2005

CALL Evaluation by Developers, a Teacher, and Students

Joan Jamieson
Northern Arizona University

Carol Chapelle
Iowa State University, carolc@iastate.edu

Sherry Preiss
Pearson Longman ELT

CALL Evaluation by Developers, a Teacher, and Students

Abstract

Researchers and teachers agree that evaluation of CALL should ideally inform pedagogical choices about how best to use CALL, but how to go about such an evaluation is not clear. This study offers an example of a context-based evaluation by operationalizing criteria for CALL evaluation and administering the instruments to three groups of stakeholders: the people who developed the content for the CALL materials, the teacher, and the students. The CALL materials were Longman English online (LEO). The setting was a community college English as a second language class in New York. Results, which focused on the agreement among stakeholders and their assessment of factors pertaining to six criteria, indicated good agreement among stakeholders and overall positive evaluations, but also identified some areas for improvement in the materials and the evaluation instruments.

CALL Evaluation by Developers, a Teacher, and Students

Joan Jamieson
Northern Arizona University

Carol A. Chapelle
Iowa State University

Sherry Preiss
Pearson Longman ELT

ABSTRACT

Researchers and teachers agree that evaluation of CALL should ideally inform pedagogical choices about how best to use CALL, but how to go about such an evaluation is not clear. This study offers an example of a context-based evaluation by operationalizing criteria for CALL evaluation and administering the instruments to three groups of stakeholders: the people who developed the content for the CALL materials, the teacher, and the students. The CALL materials were *Longman English online* (*LEO*). The setting was a community college English as a second language class in New York. Results, which focused on the agreement among stakeholders and their assessment of factors pertaining to six criteria, indicated good agreement among stakeholders and overall positive evaluations, but also identified some areas for improvement in the materials and the evaluation instruments.

INTRODUCTION

Early studies comparing CALL with classroom instruction have shown no clear advantage or disadvantage of CALL, possibly due to the difficulty of controlling the large number of factors that effect individuals' learning (e.g., Burston, 2003). Such comparison studies make use of experimental or quasi-experimental research designs to make comparisons. A characteristic of experimental approaches is manipulation or control over behavioral events (Jaeger, 1990). Indeed, in an experimental setting the researcher tries to reduce or exclude context from what is being studied in order to focus on the few variables of interest and to increase generalizability of the study. The attempted exclusion of context in the study of

CALL has proven problematic in part because it is typically the contextual factors that contribute greatly to success or failure. This is one reason that researchers have called for alternative research strategies (Chapelle, 2001, 2003; Dunkel, 1991; Pederson, 1987).

An equally important consideration pertaining to research methodology today is the fact that justification of CALL is less relevant as computer use has become commonplace for all kinds of instruction, including foreign languages. Today's more pressing questions ask for evidence for effective language learning through analysis of software design and learner engagement with learning tasks (Chapelle, 2003). Addressing such questions requires evaluation of CALL in terms of context-specific arguments supported by rationales and evidence based on theory and research in instructed second language acquisition (SLA). The question for research, then, is to what extent a particular type of CALL material can be argued to be appropriate for a given group of learners at a given point in time?

Current approaches to instructed SLA reflect a middle ground emphasizing language as the vehicle for goal-oriented activities—not the target of instruction—but warning that focus on meaning should not be taken to an extreme that ignores focus on form. In the interest of formulating perspectives from instructed SLA in a manner that would guide CALL evaluation, Chapelle (2001) defined a set of criteria, as summarized in Table 1.

Table 1
Criteria for CALL Evaluation

Language learning potential	The degree of opportunity present for beneficial focus on form
Meaning focus	The extent to which learners' attention is directed toward the meaning of the language
Learner fit	The amount of opportunity for engagement with language under appropriate conditions given learner characteristics
Authenticity	The degree of correspondence between the learning activity and target language activities of interest to learners out of the classroom
Positive impact	The positive effects of the CALL activity on those who participate in it
Practicality	The adequacy of resources to support the use of the CALL activity

Instructed SLA encourages a focus on the linguistic form of language as the need arises in the context of meaning-based instruction. Researchers advocate tasks in which language is used for a realistic purpose, while recognizing that language use should be fluent, accurate, and complex (Brown, 2000; Crookes & Chaudron, 2001; Larsen-Freeman, 2000; Lightbown & Spada, 1999; Mellow, 2002; Savignon, 2001; Skeehan, 1998). For these reasons, language learning po-

tential, meaning focus, and authenticity are three criteria for CALL evaluation (Chapelle, 2001). Learner fit is included to reflect the ways in which individuals differ such as age, learning style, and stages of development (Pienemann, 1985). Positive impact is included among the criteria for CALL evaluation in recognition of the importance of the learners' attitude in language learning as well as the learners' perceived benefits by participating in activities (Brown, 2000; Bachman & Palmer, 1996; Chapelle, 2001). Finally, practicality is included because of the necessity of having the time and money required for the CALL activity.

These criteria will help to focus evaluation of effective language learning through CALL on the materials themselves, such as their software design or their task development, and the ways in which learners interact with them. Chapelle (2001) described the use of these criteria for guiding both judgmental and empirical evaluation of CALL materials. Judgmental evaluation is based on an individual's logical analysis, whereas empirical evaluation is based on analysis of observed data, sometimes from an individual, but often from a group of individuals. These two perspectives on the six criteria were intended to offer some conceptual infrastructure for the development of research that does not rely blindly on the unanalyzed categories of "computer" and "classroom" and that does not require separation of the CALL activities from the context in which they are used.

In this study, we were interested in evaluating the overall appropriateness of CALL materials. We adopted a case study approach which used empirical data to address the following question: How appropriate are CALL materials for a group a learners in a particular context? This approach is not altogether different from experimental methods. Yin (1994) cited similarities between the experimental and case study approaches regarding the type of research questions posed and the degree of focus on contemporary rather than historic events. However, the approaches are different in an important way. As Yin points out, the two approaches differ in the extent of control an investigator has. Although in some situations a researcher can choose among research strategies, other times one research strategy has distinct advantages: "For the case study, this is when a 'how' or 'why' question is being asked about a contemporary set of events over which the investigator has little or no control" (p. 9). Yin goes on to define a case study as an empirical inquiry that "investigates a contemporary phenomenon within its real-life context, especially when the boundaries between phenomena and context are not clearly evident" (p. 13).

We defined appropriateness of CALL materials in terms of the six criteria from theory and research in instructed SLA. We conducted the analysis by considering opinions about these criteria from three distinct, yet related, groups of stakeholders: the developers of the CALL materials, the teacher who decided to include the CALL materials in her class, and the language learners who used the CALL materials. We summarized the opinions of these three groups in response to the following research questions: What was the evaluation of the CALL materials given by each of the groups of stakeholders for each of the six criteria? Did the three groups of stakeholders have the same opinions? Did the opinions of devel-

oper, teacher, and learners indicate that the CALL materials were appropriate for this context by meeting the criteria?

METHODS

Five important components in the design of a case study were explained by Yin (1994, pp. 20-27): the study's questions, its propositions, its unit(s) of analysis, the logic linking the data to the propositions, and the criteria for interpreting the findings. In this study, the questions at the end of the Introduction above addressed how to determine the appropriateness of the CALL material for a group of learners. Propositions are intended to focus the study; they may represent important theoretical issues as well as tell the researcher where to look for relevant evidence. The propositions in this study reflect the six criteria illustrated above in Table 1. The unit of analysis defines what the "case" is. In this study there were three units of analysis: the people who wrote the content for the CALL materials (the "developers"), the teacher who decided to use the CALL materials in her class (the "teacher"), and the learners who used the CALL materials (the "students"). Thus, this was a multiple-case design. When using multiple cases, Yin points out that it is important to distinguish between the replication logic of a case study and the sampling logic of an experiment or survey (pp. 45-52). Sampling logic requires that a population be defined and then that a procedure be specified for selecting a subset of that population; replication logic follows that of multiple experiments—if similar results are obtained from similar experiments, replication has occurred. With multiple cases, each case is like an experiment. In this study, each case is predicted to produce similar results. Linking data to the propositions is explained in the Instruments section below, and the criteria for interpreting the findings is described in the Discussion section.

Participants

Sometimes potential for learning is a superior criterion to representativeness; Stake (1994) recommended considering logistics, the potential reception, and resources, and then selecting a case that the researcher could spend the most time with. These factors came into play as we selected English as a second language (ESL) learners at a community college in Queens, New York City, USA. A teacher had found a description of the CALL material on the Web and she contacted the publisher about trying it with her classes.

Teacher

The teacher of the classes was a 35 year old female with a Master's degree in Teaching English as a Second Language (ESL) and several years experience teaching ESL. Interested in the field of CALL, she planned to return to school for an advanced degree in teaching and technology. In her ESL classes for the past two years, she had used Blackboard, as well as a grammar and a writing textbook with accompanying CDs. She reported that the director of the intensive ESL program was very supportive of teachers using technology.

Students

The students were enrolled in one of two classes in the intensive ESL program in the adult and continuing education division of the community college. The curriculum at the intensive ESL program was comprised of six levels of proficiency. The morning class met for 20 hours and the afternoon class met for 18 hours Monday through Thursday. Both of these classes were at Level 3, which was considered "low-intermediate." The students had classes in grammar, reading, writing, and listening and speaking. The teacher who participated in the study was their listening/speaking teacher.

Forty-two students participated in the study. There were 13 males and 24 females; 5 students did not fill out their sex on a questionnaire. They ranged in age from 18 to 43; their average age was 25. In terms of their first languages, 10 spoke Korean, 9 spoke Spanish, 7 spoke Japanese, 4 spoke Polish, and 1 person each spoke Cantonese, Chinese, Italian, Mongolian, Thai, Turkish, and Turkmen. Many of these students worked during the day or evening; some of the students were interested in attending college in the USA. Many students lived with either family or friends from their home country and did not speak much English at home.

Developers

Four of the developers of the content of the CALL materials participated in the study. Three of the developers were female and one was male. The developers were adults and firmly established professionals as experienced ESL textbook authors and editors. Two of the developers were freelance authors of the materials, and the other two developers were editors of the material employed by the publisher, one of whom is a co-author of this article.

Materials

The material used in this study was an ESL CALL series called *Longman English Online*.[1] In total this CALL courseware consisted of four levels *LEO 1*, *LEO 2*, *LEO 3*, and *LEO 4*; in this study, we used *LEO 3*. The *LEO* courseware is delivered via an authoring system that was created for Pearson Longman ELT by a group of instructional designers, programmers, and graphic artists. The screen designs and navigation icons were all tried out with small groups before the course was implemented. Consequently, *LEO*'s presentation is clear, easy to use, and professional looking.

