| Title | SLEP テストを ECA カリキュラムに組み入れるための分析 |
|---|---|
| Author(s) | ヒル・ケント，サベット・メヘラン |
| Citation | 聖学院大学論叢, 22(2)： 243-273 |
| URL | http://serve.seigakuin-univ.ac.jp/reps/modules/xoonips/detail.php?item_id=1938 |
| Rights | |

〈研究ノート〉

# Analyzing and Integrating the SLEP into the ECA Curriculum

Kent HILL・Mehran SABET

SLEP テストを ECA のカリキュラムに組み入れるための分析

ヒル　ケント\*・サベット　メヘラン\*\*

　この論文は，新入生のクラス編成と学期の事前・事後テストに使用する SLEP テストのリスニングと読解部分を English Communication Arts のカリキュラムに組み入れるための試みについて報告するものである。また，SLEP テストの詳しい分析と前年度の学生のテスト結果の項目分析も報告する。カリキュラムに組み入れるために作成した SLEP テスト教材は，信頼が置けるものであり，日本人学生の考え方に配慮して作成されたものである。SLEP テストのリスニングテストに基づいて作成されたスピーキング・アセスメントとリスニングタスクは，ダイナミックアセスメント（ヒル＆サベット，2009 年）の論理に基づき，同時に実施した。SLEP テスト模擬試験の実施前に，テストを受けるためのストラテジーが学習者に与えられた。テスト終了後，SLEP テストをカリキュラムに組み入れることによるプラス効果があるかどうかを確認するために，学生にアンケートを実施した。

　　Although language proficiency is the target of second-language learning programs, proficiency assessment through the use of standardized tests has not always been well integrated into the language-learning process（McNamara, 2005）. Even where curriculum reform involves attention to the assessment, language assessment is still carried out to serve the needs of institutional purposes. Standardized tests have typically focused on the measurement of individual knowledge and skill, and theories of proficiency have been limited in a social sense, even when they address issues of communicative competence. In the last decade or more, however, a critical voice has grown that calls for a "social turn" away from the individualistic focus in conceptualizations of

language proficiency (McNamara, 2005).

Thus, there has been a call for a reformulation of competencies needed for conducting assessment in an educational context. Inbar-Lourie (2000: 386) states, "a reformulation and the potential tensions which impact decision-making in assessment are tensions between formative and summative assessment, between criterion and norm-referenced approaches, between traditional and alternative assessment formats and between externalized standardized testing versus classroom tests." Inbar-Lourie (2000) adds that language assessors need to be familiar with contemporary theories about the learning, teaching and assessment of grammar so as to be able to design suitable assessment measures. Functioning within this framework requires not merely familiarity with its different levels and descriptors, but with language teaching skills and assessment know-how. It also assumes that language assessors, who are in many cases teachers, can adapt this framework to their teaching and assessment purposes, and thus function at a more advanced and specialized level.

The ability to integrate standardized assessment into the curriculum requires: (a) a sound understanding of key concepts in second language assessment, (b) the ability to critically evaluate existing assessments, and (c) the capacity to design or adapt assessment instruments for particular teaching contexts (Inbar-Lourie, 2000). The study reported here attempted to dynamically integrate standardized testing (i.e., the Secondary Language English Proficiency test) into Seigakuin University's English Communication Arts curriculum. However, different teachers share different willingness and capacity to embrace new ideas when attempting to dynamically integrate the standardized test into the curriculum. These differences can be attributed to different professional and personal backgrounds and experiences.

Another ethical concern with integration is the effect of a test or assessment framework on pedagogy, in other words, washback. Washback involves the positive or negative effects felt by learners in this case as a result of integrating standardized testing into the curriculum. Washback can be beneficial or detrimental to learners' development. If teachers teach to the test or assessment requirements and the consequence is narrowing of the curriculum, the effect is educationally undesirable. On the other hand, if a particular testing or assessment requirement leads to teaching practices that promote and broaden learning, the effect is positive (Leung & Lewkowicz, 2006). Washback reveals how a concern for values and social consequences opens up new questions, many of which are deeply and directly connected to language pedagogy and curriculum design. Additionally, if evidence of positive washback is found, it may also provide more information on how to further integrate assessment into the curriculum.

This all suggests that work in testing and assessment is reciprocally relevant to language pedagogy and curriculum development (Leung & Lewkowicz, 2006). Also of relevance, Chan et al. (2000) recently linked a dynamic assessment approach, which has often previously been used as a content free or content independent procedure, to the curriculum. Their contention is that curriculum tasks involve the interdependence of knowledge and cognitive processes. In the Chan et al. (2000) study, linking dynamic assessment to the curriculum allowed the investigators to identify the cognitive processes of the students and to assess the extent to which they could benefit from further instruction.

According to Lidz (2000), assessors need to go beyond what learners currently know to understand how they learn, as well as to determine any obstructions to their successful learning. Combining curriculum-based and process-based approaches within a dynamic assessment model has been proposed by Lidz (2000) as the optimal approach to assessment. Addressing cognitive processes that underlie specific curriculum objectives allows assessors to probe more deeply into the nature of learning approaches. Remaining close to the curriculum increases the relevance of the resulting recommendations for the instructional setting. The appropriateness of addressing underlying processes as a means of both understanding the learner's cognitive functioning and proposing interventions relevant to the curriculum has been documented by a number of studies (see Lidz, 2000). The materials used in the intervention should be different from those used for the pre/posttests. The interventions teach to the processes, principles, and strategies underlying the tasks that relate to successful performance without utilizing the actual test items themselves. Lidz (2000) provides these rating scale components for the interaction found within an assessment-based curriculum:

Self Regulation: Learner maintains attention and refrains from impulsive interaction with materials.

Persistence: Learner completes task without seeking to terminate prematurely.

Tolerance: When experiencing frustration, the learner is readily calmed and redirected.

Motivation: Learner shows enthusiastic reaction to the materials and task.

Flexibility: Learner does not repeat unsuccessful task solutions and develops alternative approaches to the task.

Interactivity: Learner engages in turn-taking conversational exchanges with some degree of elaboration.

Responsivity: Learner is a willing learner and open to input from assessor.

At Seigakuin University, all first-year students take the Secondary Language English Proficiency (SLEP) test in order to be placed in appropriate Reading and Speaking classes. Students are streamed into five levels: Super A, A, B, C, and D. The university has six departments: Political Science and Economics (P), Local Community Policy (L), Euro-American Studies (A), Japanese Literature (J), Child Studies (C), and Human Welfare (W). All first-year students are required to take English classes; however, which classes they take depends on their department. For example, A Department students are required to take two Speaking, one Reading, and one Listening (Cinema) class per week. However, P and L students take the Speaking classes only for both semesters, while J, C, and W students take one Speaking class in one semester and one Cinema and one Reading class in another semester.

