for writing a summary?  
 Q5. Was the length of summary (about 1/3 words of a material) adequate?  
 Q6. Was it useful for comprehension to listen to the others' summary in group training?

**Table 5: Questionnaire Result about Summary Writing**

	Yes	No	<i>p</i>
Q1	5	1	0.109
Q2	5	1	0.109
Q3	6	0	0.016**
Q4	6	0	0.016**
Q5	2	4	0.344
Q6	6	0	0.016**

\*\*  $p < 0.05$

There were significant differences in the answers to questions 3, 4 and 6. The pedagogical tools such as functional schema (Q3) and guiding questions (Q4) turned out to be useful. The result of Q6 seems to indicate that it was stimulating to the participants to know how well their rival wrote a summary.

Although Ebbitt (1978) states that the adequate length of a summary is 1/5 to 1/3 of the material, it might be better if the length of the summary (Q5) were reduced to 1/5 from 1/3. It seemed to be the heaviest burden of PILT for some participants to summarize a listening material into one-fifth its length. A one-third to one-fifth decrease would make much difference.

Activity 2. Interpretation into Japanese

- Q1. Was the interpretation training useful for understanding finer points?  
 Q2. Was there any difference between one-time and two-time listening for interpretation?  
 Q3. Did note taking cause any difference in your interpretation?

**Table 6: Questionnaire Result about Interpretation into Japanese**

	Yes	No	<i>p</i>
Q1	6	0	0.016**
Q2	5	1	0.109
Q3	4	2	0.344

\*\*  $p < 0.05$ 

According to the result of Q1, it can be inferred that the interpretation activity was effective in letting the participants focus on finer points in order to understand a text better. Advanced learners seem to have a wish to understand every part of the text like native speakers. Frequency of listening before interpretation (Q2) should be increased in the case of interpreting difficult parts. It might be said that efficient note taking leads to successful interpretation (Q3).

Activity 3. Spot Dictation

- Q1. Was the spot dictation useful for comprehending the text as a whole?  
 Q2. Was the one-time listening for the spot dictation adequate?  
 Q3. Was the number of pauses made in the spot dictation in individual training adequate?

**Table 7: Questionnaire Result about Spot Dictation**

	Yes	No	<i>p</i>
Q1	5	1	0.109
Q2	3	3	0.656
Q3	4	2	0.344

\*\*  $p < 0.05$ 

It could be because a listener might have forgotten what he had heard before writing it down (Koike 1993:57) that one-time listening was not enough (Q2). But one-time listening should be encouraged for learners to foster their concentration and retention.

Activity 4. Discussion (GT). Comment Writing (IT)

- Q1. Did discussion heighten your motivation for listening in group training?  
 Q2. Did comment writing heighten your motivation for listening in individual training?  
 Q3. Was group training more effective than individual training?

**Table 8: Questionnaire Result about Discussion and Comment Writing**

	Yes	No	<i>p</i>
Q1	4	2	0.344
Q2	2	4	0.344
Q3	4	2	0.344

\*\*  $p < 0.05$ 

There were no significant differences in the results of Q1 and Q2. Discussion in group training, however, provided the participants with opportunities to present orally their opinions in front of the others, which might have been more motivating than writing a comment without any

chances to share it with the others in individual training. Their strength of motivation also might have differed depending on a discussion topic. There was no significant tendency in the results of Q3. It shows that each learner seems to have his preferable way of practicing listening skills, so that it cannot be said which way is superior to the other.

#### 4. Conclusion

The PILT seems to have worked effectively on the improvement of the scores for the TOEIC and the STEP 1st Grade test, while not for the scores on the TOEFL, the reasons for which should be investigated by item analysis or in some other ways. Another point to notice is that the mean scores of some parts of the TOEIC and the STEP 1st Grade test indicated significant differences but others did not. The reason for this does not seem so hard to see: it may be that the style of questions which had more rewarding effects than the other types were similar to the PILT training (questions about a passage). The questionnaires revealed that the participants had approved of the PILT, the listening development program designed for the advanced learners.

These are tentative findings about the PILT. The program should be revised based on the questionnaires and future research is necessary to explore further the effects of its refined version.

*(The oral presentation on which this paper was based was made at JACET Tohoku Chapter November Meeting on the 14th of November in 1998.)*

#### Acknowledgements

Our special thanks are due to the Noshiro ESL members, who are truly enthusiastic about improving their English skills, for their sincere cooperation with the research.