The *LEO 3* package (Rost and Fuchs, 2004) is the intermediate level of this video-based multimedia, integrated skills program. Each of its 12 units continues the story of a young journalist who is covering a story about a soccer scandal. Figures 1 and 2 display screen shots from a listening comprehension activity and speaking activity, respectively.

Figure 1
Screen Shot of Listening Comprehension Section of *LEO 3*
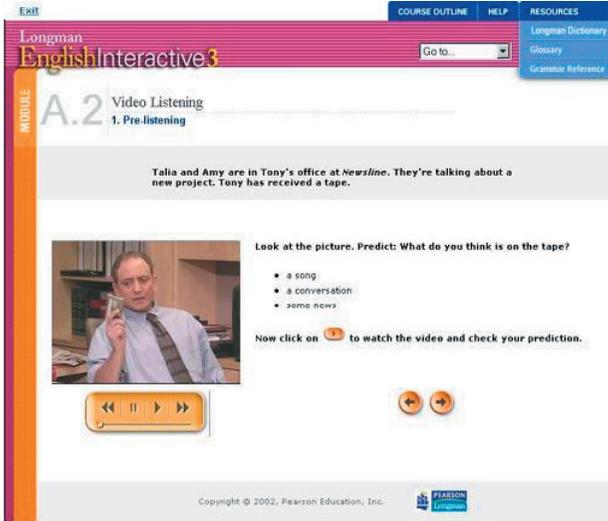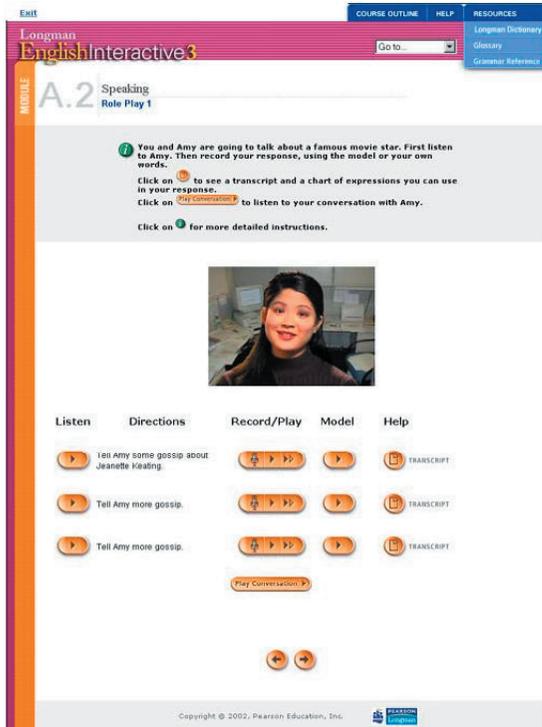

Figure 2
Screen Shot of Speaking Exercise in *LEO 3*

In Figure 1, we can see that the listening material is presented to the learner via video. The high quality video is accessible with an interface that allows for pausing, rewinding, fast forwarding, and playback. The listening sections begin with prelistening activities such as making predictions about what will come next, heightening the learners' awareness and encouraging use of background knowledge. In Figure 2, functional phrases that were presented in other parts of the lesson are practiced in a type of role-play activity. The learners first listen to an audio clip that has been used previously in the unit. Next, they read directions on the screen giving them the content (but not the linguistic form) that they should incorporate into their response. Learners are then able to record their responses and play them back so that they hear both the initial utterance and their response. Help is available in the forms of an audio model and written transcripts.

The units not only contain listening comprehension and speaking practice, but also include explanation and practice with grammar, vocabulary, pronunciation, reading, writing, and an "On the Web" activity. Each unit concludes with a review quiz. There are also longer tests at the end of every four units called "module tests," and at the end of all 12 units, called "the end-of-level test." During the time of this study, the students worked on Module A—the first four units of *LEO 3*. Although this material is no longer available via web delivery, the interested reader can view examples of the CD-ROM version of the material at the publisher's web site (http://www.longman.com/ae/multimedia/programs/lei3_4.htm).

Instruments

For each of the six criteria, questions were developed to provide us with operational definitions. In Table 2, we have briefly defined each criterion listed in the first column. In the second column, we posed questions that could be used to inform us of whether the participants thought that the criteria was successfully implemented. In order to interpret the responses to these questions, in the third column we provide the kind of answer that would provide positive evidence for quality. In this manner, we designed a plan to link our data to our propositions.

Table 2
Criteria for CALL Quality, Operationalizations as Questions, Desired Responses

Criteria	Operationalizations	Desired responses to support claims for quality
Language learning potential • sufficient opportunity for beneficial focus on form	• Will the grammar, vocabulary, and pronunciation that was studied during the week be remembered?	• Yes
	• Were the explanations clear?	• Yes
	• Were there enough exercises?	• Yes
	• Will the students' English improve as a result of *LEO 3*?	• Yes
	• Will the students' quiz scores indicate mastery of material?	• Yes

Learner fit • appropriate difficulty for learners to benefit • appropriate for characteristics of learners	• Is the material at an appropriate ability level? • Are the student characteristics as anticipated? • Is the material at the appropriate difficulty level?	• Yes, intermediate • Yes, young adult, self-motivated • Yes, somewhat difficult
Meaning focus • learner's attention primarily focused on meaning	• Will students understand or remember content?	• Yes
Authenticity • correspondence to CALL and language outside classroom	• Is the language in *LEO 3* needed for outside of class? • Is it like that used outside of class?	• Yes • Yes
Impact • learn more about strategies • lead to sound pedagogical practice • create a positive learning experience	• Do students like *LEO 3*? • Will students want to use *LEO 4*?	• Yes • Yes
Practicality • sufficient hardware • sufficient personnel • sufficient time • sufficient money	• Is the interface easy to use without help? • Are the labs and computers of sufficient quality? • Do the students have sufficient time? • Will the students get frustrated? • Will the teacher have sufficient time?	• Yes • Yes • Yes, about 3 hours/week • No • Yes, about 4 hours/week

A series of instruments which contained these questions were developed for this study as illustrated in Table 3. Questionnaires for the developers and the teacher, and questionnaires and weekly reflections for the students were designed to ask similar questions about each of the six criteria so that comparisons could be made among the three groups.

Table 3
Scheme of Data Elicitation

Criteria	Data source		
	Developer	Teacher	Student
Language learning potential	Questionnaire	Questionnaire Interview	Questionnaire Reflections Interviews Test scores
Learner fit Meaning focus Authenticity Positive impact	Questionnaire	Questionnaire Interview	Questionnaire Reflections Interviews
Practicality	Questionnaire	Questionnaire Interview	Questionnaire Interviews

Both open ended and multiple choice questions were included in the questionnaires. The multiple choice questions asked the responder to choose whether or not he or she agreed with the question, for example, "Do you think the grammar explanations are clear?" The choices were "yes," "somewhat," or "no." Only three choices were given; responses were intended to elicit opinions and general impressions. We did not want to give the impression that the results reflected a great deal of precision.

A single questionnaire was developed for the developers (see Appendix A). Two questionnaires that contained similar questions were developed for the teacher (see Appendix B) and for the students (see Appendix C)—one was administered in the beginning of the study and the other was administered at the end.

In addition to the questionnaires, interview questions were developed for the teacher and the students, and weekly reflections were developed for the students. As illustrated in Table 3, both the questionnaires and the interviews were designed to elicit information about all of the six criteria. The weekly reflections were designed to elicit information on all of the criteria except practicality. They consisted of nine sets of questions about the unit studied, the story, grammar, vocabulary, pronunciation, reading, "On the Web," speaking, and writing (see Appendix D). Items were of two types: fill in and multiple choice. The fill-in items asked students to remember what they studied in the lesson that week; these items provided information regarding language learning potential and meaning focus. The multiple-choice items elicited information on learner fit, impact, and authenticity. Scores on quizzes were also used to provide evidence for language learning potential.

Procedures

The study took place in February and March, 2003. The students had begun a new term in mid-January; the term ended in the beginning of March. Before beginning any data collection in December, 2002 and January, 2003, one of the authors spoke with the teacher and the director of the program, in order to describe the study and gain their co-operation. Once the program director and the teacher agreed to participate, the authors had to first provide documentation that they had completed an acceptable training program for Human Subjects Protection and then submit a proposal to the Institutional Review Board that oversees research done at the community college. Once the proposal was reviewed and approved, the students were informed of the study, and those who agreed to participate signed informed consent forms. Then, data collection began, following the schedule outlined in Table 4. The questionnaires were designed to take no longer than 30 minutes to complete; the weekly reflections were designed to take students 10 minutes to complete. Eight students (four from the morning class—3 females and 1 male, and four from the afternoon class—3 females and 1 male) agreed to be interviewed.

Table 4
Timetable for Data Collection

Date	Data source		
	Developer	Teacher	Student
February 7, 2003		Questionnaire Interview	Questionnaire Interviews Reflections Test scores
February 14 and 21, 2003		Interview	Reflections Test scores
February 28, 2003	Questionnaire	Interview Questionnaire	Questionnaire Interviews Reflections Test scores

During the week of February 7, two of the authors visited the classes and explained the project to the students. Those students who wished to participate signed the informed consent forms. During the first week, the students completed a questionnaire and a weekly reflection. Four students in each class were asked by the researchers to be interviewed. Each week, students filled out their weekly reflections on paper; these were then mailed by the teacher to one of the researchers. Also, as students finished a unit, their quiz scores were saved online. During the last week, the researchers again visited the classes, administered the final questionnaire, and interviewed the same students (except that one of the eight original students did not wish to be interviewed again). Participating students were given an English textbook of their choice on the last day of the project. Due to scheduling issues, one class began during the week of February 7 and worked until the week of February 28. The other class began a few days later, and continued until the week of March 7.

The teacher completed the questionnaire at the beginning and end of the project; she completed the questionnaire on the computer and emailed it to one of the researchers; she was also interviewed every week either in person or on the phone by one of the researchers. The developers were sent their questionnaire as an email attachment; all four responded within about three weeks.

Analysis

Data from the questionnaires, weekly reflections, and test scores were analyzed quantitatively. Responses made to interview questions and as comments within the questionnaires were used to support the quantitative results or to illustrate interesting perspectives that might otherwise not have come to light. Because the purpose of the study was to better understand how *LEO 3* was viewed by the developers, teacher, and students in this one community college, only frequencies were used in the analysis, showing either the number or percentages of respondents to each possible value for a question. No inferential statistics were used.

The questionnaires containing essentially the same items were administered to both the teacher and the students at the beginning and end of the study in order to try and determine any changes in their perceptions. Analysis of the pre- and postforms indicated very little change and consequently only the results from the questionnaires administered at the end of the study are reported.