The SLEP test is then administered again at the end of the academic year in order to measure learners' English proficiency upon completion of the English Communication Arts (ECA) program as well as to place learners in appropriate levels for second-year courses. The test measures English language ability in two primary areas: understanding spoken English and understanding written English (ETS, 2004). Additionally, though not the primary reason for giving the test, the results are used to promote the ECA program in order to recruit incoming students.

The SLEP test is designed for use with students whose native language is other than English but are entering grades seven through twelve or community colleges in the U. S. or Canada (ETS, 2004). As for the appropriateness of giving the SLEP as a placement test, Hughes (1989) states that no one placement test will work for every institution, and the initial assumption about any test that is commercially available must be that it will not work well. Brown (1996) separates proficiency and placement tests and says that proficiency tests are very general and are designed to assess wide bands of abilities. In contrast, a placement test must be more specifically related to a given program, particularly in terms of the relatively narrow range of abilities assessed and the content of the curriculum, so that it efficiently separates the students into level groupings within the program.

According to ETS (2004), the basic assumption underlying the SLEP test is that language ability is a critical factor in determining the degree to which students can benefit from instruction in English and their success depends on their ability to understand what is being said by teachers and fellow students, and to understand formal and informal materials written in English. The test is intended for learners in an English as a second language environment not, as is the case here, English as a foreign language; however, ECA classes are conducted mainly in English, and therefore the assumption underlying the SLEP can assist instructor/assessors to understand how

learner/test-takers may perform within a second language context, although, the ECA classes most likely have a less formal focus.  Although the test does not directly measure speaking ability, Hughes (1989) argues that if a test has enough items, it can measure students' overall ability. Gorsuch (1995) adds that the testing packages produced by more reputable companies include ample evidence supporting the reliability of their tests.  The SLEP test is produced by Educational Testing Services (ETS), which also designs the TOEFL test, and hence the SLEP test score also has an equivalent TOEFL score.

A more practical or appropriate test would be a locally developed placement test; however, as Weir (1993) and Inbar-Lourie (2000) make clear, developing such a test requires knowledge of testing principles and demands team cooperation.  This team cooperation involves a group of teachers who can create test-item specifications and item banks, pilot the test, analyze the test items, and then revise them in order to improve the test (Brown, 1996).  This same process must be repeated until desirable results are attained; unfortunately, however, many schools do not have the resources to perform such a task and as a result adopt a standardized test such as the SLEP, TOEFL or TOEIC.  Thus, in order to have the norm-referenced data provided by a standardized test, rather than develop our own test, as Leung & Lewkowicz (2006) suggest, a more effective approach was to attempt to integrate the SLEP test into the existing ECA curriculum.  This process of integration is explained in the following sections: (1) the ECA program, (2) the SLEP test and its scoring system, (3) analysis of the SLEP test, (4) dynamic assessments, (5) integration of the SLEP test, (6) test-taking strategies, (7) the practice SLEP test, and, (8) descriptors for streamed levels.

## The English Communication Arts Program

In the ECA program, the Speaking courses are taught by native or near-native English speakers and the Reading and Cinema (i.e., listening) courses are taught by Japanese instructors. Along with the SLEP pre/post-tests, students in the Speaking classes are given a program specific pre/post test which can be classified as a criterion-reference test based on the curriculum for each level.  Two speaking assessments per semester are also administered.  Figure 1 shows the overall

Figure 1. Overall annual SLEP gains

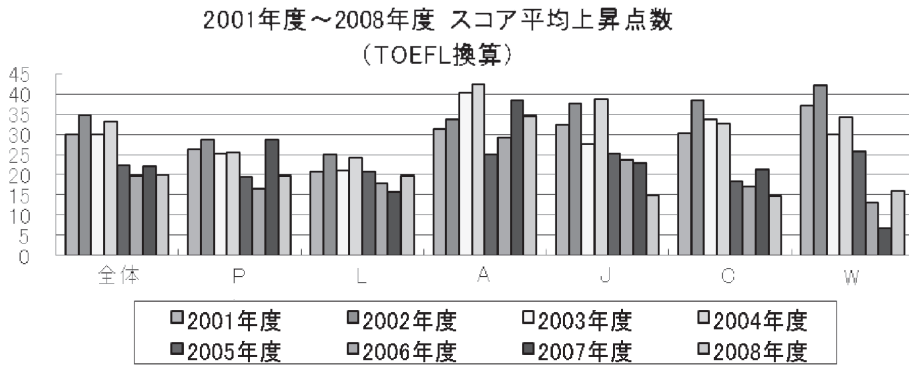|      | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 |
|------|------|------|------|------|------|------|------|------|
| 全体 | 30.0 | 34.8 | 30.1 | 33.2 | 22.3 | 19.7 | 22.0 | 20.0 |

Figure 2. Annual SLEP gains by department

annual SLEP gains for the ECA program.

Figure 2 shows the different annual SLEP gains made for each department. Some departments reduced the number of required English courses in 2005.

## The SLEP Test and Its Scoring System

The SLEP is a norm-referenced test. Assessors can therefore compare students' results with those of other students in similar situations. ETS (2004) chose the multiple-choice format of the test primarily to ensure reliability of results through standardization of administrations and to eliminate reliance on the subjective judgments of raters. The test is divided into two sections, each containing four types of questions. For the first section, i.e., the Listening Comprehension section, four types of questions all use recorded samples of spoken English, which do not rely heavily on written material to test listening comprehension. The four question types in the second section, i.e., the Reading Comprehension section, are based on written or visual materials. Section 2 includes written questions based on a cartoon, written questions based on line drawings, and a literary passage followed by questions on its content. This section also measures vocabulary and grammar (ETS, 2004).

ETS (2004) based the choice of material for the SLEP test on an analysis of actual materials designed for use in American high school classrooms. In the case of some questions, particularly the short and extended dialogs in the Listening Comprehension section, ETS used situations representative of those encountered by students in American secondary schools. Therefore, some questions are based on conversations that take place in various parts of a school and deal with events that occur in each location. Conversations also deal with extracurricular activities and

academic subjects. These subjects can be quite unfamiliar to most learners in English as a foreign language contexts and this is why the materials developed for integration of the test into the curriculum were developed to be more authentic to the learners' context.