#### Works Cited

- Donley, M. 1975. 'Précis Writing: a Rehabilitation.' *ELT Journal* 27, 3: 213-21.
- Ebbitt, Wilma R. and D.R. Ebbitt. 1978. *Writer's Guide and Index to English*(6th ed.). Scott, Foresman. 347-8, 599.
- Lee, J.F. and Riley, G.L. 1996. 'A Comparison of Recall and Summary Protocols as Measures of Second Language Reading Comprehension.' *Language Testing* 13, 2: 173-189.
- Littlewood, W. 1981. *Communicative Language Teaching*. Cambridge: Cambridge University Press.
- Rost, M. 1994. 'On-line Summaries as Representations of Lecture Understanding.' In *Academic Listening : research perspectives*, Flowerdew, J (Eds.), 93-127. Cambridge: Cambridge University Press.
- 石川 達朗. 1984. 「英検 1 級合格ヒアリングカセット」語研:50-52.
- 大友 賢二, G. Stricherz. 1984. 「TOEFL 試験問題-傾向と対策 TOEFL Listening Comprehension」語研: 95-109.)
- 奥村 清彦. 1998. 「要約指導がリスニングとライティングの能力に及ぼす効果」滋賀県立大学国際教育センター研究紀要: 39-51.



木村 恒夫.1993.「TOEIC 模試 600 問」アルク: 14-35.

小池 生夫編. 1993.「英語のヒアリングとその指導」大修館書店

佐藤 史郎. 1988.「クローズテストと英語教育」南雲堂

鳥飼 玖美子. 1997.「英語教育の一環としての通訳訓練」『言語』26, 9: 60-6. 大修館書店

三浦 信孝. 1997.「通訳理論から外国語教授法へ」『言語』26, 9:48-55. 大修館書店

### Appendix 1: Implementation of PILT

#### ◇ Group Training

No.	Date	Material ([I] = Interview, [N] = News, [A] = American E, [B] = British E)
1	April 19	[I/B] BBC 'Health Matter--'Natural Can Be Unhealthy' from English Journal. 1998, May
2	April 29	[N/A] CNN 'Appearance Effects from English Express
3	May 17	[N/A] ABC World News Tonight. 'Vitamin from Food Is Not Enough' April 7, 1998
4	June 7	[I/A] 'EE Talk--Opening the Way to Closing Deals in Real Estate'Part 1 from E.E. 1991,May
5	June 28	[I/B] BBC 'Health Matter--Stress in the Workplace' from English Journal.. 1998 June

#### ◇ Individual Training

No.	Date	Material ([I] = Interview, [N] = News, [A] = American E, [B] = British E)
1	April 23 - April 29	[N/A] ABC World News Tonight. 'School Violence' April, 1998
2	"	[N/A] ABC World News Tonight. 'Hollywood' April, 1998
3	April 30 - May 6	[N/A] CNN 'A Pinch of This, a Dash of That' from English Express
4	"	[N/A] FEN 'Hiroshima' from Hearing Marathon. 1985, November
5	May 7 - May 13	[I/B] BBC 'Health Matter-Alternatives for Osteoarthritis' from English Journal.1998 May
6	"	[I/B] BBC 'Health Matter--A Fruit-Flavored Vaccine' from English Journal.1998 May
7	May 17 - May 23	[N/A] CNN 'Homeless' from English Express
8	"	[N/A] CNN 'Korean American' from English Express
9	May 24 - May 30	[I/A] 'EE Talk--USA TODAY Gains Popularity Worldwide'Part 1 from E.E.. 1990.Aug.
10	"	[I/A] " Part 2 "
11	May 31 - June 6	[I/A] " Part 3 "
12	"	[I/B] BBC 'Health Matter-Below-the-Bra Guide for Girls' from English Journal.1998 June
13	June 7 - June 13	[I/A]'EE Talk-Opening the Way to Closing Deals in Real Estate'Part 2 from E.E. 1991,May
14	"	[I/A] " Part 3 "
15	June 21 - June 27	[I/B] BBC'Health Matter--Malnutrition - A Major Killer ' from English Journal.1998 June
16	July 1 - July 11	[I/B]BBC'Health Matter--Depression:A Global Issue' from English Journal.1998 June
17	"	[N/A] CNN 'Environmental Compatibility Goods' from English Express. '

## Appendix 2: Details of PILT

No.1

**PILT**

Program of Interpretation and Listening Training

⑮ June 21- 27  
Tape -- Side A 前半

◆ Activities ◆

1. Summarisation of the content of news  
2. Interpretation into Japanese  
3. Spot Dictation  
4. Comment

Name \_\_\_\_\_

Date when you tried: \_\_\_\_\_

No.5

◆Activity 2.  
Interpretation into Japanese

1回聞いただけで、 通訳して下さい、リスニングの際の集中力を高めるための、メモを取る必要性を感じる方はとって構いません。

No.2

◆Activity 1.  
Summarization  
いくつかの Step を経て、最終的に Summary を 120 words 程度で書いてください。  
Listening の機会 は 4 回あります。

§ Step 1.

§ Step 2. What is the interview about?

No.4

◆Activity 3.  
Spot Dictation  
Fill in the blanks by listening once.  
Take a pause at ※.

Reporter:  
The State of the World's Children is a report published each year by UNICEF, the United Nations Children's Fund. The 1998 edition has some shocking statistics. Here are just two: over 200 million children under five years old in developing countries are malnourished. Of nearly 12 million children under five who die each year in developing countries, the deaths of over six million are from malnutrition.

※:

Dan Toole, previously UNICEF's representative in Mali and Rwanda, and now in their headquarters in New York, was in London for the report's launch. I asked him how malnutrition kills so many small children.