In the weekly reflections, one item for each skill area asked the students to write down what they could remember about that section of the lesson. For example, the students were asked to write down three things that they remembered about the story, and three grammar points that were presented and practiced in the unit. These responses were later recoded as "a lot," "some," or "nothing." For the week of February 7, one of the researchers coded the responses and developed a scoring rubric. In the subsequent weeks, two raters independently scored each item about how much was remembered about the content of each section. Table 5 shows that the raters agreed over 80% of the time.

Table 5
Percentage of Agreement in Coding Amount Remembered in Weekly Reflections

	February 14	February 21	February 28	March 7
N of decisions	264	328	304	152
N of changes	47	49	43	19
% of changes	18	15	14	12
% of agreement	82	85	86	88

In those cases in which the raters did not agree, they got together and discussed the student's answer and then one of the rater's changed her coding of the student's answer, so that in the end, there was 100% agreement for coding how much of the material the student remembered for that week for each section of the CALL unit: grammar, vocabulary, pronunciation, listening, speaking, reading, writing, and "On the Web." We compared students' responses across the four weekly reflections, but large changes in response patterns were not evident. Consequently, we averaged student responses to the questions on the weekly reflection by reporting only the mode, that is, each student's most frequent response. In no case were there two modes for a response across the four weeks. Two separate questions targeted working with "On the Web" and writing activities; however, the fact that the writing exercises were within the "On the Web" sections of the lessons confused the students, so these questions were not included in the analysis.

RESULTS

Results indicated various degrees of positive judgments for most of the six criteria for CALL with authenticity and practicality evaluated the lowest and positive impact the highest. These findings are interpreted in view of the degree of agreement among the three groups of stakeholders, with the least agreement appearing on meaning focus and authenticity.

Language Learning Potential

Since language learning potential was defined as the degree of opportunity present for beneficial focus on form (Chapelle, 2001), we sought evidence from the three groups on learning of grammar, vocabulary, and pronunciation, as well as overall improvement.

Developers and the Teacher

The results shown in Table 6 indicate that the developers thought that the grammar explanations were clear and that students' English would improve by using LEO 3. As one developer wrote, "The grammar is presented and practiced in short, learnable nuggets and is very clear."

Table 6
Developers' and Teacher's Opinions Regarding Language Learning Potential

Questions	Yes, very much	Somewhat	No, not at all
1. Do you think that students will be able to remember the			
grammar	D T	DD	0
vocabulary	D T	DD	0
pronunciation that they studied during the week?	DD	D T	0
2. Do you think that grammar explanations are clear?	DDDD T	0	0
3. Do you think that there are enough exercises in each unit?	DD	DD T	0
4. Do you think that students' English will improve from using *LEO*?	DDDD T	0	0

Note: Each D indicates the response from one developer; T indicates the teacher's responses.

However, they were more mixed in their opinions about whether students would remember the language points that they had studied and about whether there were enough exercises. One developer did not answer some of the questions; she wrote on the questionnaire that there were too many variables. The teacher thought that students would remember the language points, that the explanations were clear, and that students' English would improve as a result of using *LEO 3*. She also

pointed out, however, that the students needed more supplementary review practice to recycle, review, and reinforce. The teacher thought that there would be a higher level of retention if the course material was practiced with the ancillary materials such as the communication companion.[2]

Students

In order to determine what students remembered, they were asked to write down separately what they remembered about grammar, vocabulary, and pronunciation, and their responses were compared to the material in the lessons to score each response as "a lot," "some," or "nothing."

Figure 3 shows that over 30% of the students remembered a lot of grammar, and the rest of the students remembered some; none were coded as remembering nothing about the grammar they had studied. For vocabulary, almost half of the students remembered a lot, and over 40% remembered some; almost 10%, however, remembered nothing. For pronunciation, almost 65% of the students remembered some of the pronunciation points, but over 25 % of the students remembered none of them.

Figure 3
How Much of the Language Can You Remember?
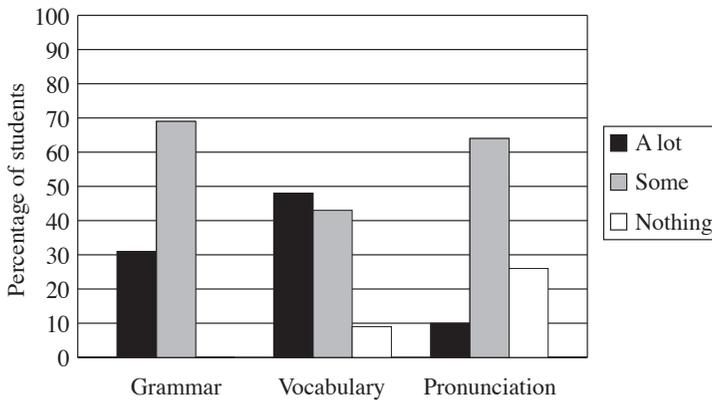
Figure 4 summarizes students' responses to questions about understanding explanations. Over 50% of the students reported that they understood the grammar and pronunciation explanations very well, but only a little over 30% said that they understood the vocabulary explanations very well. Referring to the "grammar coach," one student reported "The grammar is very clear because she was explaining good." Only one student reported that the explanations were not understandable.

Sixty percent of the students thought that they had sufficient practice, as shown in Figure 5. Several student comments in the interviews add to the survey results: "We get a lot of practice"; "There are a variety of exercises"; "The more exercises,

the more we learn"; "Because in this course I can practice English. I think practice is very important to learn English." Figure 6 shows that about 75% of the students thought that working on *LEO 3* would help them to improve their English; no one reported that *LEO 3* would not help them. In the interview, one student said "Of course I will improve with *LEO*. It is a good program."

Figure 4
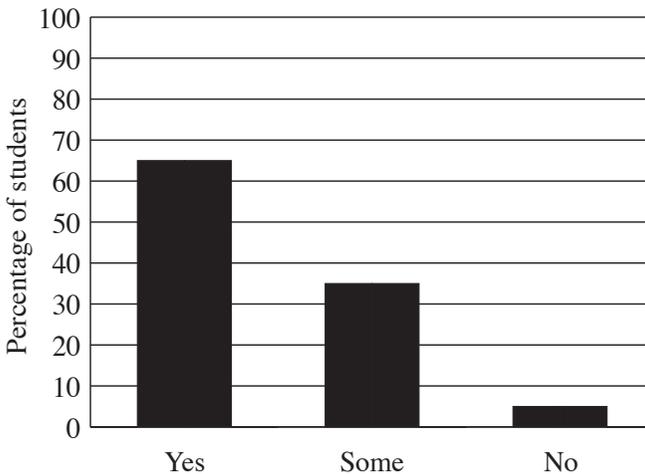Did You Understand the Explanations?
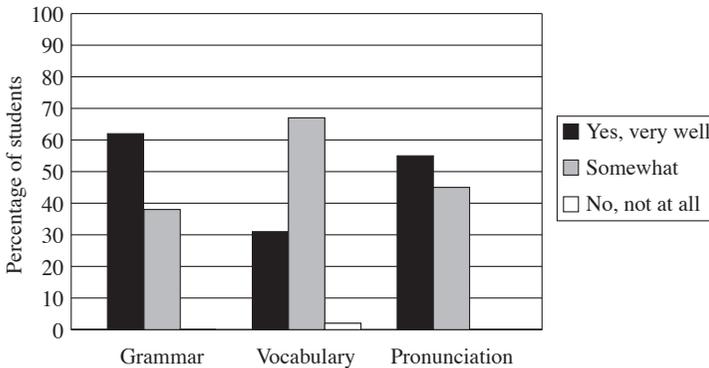
Figure 5
Did You Have Enough Practice with the Exercises?

Quiz scores were also examined to determine the degree to which students could be classified as having mastered the material in the units that they had studied. In Table 7, we see the students' performance on each of the 4 quizzes in *LEO 3*. A score of 70% was the recommended cut score which separated masters from nonmasters since this score frequently distinguishes between average and below average performance (or a grade of C vs. a grade of D) in American educational

settings. Because we desired that all students be classified as masters, these data do not provide as strong a support as we had hoped for evidence of language learning potential. Note that in Week 4, only the afternoon class completed the quiz; this accounts for the reduced number of students.

Figure 6
Will *LEO 3* Help You to Improve Your English?
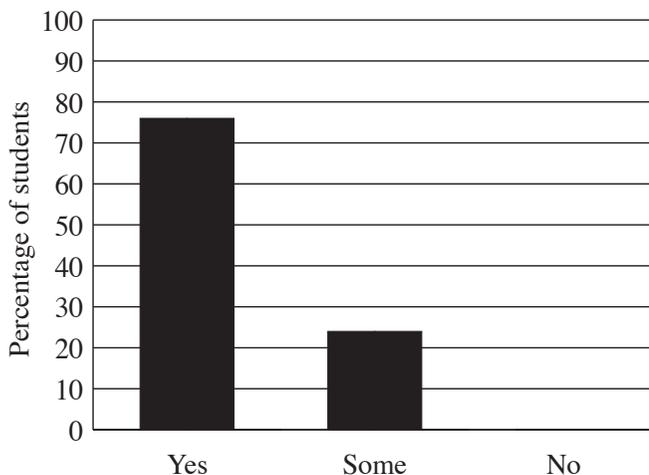
Table 7
Students' Performance on Quizzes 1-4

	Quiz 1	Quiz 2	Quiz 3	Quiz 4
Average score	75%	74%	80%	84%
Standard deviation	13%	15%	11%	9%
Number of students	42	41	38	20
Masters/nonmasters	29/13	28/13	31/7	19/1

Meaning Focus

Meaning focus was defined as the extent to which learners' attention is directed toward the meaning of the language. Four sections of each unit in *LEO 3* dealt more with understanding the meaning of the language rather than the form of the language. These sections included the video presentation of the dramatic story (listening comprehension), speaking, reading, and "On the Web."[3] For each of these sections, the developers and the teacher were asked if they thought the students would be able to understand (i.e., to follow) the content.

Developers and the Teacher

The developers indicated their intent to develop a high interest, dramatic story with high-interest readings so that students would be able to follow them. On the

questionnaire, the developers were of somewhat mixed minds when asked wheth-
er students would be able to follow (i.e., understand) the content, as illustrated in
Table 8. Three out of four developers thought that the students would understand
the story. Of the three developers who responded to questions about reading, "On
the Web," and speaking, two thought students would understand most of the read-
ing and speaking, but only one thought students would be able to follow most of
the "On the Web" activity. One developer indicated a commitment to meaning
focus with a comment about the story being less anchored to a grammar syllabus
than some materials. The teacher also thought that the students would understand
the story, the reading section, and the speaking section, but she was less confident
about their understanding the "On the Web" section. She commented that the stu-
dents seemed to enjoy and understand the story a lot.