Scoring of the SLEP test is based on raw and scaled scores. A raw score is the number of correct questions answered by a learner. Using a conversion chart, a raw score is then converted to a scaled score. The lowest scaled score is 10 points. This means that depending on the form of the test being used, any raw score between 0 and 25 basically gets a scaled score of 10. The scale also varies between the Listening and Reading sections. This means that equal number of correct raw answers in each section will not receive the same scaled scores. The SLEP test has three newer forms and each has been calibrated for difficulty, which means that no matter which form is used learners should have an equivalent result. The result of this scoring system is that there is a cut-off point that learners must score above before increasing their scores. Therefore, learners must score significantly high in the first two parts of the Listening and Reading comprehension sections in order to begin to increase their final score. Then any actual gains occur in the last two parts for both sections of the test. Although some gains can also be made in the first two parts, it is the gains in the last two sections that actually determine learners' scores.

The results of the item analysis of the 2008 posttest for the Listening section support this view. The gains made for all learners were: Part 1: 1.1, Part 2: 2.2, Part 3: 0.05, and Part 4: 0.08. In Parts 3 and 4, learners made minimal gains. After doing an earlier item analysis of the SLEP test, Culligan and Gorsuch (1999) also found that learners showed minimal gains in these parts. They claimed there were a couple of reasons for this: (1) learners already had fairly high scores on the first two parts of the Reading section and there was little room for progress; and the last two sections of the Listening section were too difficult or culturally inappropriate for learners. This once again suggests that more authentic materials developed for integration of the SLEP test into the curriculum may offer a possible solution to overcome these difficulties.

## Analysis of the SLEP Test

The first part of the Listening section includes pictures in the test book that learners must look at and then listen for the correct sentence to match to the photo. In Part 2, test-takers listen to a sentence, read four sentences in the test book, and mark the sentence that is exactly the same as they heard on the tape. In Part 3, learners read a *Wh*-question with multiple choice answers and then answer it after listening to a short dialog. In Part 4, test-takers mainly respond to questions

they hear regarding the direct speech found within extended dialogs.

In the first part of the Reading Comprehension section, learners read sentences and match them with the accompanying cartoon mind bubble. In Part 2, they look at four pictures and read a complex sentence. Then they match the appropriate picture to the sentence. In the third part, they read a passage and choose grammatically correct phrases to connect the passage. In the last part, comprehension questions are asked about the passage and, finally, they are asked to infer the general meaning of the passage.

## Analysis of the Listening Section

### *Part 1: Look at the Photo*

The choice of answers in the first part can be divided into two types: the first contains sentences with all the subjects staying the same, using present progressive aspect, and no difficult vocabulary. Here is an example:



A．The boy is standing on the left.

B．The boy is not laughing.

C．The boy is sitting in the classroom.

D．The boy is standing between the two girls.

In the second type, there is not any consistency between sentences in that all subjects may be different, different verbs may be used, and new vocabulary is introduced. Here is an example:



A．Drinks and cakes are on the table.

B．The desserts are half eaten.

C．There are a variety of desserts on the table.

E．People have gathered around the table.

Item analysis showed variability in response to these two kinds of questions. Test-takers were

much more likely to answer the first type of question correctly than the second.  Therefore, integration into the curriculum suggested encouraging learners to ask:  *Who is in the photo?*, *What are they doing?*, *What are they doing it to?*, and *Where are they?*

### Part 2: Same Sentence as You Hear

In Part 2 of the Listening section, test-takers must listen to a statement, read four sentences, and mark the exact sentence they heard on their answer sheets.  Item analysis showed that lower-level students were scoring higher on this than higher-level students and this result may have been because lower-level students were already receiving practice in the curriculum on this type of task prior to the posttest.  Additionally, this increase in pre/post-test scores indicates that this part might be an area where significant gains can be made by test-takers.  If they are aware that the semantic content of the sentences is not of importance and they need to focus on the word order of the sentence, then scores may improve with practice alone.  Here is a sample question:

On tape:

Once it's cold the plants won't grow any taller.

In the test book:

A．By the summer those plants will be much taller.

B．Those tall plants should be cut back when it's warm.

C．Once it's cold the plants won't grow any taller.

D．The taller plants keep growing all summer long.

### Part 3: Short Dialogs

In this part, students listen to short dialogs and comprehension questions and then select answers written in the test books.  The short dialogs are generally scripted as speech events, many of which are very similar to those already found in the ECA curriculum.  The speech events in the SLEP test short dialogs include:

- ・Arranging a meeting place or time.
- ・Planning a schedule.
- ・Understanding instructions.
- ・Understanding transitions.

・Giving directions.

・Talking about likes and dislikes.

・Talking about health or experiences.

・Talking about future plans.

Although the types of dialogs and vocabulary vary from test form to test form, the speech events still quite consistently matched the ECA curriculum.　Thus, providing students with more contextually authentic curriculum-based dialogs may prove to be multi-functional and beneficial to assessment.　Additionally, there is an abundance of future tense-aspect in this part.　Thus, as Inbar-Lourie (2000) pointed out, an awareness of grammatical form should be considered when attempting to integrate standardized testing into the curriculum.　Here is a sample dialog from Part 3 (although it does not feature future tense-aspect):

On tape:

Q：What didn't the boy like about his visit to the aquarium?

A：What did you do this weekend?

B：We went to the aquarium.　You should go.　It was great 'cause no matter where you stand you get a good view of everything.　Plus they have a really big range of fish and they have different shows you can go to.　The only thing was we were too late to see any of the shows. You have to get there early to get tickets for them.

Q：What didn't the boy like about his visit to the aquarium?

In the test book:

Q．What didn't the boy like about his visit to the aquarium?

A．He had trouble seeing the fish.

B．He had to stand in line a long time to get in.

C．He wanted to see more different kinds of fish.

D．He wasn't able to see the show.

After the listening, a comprehension question is asked on the recording about the dialog, which learners can also read.　At this point in the test, the test-taker is being assessed for their ability to

comprehend the: *who*, *what*, *when*, *where*, and *why* of the speech event.  Although each form of the test varies in vocabulary, this vocabulary is generally high school specific.  Therefore, worksheets were developed to integrate this vocabulary into the curriculum.