Tools:  
There's a small percentage, those you see on the television in Somalia, in Burundi, Bosnia, who are malnourished. Most die from causes like respiratory infections and diarrhoea. Malnutrition weakens their immune systems, so they get measles and die, they get other sorts of illnesses that then are complicated by diarrhoea. They are weakened already by malnutrition.

Reporter:  
What are the causes of this malnutrition?

No.3

§ Step 3-1. Please take notes for each section.

SPEAKER	Notes for Content
Reporter	
Dan Toole	
Reporter	
Dan Toole	
Reporter	
Dan Toole	
Reporter	
Dan Toole	
Reporter	
Dan Toole	

No.7

1. Check your answers.

1. malnourished  
2. linked with  
3. all  
4. die directly of  
5. related to  
6. the fact that  
7. live without it  
8. supplies  
9. malnutrition  
10. water  
11. what have come out  
12. what we call cause  
13. increased  
14. have found  
15. poverty  
16. supplementation  
17. with  
18. lower doses  
19. immune response

2. Please listen for the final time.

◆Activity 4. Comment  
Please write your comment or feelings on the news briefly.

No.4

§ Step 3-2. Please answer questions.

[Q.] Does malnutrition directly kill children in most cases of the deaths of over six million?

[Q.] There are new things that have come out as the causes of malnutrition. What is said to have a lot to do with malnutrition?

[Q.] Has malnutrition been found in developed countries?

[Q.] What is referred to as an example to solve the problem of malnutrition?

§ Step 4. Summary (about 120 words)

( words)

No.8

◆Answer Key  
Activity 1 Step 3-2

Ans 1 No, it doesn't. It is causes related to malnutrition that kill children in most of the cases.  
Ans 2 The time of a mother to feed a child is. (The capacity of a mother is.)  
Ans 3 Yes, it has  
Ans 4 Supplementation of some kinds of foods is

Activity 2 Interpretation into Japanese

1) 1300 万近くの 5 歳以下の子供が毎年貧血と脱水で死亡するのですが、その中で 6 万人以上の死が栄養失調と関連があります。

2) 毎年貧血と脱水で死亡する 1300 万近くの 5 歳以下の子供のうち、600 万人以上の死が、栄養失調と結び付けられています。

3) それから、増えているのが、新たに明らかになった事柄の中に、まさに、いわゆる(子どもの)言語の重要性があります。母親の愛情が、つまり、母親が子どもに愛情を注ぐことが重要なのです。

4) ええ、実際イギリスやアメリカであっても、栄養失調は貧血と結びついていますから、貧血のあるところほどであるよ、貧血を治す薬や水を子供に入れる方法を欠いており、栄養失調も存在するといわけですね。

5) 多かれ少なかれ、小さなカプセルがあつて、それを服用した後に子どもに投与することができます。やや少量を経腸中の女性に投与することもできます。

6) 多かれ少なかれ、生まれた後の子どもに投与できるような、小さなカプセルもあります。女性たちが経腸しているときに、比較的少量(のピルなど)を投与する方が好まれます。

## Appendix 3: Test Scores of the Participants

## Scores of TOEFL (N = 5)

	Part A		Part B		Part C		Total	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post
A	15	16	12	15	10	12	37	43
B	19	18	14	13	13	11	46	42
D	18	15	10	12	9	12	37	39
E	15	15	11	12	9	10	35	37
F	11	16	11	11	9	8	31	35
Mean	15.6	16.0	11.6	12.6	10.0	10.6	37.2	39.2
SD	3.13	1.22	1.52	1.52	1.73	1.67	5.50	3.35

## Scores of TOEIC (N = 6)

	Part I		Part II		Part III		Part IV		Total	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
A	15	16	28	29	25	24	10	13	78	82
B	15	13	22	28	21	25	15	14	73	82
C	15	15	25	24	20	22	13	17	73	78
D	15	15	17	23	22	23	11	11	65	72
E	17	15	18	20	17	18	8	16	60	69
F	15	15	15	18	13	15	7	9	50	57
Mean	15.3	14.8	20.8	23.7	19.7	21.2	10.7	13.3	66.5	73.3
SD	0.82	0.98	5.04	4.32	4.18	3.87	3.01	3.01	10.33	9.58

## Scores of STEP 1st-Grade (N = 6)

	Preliminary Test						Final Test		Total	
	Part 1		Part 2		Total		Pre	Post	Pre	Post
	Pre	Post	Pre	Post	Pre	Post				
A	5	5	5	5	10	10	12	17	22	27
B	4	4	4	4	8	8	14	25	22	33
C	4	4	5	5	9	9	16	17	25	26
D	1	5	4	5	5	10	23	23	28	33
E	0	1	2	2	2	3	6	6	8	9
F	1	2	3	4	4	6	0	5	4	11
Mean	2.5	3.5	3.8	4.2	6.3	7.7	11.8	15.5	18.2	23.2
SD	2.07	1.64	1.17	1.17	3.14	2.73	8.01	8.38	9.77	10.63