Table 8
Developers' and Teacher's Opinions Regarding Focus on Meaning

Question	Yes, very much	Somewhat	No, not at all
1. Do you think that students will be able to follow the			
story	DDD T	D	0
reading	DD T	D	0
"On the Web"	D	DD T	0
speaking	DD T	D	0
that they studied during the week?			

Note: Each D indicates the response from one developer; T indicates the teacher's re-
sponses.

Students

Figure 7 shows data from the weekly reflections suggesting that most of the stu-
dents remembered something about each of the sections, particularly about the
story. These data are supported by their interview comments: "The story is inter-
esting and very clear for me to understand"; "I understood the story and every time
I had anxiety for the next story." We interpreted this last comment as meaning that
the student could not wait for the next story. Not everyone understood, though.
One student said, "I didn't understand everything, but I adjusted somehow." In
contrast to the story, almost 40% of the students did not remember anything about
the reading and the speaking sections.

Learner Fit

Learner fit was defined as the amount of opportunity for engagement with lan-
guage under appropriate conditions, given learner characteristics. We focused on
the match between the ability level of the students and *LEO 3*, the type of learners
for whom *LEO 3* was intended, and the difficulty level of the materials.

Figure 7
How Much of the Content Can You Remember?
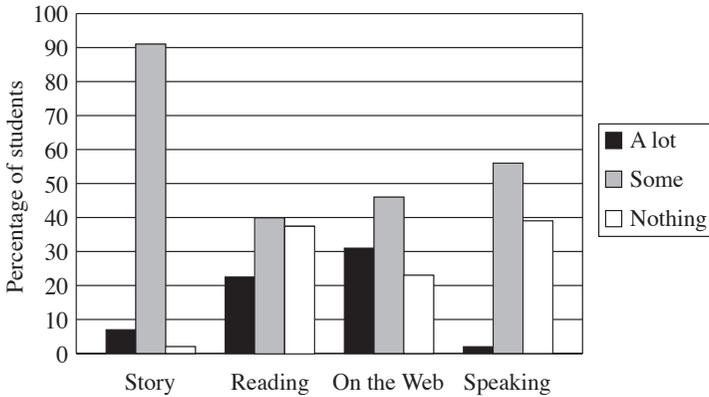
Developers

In Table 9, we see that the targeted level of the students was considered "intermediate" by two the of developers. This matches with the advertised level on the publisher's web site. Developers 1 and 3 reported rough equivalents of the Test of English as a Foreign Language (TOEFL) scores. Developer 1's range is often considered high-intermediate because a score of 500 is the lowest score required for undergraduate admissions at many universities in the USA. On the contrary, Developer 3's range of 400-450 is often considered low intermediate. Obviously, these labels are subjective; all developers agreed that the targeted level was intermediate, rather than beginning or advanced. The developers responded to the open-ended question about characteristics of the students in terms of young adults who are self-motivated, perhaps in school, and perhaps "economically well off."

Table 9
Developers' Opinions about the Students who would use *LEO 3*

Questions	Developer 1	Developer 2	Developer 3	Developer 4
1. What is the level of student you have targeted for *LEO 3*?	TOEFL 500-550	Intermediate level	TOEFL 400-440	Intermediate level
2. What kinds of students were these lessons designed for?	young adults/ adults in college, private language school, corporate program	young adult/ adult, ELT, self-directed, educated in L1, economically well off to own computer	students who have some experience in self-directed learning	

Teacher

The teacher's description of the students, as summarized in Table 10, places them at the lower end of the intended ability level. However, she felt that *LEO 3* was at an appropriate level of difficulty for her students since it was "somewhat challenging." As the students were beginning to work with *LEO 3*, the teacher said, "The course is definitely challenging for students; however most students are rising to the challenge and improving as they go." Accordingly, she described her students as motivated.

Table 10
Teacher's Description of Her Students

Questions	Responses
1. What is the level of your students' TOEFL range?	Students are in a level which is considered low intermediate; those who take the TOEFL at this level generally score below 450.
2. Do you think *LEO* was an appropriate level of difficulty?	Language was somewhat challenging … the fact that it was difficult made it more intresting.
3. Apart from language ability, how would you describe your students?	Motivated, eager to improve English and computer skills.

Students

Based on their scores on a standardized ESL test, these students had been placed in the intermediate level in the community college's curriculum. On the questionnaire, students were asked to rate their English as "excellent," "very good," "good," "fair," or "poor;" 3% said "excellent," 17% said "very good," 63% said "good," and 17% said "fair." In their weekly reflections, when students were asked if they already knew the grammar, vocabulary, and the pronunciation points that were taught, most of the students reported that they knew the language "somewhat," as shown in Figure 8. This same trend is evident in Figure 9, which illustrates weekly reflection responses to questions about the difficulty of the content in the story, reading, "On the Web," and speaking. Sixty percent or more of the students reported that the content was somewhat difficult.

Figure 8
Did You Already Know This Language?
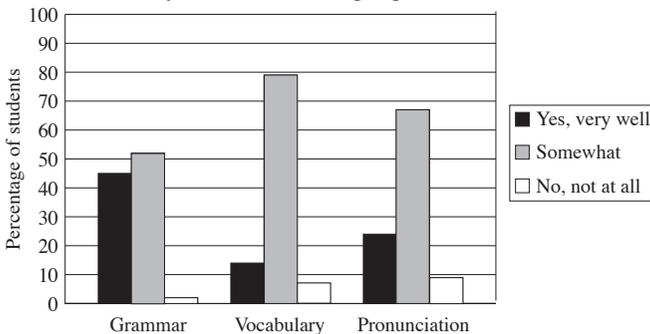
Figure 9
Was the Content Difficult?
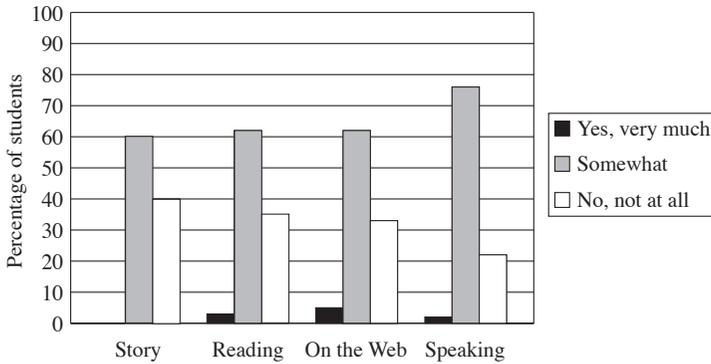
Students were directly asked about the overall difficulty of *LEO 3* on the questionnaire. Figure 10 shows that over 90% of the students thought that *LEO 3* was at a good level of difficulty—neither too easy nor too hard. In the interviews, although one student said, "Sometimes it was too fast and too hard," most students described the difficulty level as "appropriate," "good," or "cool."

Figure 10
How Difficult Was *LEO 3*?
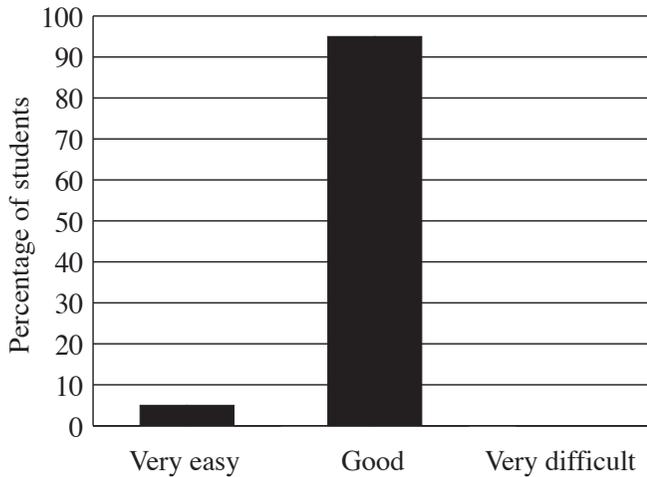
Authenticity

Authenticity was defined as the degree of correspondence between the learner activity and the target language activity of interest to the learners outside of class. The developers and teacher were asked if the language of *LEO 3* is what students will need, whereas the students were asked if the language of *LEO 3* is like the language they are exposed to outside of class.

Developers and the Teacher

Two of the developers thought that the language was needed outside of class, and one thought only somewhat (see Table 11). One developer wrote, "Can't say for sure what they'll need outside of class, but the course was designed to give students the language they could deal with at an intermediate level." As shown in Table 11, the teacher thought that the language of *LEO 3* would be needed outside of class. She wrote "The course includes exposure to both common everyday spoken English (expressions, functions, vocabulary, etc.) and more formal vocabulary in reading and 'On the Web' assignments as well as a few important academic skills such as summarizing. This skill should be taught directly in the course, and students might benefit from an additional section on academic skills and strategies for academic success."

Table 11
Developers' and Teacher's Opinions Regarding Language Used in *LEO 3* in Relation to Language in Real Life

Question	Yes, very much	Somewhat	No, not at all
1. Do you think the language students study in *LEO 3* is what they need outside of class?	DD T	D	0

Note: Each D indicates the response from one developer; T indicates the teacher's responses.

Students

The students' responses to questions about authenticity from their weekly reflections are summarized separately for the language and content topics in each of the sections of *LEO* in Figures 11 and 12. Over 60% of the students thought that the language in the grammar, vocabulary, and pronunciation sections was somewhat like the language they encountered outside of class. Almost 30% of the students thought that the grammar and pronunciation were very much like that outside of class. However, Figure 12 shows that very few students thought that they heard or read the language used in the story, reading, "On the Web," or speaking, very much outside of class. Except for speaking, over 30% of the students reported that they never heard this language outside of class. Overall though, between approximately 50% to 60% of the students said that they heard or read this language outside class sometimes.

   Comments from the students reflected this mixture of opinions. Some students said that they used the language of *LEO 3*: "I spoke a lot of English that I learned English in LEO"; "We can use in my life"; "I tell my father words like 'cheerful' and 'gossip' that I learned in LEO"; and "I use some of the words in my work." On the other hand, other students commented on the difference between LEO and the world outside the classroom: "LEO is correct, proper English. At home,

people don't use correct grammar. The English is short"; "We hear street language outside the classroom"; "I don't hear like that on the street. This is too fast and we have to catch and follow the words."