*Part 4: Extended Dialogs*

In Part 4 of the listening section, students listen to extended dialogs and answer approximately 17 questions pertaining to the one dialog.  The extended dialog is divided into sections and in most cases for each section test-takers must answer two questions.  The questions are usually related to the direct speech found in the dialog and learners hear the questions but cannot read them.  Here is a sample section and the questions which follow:

On tape:

A：Hi, Luisa.

B：Oh, Sam. Hi.  I missed you this morning in history class.  Where were you?

A：That's what I wanted to talk to you about.  You see, my homeroom teacher had to pack up a bunch of science materials from our unit on electricity, so she asked me to stick around and help her.

B：Did Mr. Jackson say it was okay to miss his class?

A：Yeah, my homeroom teacher called him and asked if he minded.  He said it was okay with him if it was okay with me.

B：Well, all we did was take notes about the Industrial Revolution.

A：I know.  That's why I was looking for you.  I wanted to see if I could look over your notes tonight.  I can bring them back tomorrow.

Q 1.  Why did Sam miss history class?

Q 2.  Why was Sam looking for Luisa?

In the test book:

A．He was sick.

B．He missed the bus.

C．<u>He was helping a teacher.</u>

D．He was doing a science experiment.

A．To give her a note.

B．To return her science book.

C．To ask to borrow her notes.

D．To see if Mr. Jackson was upset.


As the sample shows, the listener must have some understanding of the intentions and direct speech of the participants in the extended dialogs (other examples: *What does she suggest to do?* or *Where does she say she will see him?*).

## Analysis of the Reading Section

### *Part 1: Keywords*

The first part of the reading section begins with 12 questions where students must read a statement, look at the mind bubble cartoons, and choose which picture the statement is referring to. Here is an example:

Q：I hope we stay at a campground where I can go horseback riding. D
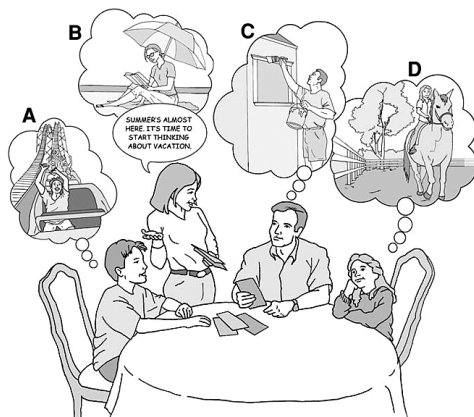


In this part, an awareness of the keywords in the sentence plays a prominent role in understanding which mind bubble refers to the sentence in question.


### *Part 2: Complex Sentences*

Part two also involves pictures. Test-takers look at four pictures, read a statement underneath

the pictures, and decide which picture it refers to.  Here is a sample:



Q．One bird is sitting in a tree but two aren't.

This part builds on and, therefore, is more difficult than the first part because learners need to comprehend something fro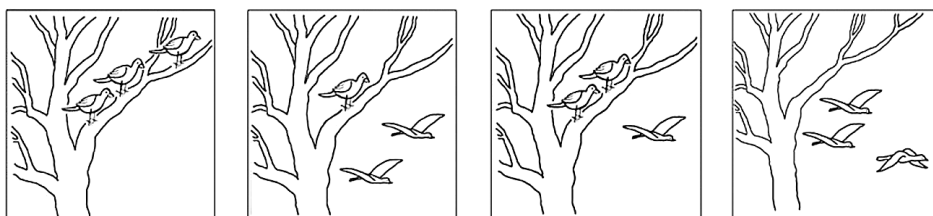m the first clause of the sentence (i.e., *one bird is sitting in a tree*) and relate it to something in the second clause of the sentence (i.e., *but two aren't*).  Thus, this part requires an understanding of complex sentences.

### Part 3 & 4: Transition Phrases and Making an Inference

The third part also builds on the previous parts in that it uses grammar to show how complex sentences are joined with transition phrases, adverbs, conjunctions, and verb tenses to form a passage.  If test-takers can spot the correct grammatical form in the boxes, similar to finding keywords, then it might be possible to answer the comprehension questions about the passage.  Here is a sample:

$$\text{Most animals} \left|\begin{array}{c} \text{live} \\ \text{they live} \\ \text{are living} \\ \underline{\text{that live}} \end{array}\right| \text{in the desert do not take in water from open sources.}$$

Part 4 mainly culminates in making an inference about the passage by building on the ability to link complex sentences and answer the comprehension questions about the passage.  As with the listening comprehension section, each part of the reading assessment builds on the other in a bottom-up fashion (i.e., key words, complex sentences, comprehension questions and making an inference).  Making an inference is the most challenging part of the test.  Here is a sample:

Example text:

*On the old Paul Jones, I got a chance to get acquainted with one of the steamboat pilots.  He taught me how to steer the boat, and thus made the fascination of river life more potent for me than ever.*

Q．What can be inferred about the author before he first went on the *Paul Jones?*

A．He had no money at all.

B．He had worked as a pilot on other steamboats.

C．He had participated in the siege of the city.

D．He had a strong interest in working on the river.

## Summary of the SLEP Analysis

This analysis of the SLEP test presented the following as the focus for attempting to integrate the test into the curriculum:

### *Listening*

・The ability to identify subjects, activities, objects and locations in sentences.

・The ability to focus on word order rather than the meaning of the sentence.

・The ability to comprehend speech events.

・The ability to understand direct speech in extended dialogs.

### *Reading:*

・The ability to focus on keywords.

・The ability to connect complex sentences.

・The ability to comprehend a passage.

・The ability to make an inference from a general understanding of the passage.

## Dynamic Assessment

The primary method of speaking assessment at our university has been through peer-to-peer interaction.  However, Hill & Sabet（2009）identified three possible areas for the application of dynamic speaking assessment（DSA）:

1. The use of three methodologies (i.e., questions, prompts and role-plays).
2. Evaluation by checking off evaluation criteria.
3. Lower-level learners negatively affecting the performance of higher-level partners.

A sociocognitive perspective to DSA (Hill, 2006) suggested the following solutions:

1. Develop transfer-of-learning between role-plays.
2. Provide mediated assistance in the form of questions and prompts.
3. Pair learners from higher to lower level and allow them to observe the previous pair's performance.

Dynamic assessment (DA) distinguishes itself from non-dynamic assessment in that it does not separate instruction from assessment. DA in the classroom is a development-oriented process of collaborative engagement that reveals the underlying causes of learners' performance problems and helps learners overcome them (Poehner, 2008). The first application, *mediated assistance*, involves interaction between an assistor and a learner to reveal problems in spoken performance. The mediation by the assessor is designed to assist learners in solving problems (Lantolf & Thorne, 2006). The second approach is to discover learners' ability to "transfer" what they have already internalized to novel problems. This *transfer-of-learning* ability to overcome performance problems also represents the primary means of assessing genuine development. Additionally, DSAs have a pair-work symmetrical nature so some transfer of learning from higher to lower-level pairs may also be observable. The Vygotskian (1986) construct of the *zone of proximal development* (ZPD) is therefore the third DSA application. However, the proximal zone between more and less competent individuals is not the only construct under observation. The same developmental and problem-solving ability between individuals can also become collective and interconnected in groups.

Prior to this attempt at integrating the SLEP test into the ECA curriculum, two previous studies (Hill & Sabet, 2008; Hill, in progress) led to the attempt to use DSAs in the integration of standardized testing into the curriculum. The DSA integration approaches to the SLEP test are as follows:

· *Transfer of Learning* (i.e., the ability to transfer what was learned in DSAs to the standardized SLEP).

・*Mediated Assistance*（i.e., corrective feedback during DSAs or short dialog practice）.

・*The Group Zone of Proximal Development*（the GZPD; i.e., allowing pairs to observe other pairs during DSAs and asking them a comprehension question about the previous pair's short dialog, which was based on the SLEP test）.

The results of the previous Hill & Sabet（2009）study showed that learners' speaking assessments significantly improved once a DSA approach was applied.　In the second study, each DSA was based and practiced on the pre/posttests in the ECA curriculum.　As learners participated in practice and the subsequent first DSA, their common errors were determined, pointed out, and then integrated into a revised written and listening posttest.　When compared with a control group who had not participated in DSAs, the study group showed significant improvement in their posttest scores as well as with their ability to correct common errors.　Additionally, results of a posttest questionnaire indicated that learners had experienced positive washback from basing the DSAs on the written and listening pre/posttests.

The results from these studies were significant enough to suggest further investigation into the integration of dynamic processes to static or standardized forms of assessment（i.e., the SLEP test）.　Static refers to viewing learning as a *product* rather than underlying cognitive *processes* involved in learning development; however, this understanding of learning a dynamic process is still unfamiliar to many teachers.　Attempts to dynamically integrate the SLEP test into the ECA program therefore met with varying degrees of success because of the predominance of static forms of assessment and an inability to integrate new assessment and instruction approaches into classes.

## Integrating the SLEP

Teachers can get a better measurement of learners' ability when learners understand the test format and have prior experience with the test（Westrick, 2005）.　Aneiro（1989）also reported a significant correlation between low anxiety and high listening ability, suggesting that the use of affective strategies could facilitate listening ability.　Since dynamic approaches to assessment were unfamiliar to some teachers, they felt that focusing on non-dynamic skill-building approaches and exercises was the best way to integrate the SLEP test into the curriculum.　Although all teachers, who are also assessors, did not take part in the dynamic aspect of integrating the test into the curriculum, the study implementing them was still conducted in some classes and, therefore, the

dynamic integration study is reported on here.

Another factor in providing extra practice for the test is to familiarize the learners with the test's question types in order to reduce any anxiety learners may feel towards them. When taking the actual test, if learners know what to expect, and have practiced for it, then their chances of scoring higher may increase. Naturally, seeing higher gains is one of the main goals for integrating the SLEP test into the curriculum. Analysis of the test indicated integration of the test may be most effective with the following sequence of instruction:

- ・Reading practice.
- ・Vocabulary practice in the form of worksheets.
- ・Dynamic short dialog practice.
- ・DSAs based on the short dialogs.
- ・Dynamic extended listening practice.
- ・Test-taking strategies and tips.
- ・A practice SLEP.

Integration of the materials was done in two sessions (i.e., at mid-term and at the end of term) of five classes each: 1) reading practice, 2) vocabulary worksheets, 3) short dialog practice, 4) a DSA based on short dialogs, and 5) extended listening practice. The test-taking strategies and the practice SLEP were explained and used after the second integration session at the end of the term.

## Reading Practice

The reading teachers also made and used their own grammar-based reading activities in their classes. However, because some departments do not take reading classes, speaking-based reading tasks were developed to dynamically practice the SLEP test in ECA speaking classes. These tasks mostly involved pair-work and focused on keywords, complex sentences, transition phrases, comprehension questions, and making an inference.

### *Keywords*

In the first part of the reading practice, learners were asked to work with their partner to fill in the blanks in the word bubbles. Student A had half the words necessary to complete the sentence and Student B had the other half, so they had to work together to complete the task. Here is an

example:

１.

He read (1)_____ to us
and we enjoyed the
(2) _____ in the story.

| Student A | Student B |
|---|---|
| characters | stories |

### *Complex Sentences*

In the second section, both Student A and Student B were asked to look at photos and work with their partner to join the clauses together to make a sentence that matches the sentence in the picture. Student A had one half of the clause and Student B had the other half. Here is an example:

１.

| Student A | Student B |
|---|---|
| c. The man is sitting . . . | e. and having his shoes polished. |

### *Transition Phrases and Making an Inference*

In the third part, partners were first asked to work together to pick the best grammatical choice from the box to complete the sentences in the passage. Again each partner had half of the possible correct answers. They were told that the vocabulary could be difficult, but to try to think of the correct grammar to connect the passage. Here is an example:

Barack Obama attended Columbia University, but found New York's racial tension inescapable. He became a community organizer for a small Chicago church-based group for three years [ 1. ] poor South Side residents cope with a wave of plant closings.

Student A                         Student B

１．a. helped     b. helpless    １．c. helpful     d. helping

Then they were asked to work with their partner to answer the comprehension questions pertaining to the passage. Finally, they were given the Japanese translation for *infer* (i.e., 推論する) and encouraged to work with their partner to answer the inference question.

The reading practice was carried out over two classes with the lowest streamed level of learners. The first class focused on understanding keywords, complex sentences, and reading the passage. In the second, learners answered the comprehension and inference questions. In these classes, the instructor, or assistor, provided mediated assistance with each pair as they completed the tasks. Additionally, each pair observed the previous pair in order to develop the GZPD. With mediated assistance from the assistor and observation of the previous pair, many of the lowest-level learners were able to grasp the final inference about the general meaning of the passage. Here is the inference question:

Q. What can you infer (推論する) about Barack Obama's activities?

a. He is looking for a good job.
b. He wants people to lose their jobs.
c. He's interested in people's civil rights.
d. He wants to win elections.