Figure 11
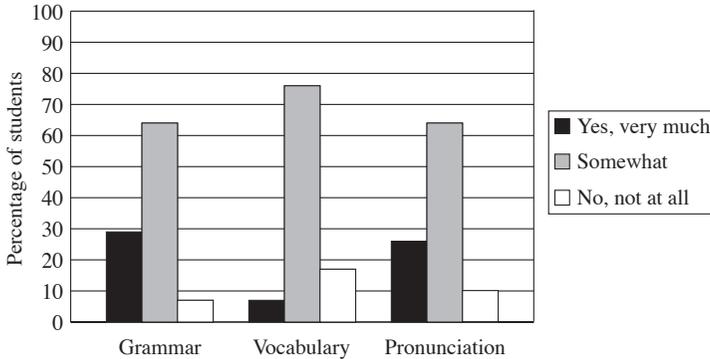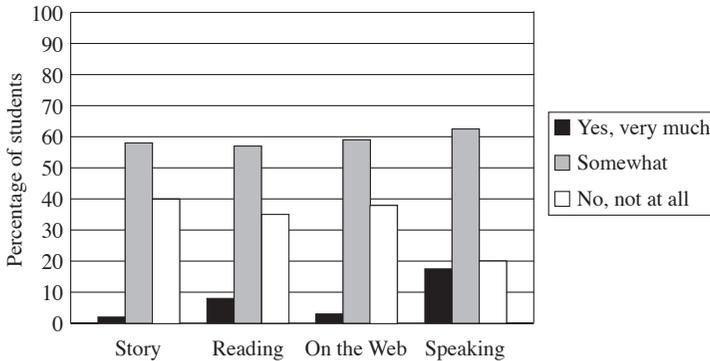Was the Language of *LEO 3* Like That Outside of Class?

Figure 12
Did You Hear or Read This Language Outside of Class?

Positive Impact

Positive impact was defined as the positive effects of the CALL activity on those who participate in it, so we asked developers and teachers whether or not students (would) like *LEO 3* and want to use *LEO 4*.

Developers and Teacher

Perhaps not surprisingly, the developers were unanimous in their perceptions as to students' positive attitudes toward *LEO 3*, as shown in Table 12. The developers commented that they thought students would find that learning can be fun, that students would enjoy the story because it is suspenseful, and that students could learn about life in the US, computer use, body language, learning strategies, and

gaining control over their own learning. Similarly, the teacher thought that students liked using *LEO 3* and that they would want to use *LEO 4*. She commented that students were excited and engaged and that students also gained facility with computers by working on *LEO 3*.

Table 12
Developers' and Teacher's Opinions Regarding Students' Attitudes toward *LEO 3*

Questions	Yes, very much	Somewhat	No, not at all
1. Do you think students will like using *LEO 3*?	DDDD T	0	0
2. Will students want to use *LEO 4* after having used *LEO 3*?	DDDD T	0	0

Note: Each D indicates the response from one developer; T indicates the teacher's responses.

Students

The students were asked questions in their weekly reflections about how they liked the material in that week's lesson, and they were asked more general questions in the questionnaires. Figure 13 shows that the majority of the students enjoyed the language practice very much, and a sizable minority (between about 35%-45%) said that they enjoyed the practice somewhat. When asked if they enjoyed the content of the lessons, students' opinions were evenly split between very much and somewhat, as shown in Figure 14. Very few students reported that they did not like the practice or content. These opinions are reinforced by the results in Figure 15, which show almost 80% of the students reported that they liked using *LEO 3* and by the results in Figure 16 showing that over 80% of the students said that they would like to use *LEO 4*.

Figure 13
Did You Enjoy the Practice?
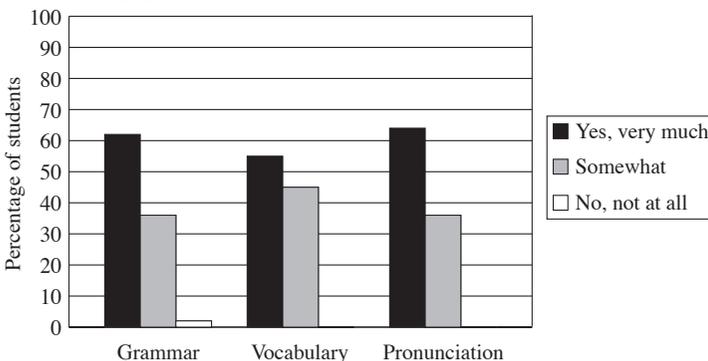
Figure 14
Did You Enjoy the Content?
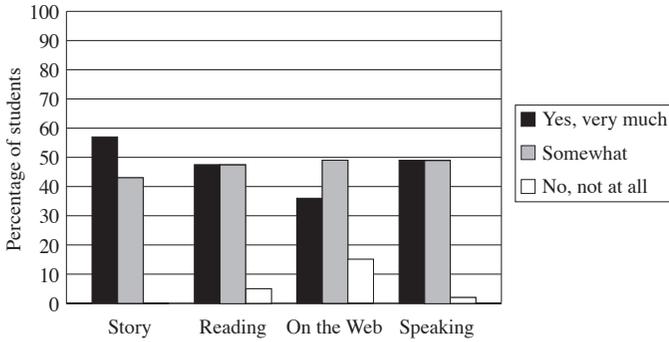
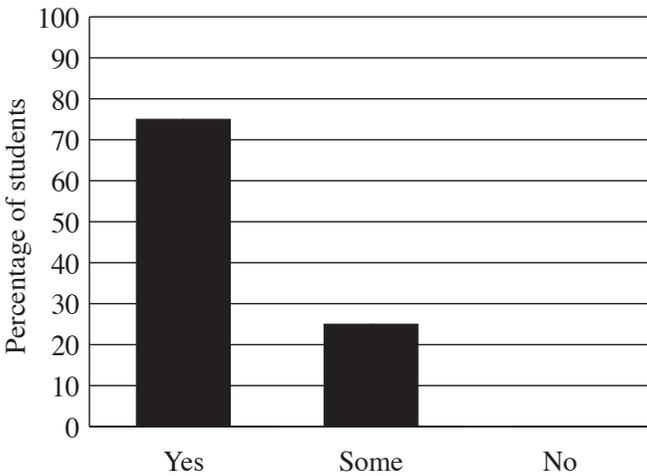

Figure 15
Did You Like Using *LEO 3*?
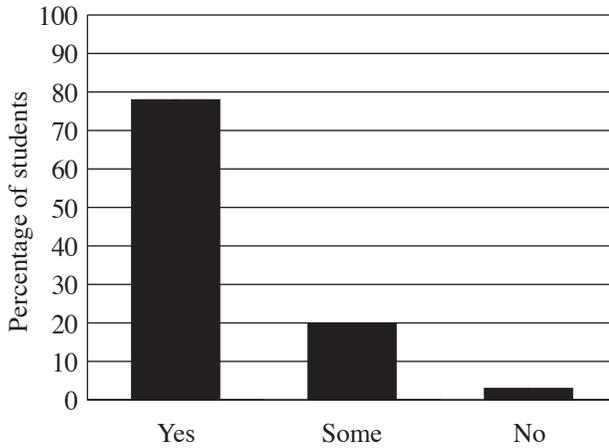
Figure 16
Would You Like to Use *LEO 4* after *LEO 3*?
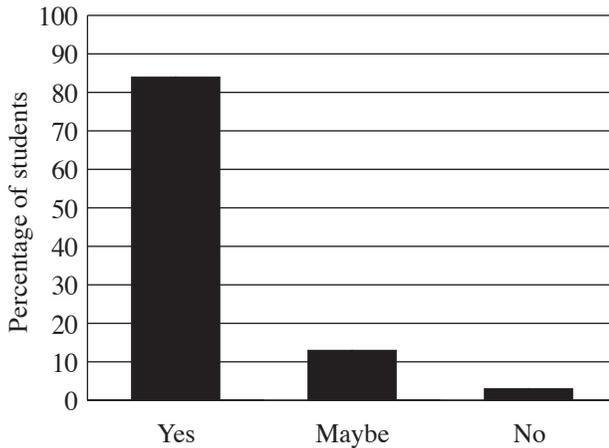
In the interviews, students provided comments regarding the impact of using *LEO*: "I liked the quizzes too much because you can know your performance"; "I can concentrate more"; "My mind has opened in order to improve my English"; "I could learn other things like American culture, computer use, web site, grammar, and vocabulary"; "Sometimes good, sometimes bad, but I can't do it from my home"; "Sometimes I go out in the street and pay attention more, but there are a lot of immigrants, so it's not good to learn"; and lastly, "I enjoy the computer, but I also like having a teacher."

Practicality

Practicality was defined as the adequacy of resources to support the CALL activity. The three groups responded to questions about the computer interface, the labs, time, and levels of frustration.

Developers

As depicted in Table 13, for the most part the developers thought that the computer interface would be easy to use and that the students would not need help. Their comments included the following: "The UI is simple"; "It is necessary to have some orientation and guidance to use it effectively"; and "The role play might be confusing at first." They were not as confident that the computers and the labs would be able to run *LEO 3* without problems, and they thought that students would get frustrated at times. As one developer wrote, "The technology needs to be right." The views of the developers concerning the amount of teacher time varied widely; for this teacher with about 45 students, developers responses ranged from 0 to 45 hours per week! The developers exhibited much more agreement about the amount of student time that would be required per week.

Teacher

As seen in Table 14, the teacher thought that the computers and the lab were of sufficient quality to run *LEO 3* and that students had enough time in the lab to work on *LEO*, although she stated that the students needed more computer time. She also wrote that the class needed workbooks to practice at home. The teacher thought that the interface needed some explanation and that students sometimes got frustrated. She wrote, "I have some practical suggestions to improve the user interface of the role play." During interviews with the teacher, this frustration was also mentioned. When asked if she had done the orientation as had been recommended, the teacher said no. One of the developers firmly believed that the teacher's frustration could have been eliminated completely with some orientation and spot training.

The teacher spent about 5 to 10 hours correcting and checking assignments. During interviews and in the comments on her questionnaire, she expressed displeasure stating that the time she had to spend on *LEO 3* exercises was excessive.[4] Finally, she thought that most of her students finished a unit in about 3 hours.

Table 13
Developers' Opinions Regarding the Practicality of Using *LEO 3*

Questions	Yes, very much	Somewhat	No, not at all
1. Will the interface be easy to use?	DDD	0	0
2. Will students be able to work without help?	DDD	D	0
3. Will the the computers and the lab be of sufficient quality?	D	DD	0
4. Will the students get frustrated?	0	DDD	0
5. How much time will a teacher spend a week?	• 30-60 minutes per student, assuming writing/speaking submits and chat • 30 minutes class time; an additional hour per 12 students • 0-4 hours		
6. How much time will it take a student to finish a unit?	• 3, 1 hour sessions • 4-6 hours • 2-3 hours		

Note: Each D indicates the response from one developer.

Table 14
Teacher's Opinions Regarding the Practicality of Using *LEO 3*

Questions	Yes, very much	Somewhat	No, not at all
1. Was the interface easy to use?		T	
2. Did your students have sufficient time in the computer lab?	T		
3. Were the computers and the lab of sufficient quality?	T		
4. Did the students get frustrated?		T	
5. How much time do you spend a week outside of class?	• 2-4 hours on speaking assignments • 2-4 hours checking progress reports and reading emails • 1-2 hours setting up Discussion Board and chats and looking at writing		
6. How much time does it take most of your students to finish a unit?	• about 3 hours		

Note: Each T indicates the teacher's responses.

Students

Figure 17 shows that over 70% of the students thought they understood the directions, but only a little over 40% of the students reported that they did not need help when they were using *LEO 3* (see Figure 18). Their comments revealed that this was particularly true when they were first using *LEO*: "When I started LEO, I needed to ask my teacher more"; "Role playing was confusing at first, but then it got easier"; "The teacher was always available to help"; and "The directions were good." About half of the students felt that they were spending an appropriate amount of time using *LEO 3*, but over 40% would have liked to spend more time, as seen in Figure 19.

Figure 17
Are You Able to Understand the Computer Directions in *LEO 3*?

Figure 18
Did You Need Help Using *LEO 3*?
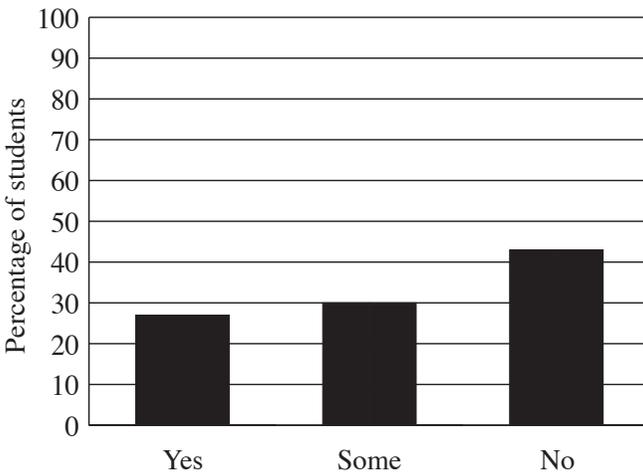
Figure 19
Would You Like to Spend More or Less Time Using *LEO 3*?
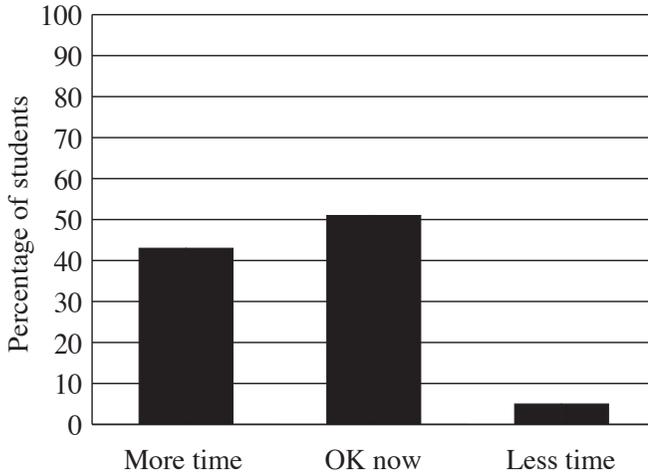
As one student said, "We needed more time and a chance to practice at home." Another student stated: "We couldn't use computers whenever we wanted." About 85% of the students reported that they spent between 1-3 hours completing a unit, as illustrated in Figure 20.

Figure 20
How Long Does It Take You to Finish One Unit?
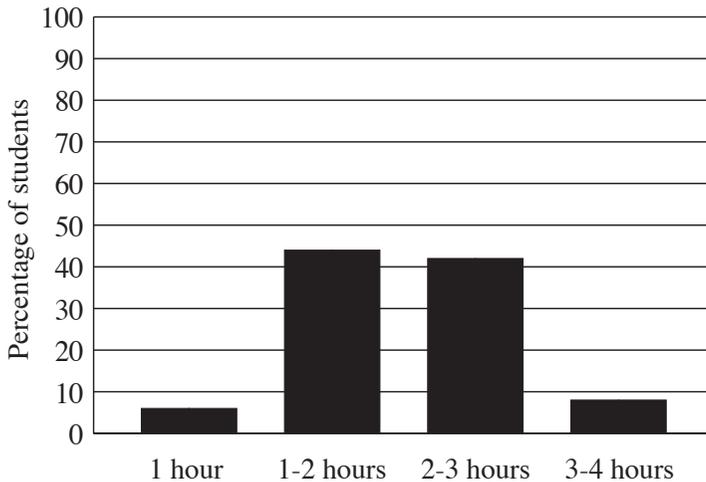
DISCUSSION

Table 15 summarizes our evaluation of the extent to which the developers, the teacher, and the students provided responses that would support a positive evaluation of *LEO 3* for this context. It also indicates the level of agreement among the three sources and synthesizes the responses to result in an overall evaluation. Level of agreement was estimated on a 5-point scale: excellent, good, average, fair, and poor. If all three groups gave the same rating on each point, their agreement was "excellent;" if the agreement levels across points were sometimes excellent and sometimes average, agreement was "good;" if 2/3 agreed on all points, their agreement was "average;" if the agreement levels were sometimes good and sometimes poor, we described it as "fair;" if there was no agreement, we described it as "poor."

Language Learning Potential

Five of the 13 evaluations of language learning potential were positive (+). In 5 other cases, the responses (+/~) were positive but not as strong. In the remaining 3 cases, the responses (~) were neutral. No negative evaluations were made in any of the categories. We judged the level of agreement among the three groups to be good because no strong disagreements were evident. Table 15 illustrates only one case in which there was complete agreement: Developers, the teacher, and the students were all of the opinion that working with this CALL material would result in students' improving their English. For the other three questions, two out of the three groups agreed with each other. In no case was there serious disagreement among the groups. In our opinion, this pattern of responses across several questions provided empirical support for the argument that *LEO 3* had good language learning potential in this situation.

Meaning Focus

The developers and the teacher agreed that the students would understand some parts of the content better than other parts. The responses in the weekly questionnaires revealed that most of the students could remember some of the content, and about equal percentages of the students remembered either a lot or nothing. No group responded as desired; still, these data do provide moderate support that this CALL material encouraged a focus on meaning. In hindsight, there was an obvious problem with the way in which we operationalized meaning focus. We really asked different questions when we asked developers and teacher if students would understand and then asked the students what they remembered. With this as the only sources of evidence about meaning focus, agreement was only average. In future versions of the instruments, students should be asked about their understanding in addition to being asked to display what they remembered.

Table 15
Summary of Responses for the Six Criteria for CALL Quality

Criteria	Questions (desired response)	Three evaluation perspectives			Agreement	Summary
		Developers	Teacher	Students		
Language learning potential	Remember language (yes)	+/~	+/~	~	Good	Good
	Explanations clear (yes)	+	+	+/~		
	Enough exercises (yes)	+/~	~	+/~		
	Improve English (yes)	+	+	+		
	Quiz scores (masters)			~		
Meaning focus	Remember/understand content (yes)	+/~	+/~	~	Average	Good
Learner fit	Ability level (intermediate)	+/~	+/~	+/~	Good	Good
	Learner characteristics (young adult, self-motivated)	+	+	+/~		
	Appropriate difficulty (somewhat difficult)		+	+		
Authenticity	Language used is needed outside/like that used outside (yes)	+/~	+/~	~/-	Average	Average
Positive impact	Students like LEO 3 (yes)	+	+	+/~	Good	Excellent
	Students want to use LEO 4 (yes)	+	+	+		
Practicality	Interface easy to use without help (yes)	+	~	+/~	Good	Average
	Labs/computers OK (yes)	~	+	+		
	Sufficient time (yes, about 3 hours/week)	+	+	+/~		
	Students frustrated (no)	~	~	~		
	Teacher time (about 4 hours/week)	-	-			
Overall					Good	Good

Note: + group responded as desired; +/~ desired response mixed with some/somewhat; ~ response neither desired nor undesired-some/somewhat; ~/- undesired response mixed with some/somewhat; - group responded opposite to what was desired.

Learner Fit

Looking at Table 15, we see that this CALL material seems to have been a good fit for these learners. In eight cases in which evaluations were made, four of the times the responses were what was desired to provide strong evidence (+), and four of the times the responses were close, but mixed (+/~).The level of agreement among our three groups was good here, indicating good learner fit overall.

Authenticity

The operational questions for authenticity suffered from a similar problem to those for meaning focus. The developers and teacher were asked if the CALL material's language was needed outside of class, whereas the students were asked if the language was like what they heard or read outside of class. However, this may not account for the difference of opinions we see reflected in the responses, as shown in Table 15, which resulted in an "average" level of agreement. Whereas the developers and the teacher agreed that students needed a lot of the kind of language that was presented in *LEO 3*, the students thought that some, but not a lot of *LEO 3*'s language was what they were exposed to when they left the classroom. Considering the definition of authenticity, in future versions the question for the developer and teacher might be worded more closely to the question that the students responded to. At the same time, clearly both perspectives on language use out of the classroom, what learners need versus what they are exposed to, are relevant, and therefore a more complex conceptualization of authenticity may also be worth considering.

Positive Impact

Developers and the teacher agreed that students would like to use this CALL material, and about 90% of the students liked working with it very much or somewhat. All three groups reported that after having worked with *LEO 3*, it would be desirable to work with *LEO 4*. Responses indicated good agreement among the groups regarding positive impact. In 5 of the 6 evaluations, the responders agreed with the desired response (+) which provides strong evidence for the positive impact of this CALL material for these particular students.

Practicality

In the 14 decisions that were made regarding the practicality of *LEO 3* in this situation, 5 responses were what we desired (+); in 2 cases the responses were close to desired (+/~); in 5 cases the responses were not as desired but were not negative either (~); but in 2 cases the responses were negative (-). On a positive note, the estimated student time seems reasonable although it should be noted that the teacher assigned the "On the Web" activities via Blackboard so that the students could work on them at home. It seems that students needed help when beginning to use this CALL material and that they exhibited a certain level of frustration as they learned to use the interface. However, once over this initial hurdle, frustra-

tion as well as the need for help diminished. The publisher did provide teacher training for the teacher and encouraged her to use the online orientation to train and educate the students, but the teacher did not follow the publisher's recommendation. All in all though, these results do not allow a positive evaluation in terms of practicality.

The responses about practicality in Table 15 show some agreement among groups, but not as much as we have seen for the other criteria. Indeed, the only two questions which show agreement revealed that the students do get somewhat frustrated at times and that the amount of teacher time that was envisioned (4 hours per week) was impractical. Although the developers were not sure that students' computers and schools' computer labs would be of suitable quality, in this situation both the teacher and students agreed that they were.

Summary

The evaluation of overall appropriateness of the CALL material is based on the results obtained for each of the six criteria separately. The quality of this evaluation rests in part on the level of agreement among the stakeholders, which was good, although not excellent. This level of agreement should give some confidence about the use of the summary of their collective decisions to arrive at an overall evaluation of the CALL material for this particular teacher and her group of students.