### *Vocabulary Worksheets*

Most of the words used in the SLEP test are American English and they correspond with academic topics and subjects found in a typical high school. As stated previously, the SLEP test is usually administered to non-native speakers in an ESL environment in the U. S. or Canada. The following are some of the vocabulary found in Form 4 of the SLEP:

student council, period, biology, project, rehearsal, review, assignment, research, particularly, computer lab

The next step to integration into the curriculum involved developing vocabulary worksheets that introduced the students to this academic vocabulary and reinforced it through Student A/Student B exercises. Students first worked together to match the word with the Japanese equivalent. They then matched the word to its meaning in English. Finally, they filled in the blanks with the words to short Student A/Student B conversations, the content of which was based more on the students' actual learning context. They then practiced the conversations together. In this way, it was hoped these short adjacency pairs would be more socially valid to learners, so they might be able to co-construct a better understanding of the meanings of the words from their context. They were also intended to prepare students for the longer conversations found in the short dialogs. Here is an example:

| *student council* | Q. A：Did you hear Satoshi might get suspended from school? |
| | B：Yes, they're going to discuss it at the next . . . meeting. |

### Dynamic Short Dialog Practice

Many of the short dialogs or speech events in the SLEP test matched the ones found in the ECA Speaking curriculum. Therefore, new short dialogs that matched the ECA curriculum were developed based on the SLEP test. Once again, although these short dialogs were analogous to the test, they were based on the perspective of a typical Japanese university student. These dialogs were then used in DSAs. The dialogs also recycled the academic vocabulary found in the vocabulary worksheets.

Prior to the DSA, learners practiced each dialog with a partner and the teacher provided corrective feedback in the form of recasts. Corrective feedback (i.e., the strategic use of reformulations of learners' erroneous utterances) is another form of mediated assistance in DSA. Higher-level learners were paired with lower-level learners and they practiced the dialogs in pairs, building their comprehension through procedural knowledge. They then developed their own dialogs using the previous practice dialog as a prompt. They used their own dialogs in the DSA and each pair observed the previous pair's performance. After completing their dialog, the pair observing the previous pair was asked to respond to a comprehension question about the dialog

similar to that found in the SLEP test.  Here is an example dialog and the prompt used by learners:

On tape:

Q：What didn't he like about Disneyland?


A：What did you do this weekend?

B：We went to Disneyland.

A：How was it?

B：It was great.  You should go.

A：Why?

B：Well, because there are so many attractions and different rides you can go on.

A：Was there anything you didn't like?

B：Well, the only thing I didn't like was the long line-ups for some rides.

A：Yeah, I've heard that you have to get there early for those.


Q：What didn't he like about Disneyland?


In the test book:

A．The attractions.

B．The different rides.

C．The long line-ups.

D．Getting there early.


Prompt:

Q：What didn't he/she like about _____?


A：What did you do this weekend?

B：We went to _____.

A：How was it?

B：It was great.  You should go.

A：Why?

B：Well, because there are so many _____ and _____.

A：Was there anything you didn't like?

B：Well, the only thing I didn't like was the _____ .

A：Yeah, I've heard that you have to _____ .


Q：What didn't he/she like about _____?


Once again, during the DSA, most of the lower-level learners observing the previous pair were able to respond correctly to the comprehension question about the dialog.


### Dynamic Extended Listening Practice

extended listening dialogs were developed based on the extended listening dialogs found in the final section of the SLEP test. This activity was intended to develop learners' declarative knowledge based on the procedural knowledge developed in the short dialogs. In order to practice the extended listening, learners were paired according to their pre/posttest scores and two pairs sat at one table. One pair read a section from the dialog, while the other pair listened. The pair who listened could only read the multiple choice answers. After the first pair finished reading the dialog, the assessor read the comprehension questions and the pair listening tried to answer the questions from the multiple choice answers. If they could not answer, the first pair read the section of the dialog once more.

This approach is intended to develop learners' ability to anticipate answers. This is a more active comprehension process, building from one listening comprehension question to anticipate the next. The idea here is that listening comprehension is not just a matter of "listen for it" but an active dynamic process of trying to anticipate what will be said next. The answer to each question should build on the previous one to create a forward-looking pursuit of understanding. The second time a dynamic extended dialog was practiced the role of reader and listener was then reversed. Here is an excerpt from the more authentic extended dialog practice:


On tape:

Masayuki: Hi Hiroyuki, are you going to the English Speaking Society club meeting today?

Hiroyuki: I wanted to but I'm not sure if I'm ready for our speech contest.

M：Why not? Didn't you practice?

H：Yes, but not enough.

M：Oh that's too bad. So what are you going to do then?

H：I don't know . . . should I say I'm sick?

M：Hmm, I don't think most people would believe that.

H：Okay, well then I was thinking of practicing my speech before the meeting and maybe coming to it late.

M：That sounds better.  I'll let everyone know you might be late.


Q 1.  Why doesn't Hiroyuki want to go to the English Speaking Society club meeting?

Q 2.  What does Hiroyuki want to do before the meeting?


In the test book:

1．A．He doesn't like rules.

　　B．He's not going to the English club meeting.

　　C．He doesn't like to practice.

　　D．<u>He's not ready</u>.


2．A．<u>Practice his speech.</u>

　　B．Get sick.

　　C．Not believe it.

　　D．Be late.


## Test-Taking Strategies


Because of the various levels of learners and classes, as well as the dynamic and non-dynamic attempts to integrate the SLEP test into the curriculum, a set of test-taking strategies was developed and provided to learners in hopes that it would result in positive washback after taking the test.  The test-taking strategies were also translated into Japanese for the learners.  The following are the test-taking strategies and tips explained to learners prior to taking the practice SLEP:


### Listening Strategies
*Section 1: Photos*

*Strategies:*　1．Look at the photo.

　　　　　　　2．Who is in the photo?

　　　　　　　3．What are they doing?

4．What are they doing it to?

5．Where are they?

*Section 2: Same Sentence as You Hear*

*Strategies:*　1．Read the sentences.

　　　　　　　2．Focus on the order of the words.

　　　　　　　3．The meaning of the sentence is not important.

*Section 3: Short Dialogs*

*Strategies:*　1．Read the question carefully: Does it ask: "*What, Where, Why, When* or *How*?

　　　　　　　2．Anticipate the answer from keywords in the four possible answers.

　　　　　　　3．Identify who is speaking and the topic of the dialog.

　　　　　　　4．When you hear the keywords repeated, listen for the details to answer the question: *What, Where, Why, When* or *How*?