CONCLUSION

Based on data from developers, the teacher, and the students, we conclude that *LEO 3* exhibited desirable qualities which were appropriate for these two community college ESL classes. The CALL material had good language learning potential, meaning focus, and learner fit, as well as having an excellent positive impact in this particular setting. These findings, of course, are bound to the context we have studied and are therefore useful for the situation-specific argument that the teacher may wish to make for using the materials again or for discussing them with other teachers working within a similar setting. Yet, as Jaeger (1993) pointed out in his discussion of external validity:

> Even in the most immediately focused evaluation study, the interest of the researcher is rarely confined exclusively to the subjects on whom he or she has collected data, in the specific setting and period of time when the data were collected. More likely, the researcher will draw conclusions that carry an implicit, if not explicit, assumption that the results of the study can be generalized. The nature of the generalization made or implied will vary substantially in different studies, but often extends to "future times," or "subjects like these," or "settings or contexts like these" ... . Even when a researcher interprets his or her results only for the subjects, time period, and context used in a particular study, consumers of the results inevitably engage in generalization. (pp. 122, 124)

In other words, despite the fact that our findings were based on two classes of ESL adults attending a community college in New York, with all of the particular circumstances of the research setting, we would suggest that the results are probably relevant to other potential uses of this CALL material. Consistent with the practices of case studies, we have attempted to describe the case in sufficient descriptive narrative so that readers can vicariously experience these happenings, and draw their own conclusions (Stake, 1994, p. 242). In view of the probability that each teacher cannot participate in a research project such as this to determine the appropriacy of CALL materials for his/her class, ideally readers can assess the transferability of the circumstances and findings to other situations.

Although we have used numeric summaries of opinions rather than narratives for the most part, we hope that a teacher might be able to look at a study about CALL use, such as this one, and determine the degree to which the students and the context is similar to or different from the situation he/she is in. To the degree that the situations are similar, the teacher might be willing to interpret the findings as potentially relevant for his/her situation. Moreover, others in a variety of contexts may find our experience helpful in using the criteria and instruments in this study as a means of framing research on CALL in a manner that does not rely on CALL-classroom comparisons.

NOTES

[1] Originally distributed via the Web, this courseware is now distributed on CD-ROMs. It has been renamed *Longman English interactive*.

[2] The *Longman English interactive 3 communication companion* is a supplementary text that instructors can use in the classroom to provide communicative, face-to-face activities. In May 2005, another supplementary, classroom-based text, *Longman English interactive 3 activity and resource book*, was published to provide further practice and review of the CD-ROM course. Feedback from students in this project inspired these new workbooks.

[3] "On the Web" activities are no longer included in the CD-ROM format of these CALL materials. However, they are included in the *Teacher's Guide*s, available online, and they include reproducible handouts with writing assignments for the students.

[4] The online submissions are no longer part of the CD-ROM materials, and so this aspect of teacher time has been eliminated from the CALL materials. However, the *Teacher's Guides* contains suggestions for the teacher to assess speaking in the course, as well as additional activities for speaking and listening.

REFERENCES

Bachman, L., & Palmer, A. (1996). *Language testing in practice*. Oxford: Oxford University Press.

Brown, H. D. (2000). *Principles of language learning and teaching* (4th ed.). New York: Longman.

Burston, J. (2003). Proving IT works. *CALICO Journal, 20* (2), 219-226.

Chapelle, C. A. (2001). *Computer applications in second language acquisition*. Cambridge, UK: Cambridge University Press.

Chapelle, C.A. (2003). *English language learning and technology: Lectures on teaching and research in the age of information and communication*. Amsterdam: John Benjamins Publishing.

Crookes, G., & Chaudron, C. (2001). Guidelines for language classroom instruction. In M. Celce-Murcia (Ed.), *Teaching English as a second or foreign language* (3rd ed.) (pp. 29-42). Boston: Heinle & Heinle Thomson Learning.

Dunkel, P. (1991). The effectiveness research on computer-assisted instruction and computer-assisted language learning. In P. Dunkel (Ed.), *Computer-assisted language learning and testing: Research issues and practice* (pp. 5-36). New York: Newbury House.

Jaeger, R. (1990). *Statistics*. Thousand Oaks, CA: SAGE Publications.

Larsen-Freeman, D. (2000). *Techniques and principles in language teaching* (2nd ed.). New York: Oxford University press.

Lightbown, P., & Spada, N. (1999). *How languages are learned* (2nd ed.). New York: Oxford University Press.

Mellow, J. D. (2002). Toward principled eclecticism in language teaching: The two-dimensional model and the centering principle. *TESL_EJ Online Journal*. Retrieved August 6, 2003 from http://www.kyoto-su.an.jp/information/test-ej/ej20/al.html

Pederson, K. M. (1987). Research on CALL. In W. F. Smith, (Ed.), *Modern media in foreign language education: Theory and implementation* (pp. 99-132). Lincolnwood, IL: National Textbook Company.

Pienemann, M. (1985). Learnability and syllabus construction. In K. Hyltenstam & M. Pienneman (Eds.), *Modeling and assessing second language acquisition* (pp. 23-75). Clevedon, Avon: Multilingual Matters.

Rost, M., & Fuchs, M. (2004). *Longman English interactive*. New York: Pearson Education.

Savignon, S. (2001). Communicative language teaching for the twenty-first century. In M. Celce-Murcia (Ed.), Teaching English as a second or foreign language (3rd ed.) (pp. 13-28). Boston: Heinle & Heinle Thomson Learning.

Skeehan. P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.

Stake, R. E. (1994). Case studies. In N. Denzin & Y. Lincoln (Eds.), *Handbook of qualitative research* (pp. 236-247). Thousand Oaks, CA: SAGE Publications.

Yin, R. (1994). *Case study research: Design and methods*. Thousand Oaks, CA: SAGE Publications.

APPENDIX A
Developers' Questionnaire

A Questionnaire about LEO 3

In order to compare the perspectives of developers, teachers, and students, we appreciate your taking the time to fill out this questionnaire. Please fill this out online, save it, and e-mail it back to Joan.Jamieson@nau.edu as an attachment.

Part One.	**Please consider these questions about LEO 3. Click on the shaded box and type in your comments.**
What is the level of student you have targeted for LEO3? Could you make a guess at an appropriate TOEFL range? ooooo¶ ¤	
What kinds of students were these lessons designed for? ooooo¶ ¤	
What aspects of the lessons would you expect to be especially appropriate for the targeted learners? Why? ooooo¶ ¤	
How much of each unit do you think students should do? ooooo¶ ¤	
How long do you think it will take a student to finish one unit? ooooo¶ ¤	
What do you think students will learn from LEO besides English? ooooo¶ ¤	
How much teacher time will be involved each week? ooooo¶ ¤	

Part Two.	**In this section, please consider the questions. Click on the box that most closely reflects your opinion. Add any comments you think will clarify your opinion.**			
Do you think that students' English will improve from working on LEO? **Why?** ooooo¤	Yes, very much ☐	Somewhat ☐	No, not at all ☐	
Do you think the students will be able to remember the **grammar** that they studied during the week? Comments: ooooo¤	Yes, very much ☐	Somewhat ☐	No, not at all ☐	

Do you think the students will be able to remember the **vocabulary** that they studied during the week? Comments: ○○○○○¤	Yes, very much ☐	Somewhat ☐	No, not at all ☐
Do you think the students will be able to remember the **pronunciation** that they studied during the week? Comments: ○○○○○¤	Yes, very much ☐	Somewhat ☐	No, not at all ☐
Do you think the students will be able to remember the **speaking phrases** that they studied during the week? Comments: ○○○○○¤	Yes, very much ☐	Somewhat ☐	No, not at all ☐
Do you think the students will be able to follow the **reading topic** that they studied during the week? Comments: ○○○○○¤	Yes, very much ☐	Somewhat ☐	No, not at all ☐
Do you think the students will be able to follow the **"On the Web" activity** that they studied during the week? Comments: ○○○○○¤	Yes, very much ☐	Somewhat ☐	No, not at all ☐
Do you think the students will be able to follow the **writing topic** that they studied during the week? Comments: ○○○○○¤	Yes, very much ☐	Somewhat ☐	No, not at all ☐
Do you think the students will be able to follow **the story** that they studied during the week? Comments: ○○○○○¤	Yes, very much ☐	Somewhat ☐	No, not at all ☐
Do you think that the grammar explanations are clear? Comments: ○○○○○¤	Yes, very much ☐	Somewhat ☐	No, not at all ☐
Do you think that there are enough exercises in each unit? Comments: ○○○○○¤	Yes, very much ☐	Somewhat ☐	No, not at all ☐
Do you think that the language students study in LEO is what they need outside of class? Comments: ○○○○○¤	Yes, very much ☐	Somewhat ☐	No, not at all ☐
Do you think that students will like using LEO? Comments: ○○○○○¤	Yes, very much ☐	Somewhat ☐	No, not at all ☐

Will students want to use LEO 4 after having used LEO 3? Comments: ○○○○○¤	Yes, very much ☐	Somewhat ☐	No, not at all ☐
Do you think the computer interface will be easy for students to use? Comments: ○○○○○¤	Yes, very much ☐	Somewhat ☐	No, not at all ☐
Do you think that the computers will be of sufficient quality and the labs will be accessible to students? Comments: ○○○○○¤	Yes, very much ☐	Somewhat ☐	No, not at all ☐
Do you think students might get frustrated? Where in the lessons? Comments: ○○○○○¤	Yes, very much ☐	Somewhat ☐	No, not at all ☐
Do you think that student will be able to use LEO by themselves? Comments: ○○○○○¤	Yes, very much ☐	Somewhat ☐	No, not at all ☐
Part Three. **Please add any other comments you would like to make.**			

APPENDIX B

Teacher's Questionnaire

Questionnaire #2 about LEO 3

We appreciate your taking the time to fill out this questionnaire, which we will use to help understand the teacher's perspective on LEO 3.

Part One. **Please consider these questions about LEO 3. Type in your comments in the shaded box.**
Do you see some aspects of LEO3 that were well suited to your students and others that were not? Please explain. ○○○○○¶ ¤

Did you tell students that they have to do some parts of LEO (for example, Listening) as opposed to letting them do whatever they want? If so, which parts? ○○○○○¶
¤

How long do you think it took a student to finish one unit? ○○○○○¶
¤

What do you think students learned from LEO besides English? ○○○○○¶
¤

Do you think that LEO 3 was at the appropriate level of difficulty for your students? Why or why not? ○○○○○¶
¤

How much time did you devote to LEO in your weekly syllabus? ○○○○○¶
¤

How much time did you spend outside of class on LEO related assignments? ○○○○○¶
¤

Part Two.
In this section, please consider the questions. Put an X in the box that most closely reflects your opinion. Type any comments you think will clarify your opinion in the shaded box.