*Section 4: Extended Dialogs*

*Strategies:*　1．Listen carefully for the question.

　　　　　　　2．Anticipate the answer from keywords in the four possible answers.

　　　　　　　3．Identify who is speaking and who is saying what to whom.

　　　　　　　4．When you hear the keywords repeated, listen for the details to answer the question.

　　　　　　　5．Remember the same speakers are talking for the whole extended dialog.

　　　　　　　6．Many of the questions ask who is saying what to whom.

### Reading Strategies

*Section 1: Keywords*

*Strategy:*　1．Look for keywords in the sentences to match them to the pictures in the mind bubbles.

*Section 2: Complex Sentences*

*Strategies:*　1．Look at the sentence. There are two parts to the sentence.

　　　　　　　2．Look at both parts of the sentence.

　　　　　　　3．Match the sentence to the only photo which is the same as the whole sentence.

4. The other three photos will not match both parts of the whole sentence.

*Sections 3: Transition Phrases*

*Strategies:* 1. Look at the words in the boxes.

2. If you don't know the word, use the words that come before and after the box to guess the meaning of the word in the box.

3. Think of the word's grammar to correctly connect the words that come before and after.

*Section 4: Making an Inference*

*Strategies:* 1. Read the passage from beginning to end.

2. Read the comprehension questions and look for the details in the passage.

3. Read the passage from beginning to end again to understand the passage's general meaning.

4. Answer the last few questions.

### Tips for Test-Taking Success

These tips are based on the behavior rating scales in Lidz (2000):

1. *Attention:* Maintain attention to the test from beginning to end. Answer all the questions.

2. *Persistence:* Complete the test. Try to answer each question the first time.

3. *Management:* Quickly fill in the blanks and spend your time by reading the multiple choice answers for the next question.

4. *Flexibility:* Remember what you practiced for the SLEP in class. If you understand the directions and examples, start reading the first questions or answers in that part.

5. *Tolerance:* Stay calm and focused on doing your best.

6. *Motivation:* Be open to listening and reading English. Before the test, listen to something in English.

## The Practice SLEP

In addition to these strategies and tips, learners were also given a practice SLEP test for both Listening and Reading sections. This practice test was especially important for the first two

sections where practice may have the most significant gains. Once the practice test was over, correct answers were given and explained to the practice test-takers.

## Descriptors for Streamed Levels

Because of the difficulty of the SLEP test and since all levels of learners could not be expected to perform the same, these descriptors were developed for the different streamed levels found in the ECA program（each descriptor entails the ability of the lower level）:

### *Listening*

*A Level:*　　Ability to comprehend extended dialogs and answer questions.

*B Level:*　　Ability to comprehend short dialogs and answer questions.

　　　　　　Ability to comprehend some extended dialogs and answer questions.

*C/D Level:* Ability to listen to some short dialogs and answer questions.

### *Reading*

*A Level:*　　Ability to make an inference from reading a passage.

*B Level:*　　Ability to answer comprehension questions from a passage.

*C/D Level:* Ability to choose some correct grammar in a passage.

## Discussion

Because learners have not yet taken the post-SLEP test, the data of the results of integration have not been collected and cannot be reported here. Therefore, this discussion will only briefly comment the main points of integration, limitations of integration, and future areas for integration. That said, integration of the pre/post-tests with DSAs in previous studies saw significant increases in scores so the same results are hoped for with the integration of the SLEP test into the ECA curriculum. Analysis of the SLEP has also been useful for developing materials to integrate into the curriculum. The use of more procedural knowledge or speaking-based processes seems to result in significant increases in learners' declarative or written knowledge. This also seems to result in learners being more motivated to undertake standardized tests. See Appendix A for a questionnaire administered after the post-SLEP which was intended as a form of triangulation between integration materials and SLEP tests results, and also as a means to determine any

positive washback effects with learners.

The main limitation of this study is the lack of some teachers to adopt more dynamic approaches to integrating the curriculum into their classes. This limitation may mean that the materials they used taught more *to* the test rather than *within* it. Thus, future integration of the SLEP test into the curriculum should involve developing a method to increase teachers' awareness of, and receptivity to, the benefits of more social and dynamic approaches to assessment and instruction. The other main point of interest will be the results of the questionnaire. If results of the questionnaire indicate a positive washback effect, then this could be an indication that integration of the SLEP test into the curriculum has been educationally beneficial.

## Conclusion

Integration of the SLEP test into the curriculum shows more clearly how instruction and assessment are one in the same thing. Dynamically using materials that are different from the actual test items, results in teaching within the test rather than to it. The addition of a dynamic element to assessment tunes into learners' development processes and allows them to perform at their potential level. Integrating dynamic approaches to standardized tests also removes some of the static individualistic focus found within them. Awareness of the social aspect of assessment can lead to further integration between assessment and the curriculum. After having completed the SLEP posttest in January 2010, learners' gains will then be compared with those from previous years. If, along with positive washback, significant gains occur, then the next step will be to further integrate the SLEP test practice materials into the ECA curriculum. Finally, if integration does show significant gains, then this may be an indication that the producers of standardized tests should also be looking for ways to remove the individualistic focus from their tests and to further integrate the social and dynamic aspects of assessment into them.

### References

Brown, J. D. (1996). *Testing in language programs*. New York: Prentice Hall.

Culligan, B & Gorsuch, G (1999). Using commercially produced proficiency test in a one-year core EFL curriculum in Japan for placement purposes. *JALT Journal*. 21 (1), 7-28.

Educational Testing Services. (2004). *SLEP manual*. Princeton, NJ: ETS.

Gorsuch G. J. (1995). Tests, testing companies, educators, and students. *The Language Teacher*. 19 (10).

Hill, K. (2006). A sociocognitive perspective: The best of both worlds. *TESOL Quarterly*. 40 (4), 819-825.

Hill, K. & Sabet, M. (2009). Dynamic speaking assessments. *TESOL Quarterly*, 43 (3), 537-545.

Hughes, A. (1989). *Testing for language teachers*. Cambridge: Cambridge University Press.

Inbar-Lourie, O. (2008). Constructing a language assessment knowledge base: A focus on language assessment courses. *Language Testing*. 25(3), 385-402.

Lantolf, J. & Thorne, S. (2006). *Sociocultural theory and the genesis of second language development*. Oxford: Oxford University Press.

Leung, C. & Lewkowicz, J. (2006). Expanding horizons and unresolved conundrums: Language testing and assessment. *TESOL Quarterly*. 40(1), 211-234.

Lidz, C. S. (2000). The application of cognitive functions scale: An example of curriculum-based dynamic assessment. In C. S. Lidz and J. G. Elliott (Eds.), *Dynamic assessment: Prevailing models and applications*. New York: Elsevier. pp. 407-439.

McNamara, T. (2005). The social turn in language assessment. In E. Hinkel (Ed.), *The handbook of research in second language teaching and learning*. Mahwah, NJ: Lawrence Erlbaum. pp. 775-778.

Poehner, M. 2008. *Dynamic assessment: A Vygotskian approach to understanding and promoting L 2 development*. Berlin: Springer.

Taguchi, N. (2001). L 2 Learners' strategic mental processes during a listening test. *Jalt Journal*. Volume 23(2), 176-201.

Vygotsky, L. (1986). *Thought and language*. Cambridge, Mass.: The MIT Press.

Weir, C. (1993). *Understanding and developing language tests*. New York: Prentice Hall.

Westrick, P. (2005). Score reliability and placement testing. *JALT Journal*. 27(1), 71-93.

# Appendix A

**Integrating the SLEP Questionnaire**

1．Do you think using the SLEP practice materials in class may have improved your SLEP score?

（授業で SLEP 練習問題を使用したことで，あなたの SLEP テストの点数は上がったと思いますか？）

A．yes very much　　B．yes a little　　C．not yes or no　　D．no not much

E．no not at all

（A．とてもそう思う．　　B．多少そう思う．　　C．どちらでもない　　D．あまりそう思わない

　E．全くそう思わない）

2．Did you understand the SLEP Test better after doing this kind of practice?

（SLEP テストの練習をしたことで，SLEP テストの形式が良く分りましたか？）

A．yes very much　　B．yes a little　　C．not yes or no　　D．no not much

E．no not at all

（A．とてもそう思う．　　B．多少そう思う．　　C．どちらでもない　　D．あまりそう思わない

　E．全くそう思わない）


3．Do you think the SLEP practice materials were useful in ECA speaking classes?
（SLEP 練習問題は，スピーキングのクラスでは役に立ったと思いますか？）


A．yes very much　　　B．yes a little　　　C．not yes or no　　　D．no not much

E．no not at all

（A．とてもそう思う．　　B．多少そう思う．　　C．どちらでもない　　D．あまりそう思わない

　E．全くそう思わない）


4．Do you think the SLEP practice materials were useful in ECA reading classes?
（SLEP 練習問題は，リーディングのクラスでは役に立ったと思いますか？）


A．yes very much　　　B．yes a little　　　C．not yes or no　　　D．no not much

E．no not at all

（A．とてもそう思う．　　B．多少そう思う．　　C．どちらでもない　　D．あまりそう思わない

　E．全くそう思わない）


5．Do you think this SLEP practice in class also improved your English ability?
（授業で SLEP 練習問題をやったことで，あなたの英語力は向上したと思いますか？）


A．yes very much　　　B．yes a little　　　C．not yes or no　　　D．no not much

E．no not at all

（A．とてもそう思う．　　B．多少そう思う．　　C．どちらでもない　　D．あまりそう思わない

　E．全くそう思わない）


6．Did this SLEP practice in class give you more confidence to take the SLEP test?
（授業で SLEP 練習問題をやったことで，もっと自信をもって SLEP テストを受けることが出来た
と思いますか？）


A．yes very much　　　B．yes a little　　　C．not yes or no　　　D．no not much

E．no not at all

（A．とてもそう思う．　　B．多少そう思う．　　C．どちらでもない　　D．あまりそう思わない

　E．全くそう思わない）


7．Did this SLEP practice in class encourage you to continue studying English?

（授業で SLEP 練習問題をやったことで，これからも英語を続けて勉強したいと思いましたか？）


A．yes very much　　　B．yes a little　　　C．not yes or no　　　D．no not much

E．no not at all

（A．とてもそう思う．　　B．多少そう思う．　　C．どちらでもない　　D．あまりそう思わない

　E．全くそう思わない）


8．Would you like to spend more or less time doing this SLEP practice in ECA speaking classes?

（スピーキングのクラスでは，SLEP 練習問題にもっと時間をかけたいですか？　少ない方が良い

ですか？）


A．a lot more　　　B．a little more　　　C．neither more or less　　　D．a little less

E．a lot less

（A．もっと沢山時間をかけたい．　　B．もう少し時間をかけたい．　　C．どちらでもない

　D．もう少し少ない方がよい　　E．ずっと少ない方がよい）


9．Would you like to spend more or less time doing this SLEP practice in ECA reading classes?

（リーディングのクラスでは，SLEP 練習問題にもっと時間をかけたいですか？　少ない方が良い

ですか？）


A．a lot more　　　B．a little more　　　C．neither more or less　　　D．a little less

E．a lot less

（A．もっと沢山時間をかけたい．　　B．もう少し時間をかけたい．　　C．どちらでもない

　D．もう少し少ない方がよい　　E．ずっと少ない方がよい）


10．Do you think there is enough SLEP practice materials in class?  Would you like:

（授業で使う SLEP 練習問題は十分あったと思いますか？　それとも次のうちどれがよいです

か？（A．スピーキングの練習を多く，B．スピーキングの練習を少なく，A．リスニングの練習を多く，B．

リスニングの練習を少なく，A．リーディングの練習を多く，B．リーディングの練習を少なく）

A．more speaking practice　　B．less speaking practice

A．more listening practice　　B．less listening practice

A．more reading practice　　B．less reading practice

11. Did you try your best when using the SLEP practice materials in class?

（あなたは，授業で SLEP 練習問題を一生懸命学習しましたか？）

A．yes very much　　B．yes a little　　C．neither yes or no　　D．no a little

E．no very much

（A．とてもそう思う．　　B．多少そう思う．　　C．どちらでもない　　D．あまりそう思わない

　E．全くそう思わない）

12. Do you think you did better or worse at the SLEP Test this time than when you did it last April?

（前回に比べて，あなたは今回の SLEP テストの結果の方が良いと思いますか？　悪いと思いますか？）

A．a lot better　　B．a little better　　C．the same　　D．a little worse

E．a lot worse

（A．すごく良い　　B．少し良い　　C．同じ　　D．少し悪い　　E．ずっと悪い）

If you have any comment about the good points of the practice SLEP materials or about how the materials could be improved, please write your comments in the box.

（SLEP test 練習問題の良い点，または改善点についてコメントがありましたら下に書いてください。）