	Yes, very much	Somewhat	No, not at all
Do you think that students were able to use LEO by themselves? Comments: ○○○○○¤	☐	☐	☐
Do you think that your students' English improved from working on LEO? Why or why not? ○○○○○¤	☐	☐	☐
Do you think the students remembered the **grammar** that they studied? Comments: ○○○○○¤	☐	☐	☐
Were the grammar points in LEO the same ones that you covered in class? Comments: ○○○○○¤	☐	☐	☐
Do you think the students remembered the **vocabulary** that they studied? Comments: ○○○○○¤	☐	☐	☐

	Yes, very much	Somewhat	No, not at all
Do you think the students remembered the **pronunciation** that they studied? Comments: ○○○○○¤	☐	☐	☐
Do you think the students remembered the **speaking phrases** that they studied? Comments: ○○○○○¤	☐	☐	☐
Do you think the students understood the **reading topic** that they studied? Comments: ○○○○○¤	☐	☐	☐
Do you think the students understood the **"On the Web" activity** that they studied? Comments: ○○○○○¤	☐	☐	☐
Do you think the students understood the **writing topic** that they studied? Comments: ○○○○○¤	☐	☐	☐
Do you think the students understood **the story** that they studied? Comments: ○○○○○¤	☐	☐	☐
Did you ever talk about the story in class or hear your students talking about it? Comments: ○○○○○¤	☐	☐	☐
Did you think that the grammar explanations were clear? Comments: ○○○○○¤	☐	☐	☐
Did you think that there were enough exercises in each unit? Comments: ○○○○○¤	☐	☐	☐
Do you think that the language students studied in LEO is what they need outside of class? Comments: ○○○○○¤	☐	☐	☐
Do you think that students liked using LEO? Comments: ○○○○○¤	☐	☐	☐
Do you think students would like to use LEO 4 after having used LEO 3? Comments: ○○○○○¤	☐	☐	☐

Do you think the quizzes and module tests were helpful? Comments: ○○○○○¤	Yes, very much ☐	Somewhat ☐	No, not at all ☐
Did you make use of the chat and discussion board activities? Comments: ○○○○○¤	Yes, very much ☐	Somewhat ☐	No, not at all ☐
Do your students have sufficient time when they can use the computer lab to work on LEO? Comments: ○○○○○¤	Yes, very much ☐	Somewhat ☐	No, not at all ☐
Do your students ever get frustrated with LEO? If so, why? Comments: ○○○○○¤	Yes, very much ☐	Somewhat ☐	No, not at all ☐

Part Three.
Please add any other comments you would like to make.

○○○○○¶
¶
¶
¶
¶
¤

APPENDIX C

Students' Questionnaire

Name: _____

Part One.
Please read these questions about LEO 3. Write your answer or put an X in the box.

What do you think is the best way for you to study English?

¤

How old are you? _____ Are you a male ☐ or a female ☐?

What is your first language? _____
How many languages do you know? _____

| How long did it take you to finish one unit? |
| less than 1 hour ☐ 1-2 hours ☐ 2-3 hours ☐ 3-4 hours ☐ 4-5 hours ☐ |
| more than 5 hours ☐ |

| What did you learn from LEO besides English? |

| How difficult was LEO 3 for you? |
| Very easy (you already knew it) ☐   Good ☐ |
| Very difficult (you did not understand) ☐ |

| Did you want to spend more or less time using LEO? |
| more time ☐   you spent a good amount of time ☐   less time ☐ |

| What is your English level? Excellent ☐   Very good ☐   Good ☐   Fair ☐ |
| Poor ☐ |

Part Two.
Please read the questions. Put an X in the box that is closest to your opinion. Write any comments you have.

	Yes	Some	No
Did you need help to use LEO? Comments:	☐	☐	☐
Did your English improve from working on LEO? Why or why not? Comments:	☐	☐	☐
Did you try to understand the story in LEO? Comments:	☐	☐	☐
Did you understand the grammar explanations in LEO? Comments:	☐	☐	☐
Did you like using the Discussion Board? Comments:	☐	☐	☐

	Yes	Some	No
Did you like the quizzes? Comments:	☐	☐	☐

	Yes	Some	No
Were the computer directions easy to understand in LEO? Comments:	☐	☐	☐

	Yes	Maybe	No
Did you get enough practice with the grammar, vocabulary, and pronunciation in LEO? Comments:	☐	☐	☐

	Yes	Some-what	No
Did the quizzes help you to learn English? Comments:	☐	☐	☐

	Yes	Maybe	No
Is what you learned in LEO useful outside of class? Comments:	☐	☐	☐

	Yes	Some-what	No
Did you like using LEO? Comments:	☐	☐	☐

	Yes	Maybe	No
Would you like to use LEO 4 after LEO 3? Comments:	☐	☐	☐

APPENDIX D

Students' Weekly Reflection

Name: _____

Directions: Please answer the questions below about LEO. Write short answers on the lines about what you did this week. If you do not remember, do not write anything on the line. For the other questions, circle either "Yes, very much," "Somewhat," or "No, not at all"; we want to know your opinions.

1. **What unit or units of LEO3 did you work on this week? Circle your answer.**

    Unit A1   Unit A2   Unit A3   Unit A4

    Unit B1   Unit B2   Unit B3   Unit B4

    Unit C1   Unit C2   Unit C3   Unit C4

2. **The Story**

2.1 What happened in the story? Write 1-3 sentences if you remember.

    _____

    _____

    _____

Circle your answers.

2.2 Was listening to the story   Yes, very much   Somewhat   No, not at all
     difficult?

2.3 Did you enjoy the story?   Yes, very much   Somewhat   No, not at all

2.4 Do you hear the English in   Yes, very often   Sometimes   No, never
     the story outside of class?

3. **Grammar**

3.1 What grammar points did you study this week? Write them if you remember.

    _____

    _____

Circle your answers.

3.2 Did you already know this   Yes, very much   Somewhat   No, not at all
     grammar?

3.3 Did you understand the   Yes, very much   Somewhat   No, not at all
     grammar explanations?

3.4 Did you enjoy the prac-   Yes, very much   Somewhat   No, not at all
     tice?

3.5 Do you hear or read this   Yes, very often   Sometimes   No, never
     grammar outside of class?

4. Vocabulary
4.1 What vocabulary did you study this week? Write 3 words if you remember.

_____

_____

Circle your answers.

4.2 Did you already know these words?	Yes, very well	Somewhat	No, not at all
4.3 Did you understand the definitions?	Yes, very well	Somewhat	No, not at all
4.4 Did you enjoy the practice?	Yes, very well	Somewhat	No, not at all
4.5 Do you hear or read these words outside of class?	Yes, very often	Sometimes	No, never

5. Pronunciation
5.1 What pronunciation points did you study this week? Write them if you remember.

_____

_____

Circle your answers.

5.2 Did you already know these rules?	Yes, very well	Somewhat	No, not at all
5.3 Did you understand the explanations?	Yes, very well	Somewhat	No, not at all
5.4 Did you enjoy the practice?	Yes, very well	Somewhat	No, not at all
5.5 Do you hear or read these sounds outside of class?	Yes, very often	Sometimes	No, never

6. Reading
6.1 What was the reading about?

_____

_____

Circle your answers.

6.2	Did you think it was difficult?	Yes, very well	Somewhat	No, not at all
6.3	Did you enjoy the reading?	Yes, very well	Somewhat	No, not at all
6.4	Do you read about similar topics outside of class?	Yes, very often	Sometimes	No, never

7. On the Web

7.1 What was this activity about?

_____

_____

Circle your answers.

7.2	Did you think it was difficult?	Yes, very well	Somewhat	No, not at all
7.3	Did you enjoy the activity?	Yes, very well	Somewhat	No, not at all
7.4	Do you read about similar topics outside of class?	Yes, very often	Sometimes	No, never

8. Speaking

8.1 What speaking phrases did you study this week? Write them if you remember.

_____

_____

Circle your answers.

8.2	Did you already know these expressions?	Yes, very well	Somewhat	No, not at all
8.3	Did you understand the explanations?	Yes, very well	Somewhat	No, not at all
8.4	Did you enjoy the practice?	Yes, very well	Somewhat	No, not at all

8.5 Do you hear these phrases of English outside of class?	Yes, very often	Sometimes	No, never
8.6 Is your teacher's correction helpful?	Yes, very much	Somewhat	No, not at all

9. Writing
9.1 What was the writing about?

_____

_____

Circle your answers.

9.2 Did you think the writing was difficult?	Yes, very much	Somewhat	No, not at all
9.3 Did you enjoy the writing?	Yes, very much	Somewhat	No, not at all
9.4 Do you write about similar topics outside of class?	Yes, very often	Sometimes	No, never
9.5 Is your teacher's correction helpful?	Yes, very much	Somewhat	No, not at all

Thank you!

AUTHORS' BIODATA

Joan Jamieson (Ph.D. in Second Language Acquisition and Teacher Education, University of Illinois) is Professor of English/Applied Linguistics at Northern Arizona University. In the MA-TESL program and the doctoral program in Applied Linguistics, she teaches classes in CALL, second language testing, and research design and statistics for applied linguistics. Joan has been involved in several research and development projects in these areas, as reflected in her publications.

Carol A. Chapelle (Ph.D. in Second Language Acquisition and Teacher Education, University of Illinois) is Professor of TESL/Applied Linguistics at Iowa State University. She teaches classes in CALL, second language testing, and second language acquisition in an MA program in TESL/Applied Linguistics with specialization in CALL and in a Ph.D. program in Applied Linguistics & Technology. Carol has published widely in these areas. She has served as editor of *TESOL Quarterly*, and she is second vice president of the American Association for Applied Linguistics.

Sherry Preiss (MAT in TESOL from the School for International Training in Brattleboro, Vermont) is Vice President of Adult and Higher Ed Publishing at Pearson Longman ELT based in White Plains, NY. Sherry has taught for many years in a wide variety of programs—private language schools, adult education programs, colleges and universities, and corporate language programs. She is the author of the Longman textbook, *NorthStar: Advanced listening and speaking*.

AUTHORS' ADDRESSES

Joan Jamieson
Department of English
Northern Arizona University
Flagstaff, AZ 86011-6032
Phone: 928/523-6277
Fax: 928/523-7074
Email: joan.jamieson@nau.edu

Carol A. Chapelle
Department of English
203 Ross Hall
Iowa State University
Ames, IA 50011
Phone: 515/294-7274
Email: carolc@iastate.edu

Sherry Preiss
Pearson Longman ELT
10 Bank Street, 9th Floor
White Plains, NY 10606
Phone: 914/287-8163
Email: Sherry.Preiss@PearsonEd.com