

LCTL Teachers' Assessment Knowledge and Practices: An Exploratory Study

Megan Montee, Aileen Bach, Anne Donovan, Lynn Thompson
Center for Applied Linguistics

Abstract

Language assessment allows teachers to gather information about student learning and adjust their instructional practices accordingly. When integrated with instruction, assessment can support student-centered teaching by helping instructors understand what students have learned or are able to do and what they still need to know (Shepard, 2000). However, language teachers often receive limited pre-service training in assessment and testing (Malone, 2008). To date, limited empirical work has been conducted about LCTL teacher education and language assessment literacy.

This paper explores a teacher training program designed to address the assessment needs of LCTL educators. This program provides a model of evidence-based assessment training that may be useful to other teacher educators. Furthermore, data generated from the project contributes to research about LCTL teacher education and assessment literacy and suggests future directions for work in this area.

Introduction

Language assessment allows teachers to gather information about student learning and adjust their instructional practices accordingly. When integrated with instruction, assessment can support student-centered teaching by helping instructors understand what students have learned or are able to do and what they still need to know (Shepard, 2000). However, in spite of the importance of language assessment for teaching and learning, teachers may view assessment as a hindrance or distraction. Additionally, assessment practices typically require some degree of technical knowledge or

training in order to implement, but language teachers often receive limited pre-service training in assessment and testing (Malone, 2008).

Limited empirical work has been conducted regarding teacher education and language assessment. However, there is a growing consensus in the field of language testing on the need for assessment literacy. With respect to teachers of Less Commonly Taught Languages (LCTLs), questions arise about how the assessment literacy needs of this group may differ from those of language teachers more generally. Research in this area will contribute to better teacher education resources for LCTL instructors and thus better assessment practices. Furthermore, such research can provide insight into the ways in which LCTL teachers mediate the relationship between assessment and instruction.

This paper explores a teacher training program designed to address the assessment needs of LCTL educators working in short-term foreign language programs offered through the STARTALK program. STARTALK (www.startalk.umd.edu) is a component of the U.S. National Security Language Initiative and helps meet the goal of expanding the teaching and learning of languages not widely taught in the U.S. STARTALK is made up of a network of summer programs. There are two primary types of STARTALK summer programs: K-16 student programs and teacher training programs. Additionally, some programs combine student and teacher programs so that classrooms serve as training sites for teachers. The program discussed in this paper was a teacher training program funded by the STARTALK initiative, and most of the program participants worked as instructors in STARTALK student programs.

The assessment training program discussed here provides a model of evidence-based professional development that may be useful to other LCTL educators. Furthermore, data generated from the project contributes to research about LCTL teacher education and assessment literacy. We describe both the structure and content of the course as well as a pilot study designed to explore the needs of LCTL teachers who participated in the course. The results of this preliminary study will inform the design of a follow-up study that further explores this topic.

Background

Two areas of research informed this project: assessment literacy and LCTL teacher education.

Assessment literacy

The term assessment literacy has emerged to describe the knowledge and skills teachers need to be able to plan for, administer, interpret and apply the results of assessments accurately and effectively (Boyles 2005; Malone, 2008; Stiggins, 1999; Stoyhoff and Chapelle, 2005; Taylor, 2009). Given the current focus on accountability in the United States, there are growing concerns about the potentially negative effects of standardized testing on instruction. However, as Shepard (2000) has noted, teacher education can focus on reconceptualizing classroom assessment practices “to be more effective in moving forward the teaching and learning process” (p. 6). From this perspective, classroom assessment can be a valuable source of information for teachers and students. The realization of this type of dynamic, continual assessment will depend on providing teachers with the knowledge and resources to integrate assessments into the learning process.

Some previous studies have examined teachers’ beliefs and practices. In a study of English language teachers in England, Rea-Dickins (2001) developed a framework for identifying teachers’ formative, classroom-based assessment practices. Rea-Dickins found that these practices were often informal and embedded within instruction. Davison (2004) conducted a study of English teachers in Australia and Hong Kong and found that contextual differences between teachers in the two countries led to differences in assessment practices. In particular, Davison focused on how teachers interpreted criteria for a written test and used these criteria to make scoring decisions. Individual variation between teachers was greater in Hong Kong, possibly due to a lack of professional development in the area of assessment.

While limited empirical research exists about teachers’ assessment literacy, several researchers have documented current issues and trends. Inbar-Lourie (2008) has noted a shift in pre-service courses toward assessment rather than testing. Whereas a testing

perspective emphasizes technical knowledge often associated with standardized testing, assessment courses often focus on classroom-based practices and the integration of assessment and instruction. Assessment is a broader term that includes traditional tests as well as other modes of evaluating students (e.g., portfolios, student self-assessment). This shift toward classroom-based assessment is a promising area of development for teacher education.

As the knowledge-base of assessment literacy continues to evolve and expand (Taylor, 2009), teacher educators face the challenge of preparing teachers for a variety of classroom and assessment contexts. A deeper understanding of teacher knowledge and practice will form a basis for future assessment literacy efforts.

LCTL Teacher Education

LCTL teacher education shares many issues with language teacher education in general while also facing some urgent challenges specific to these languages. Resources for teaching LCTLs are often limited, and teachers may have received limited training in pedagogy (Wang, 2009). In addition, the backgrounds of LCTL teachers are often very different than those of other teachers even within the language profession. Using data from 2009 STARTALK programs, Wang shows that LCTL teachers in this program were predominantly female native speakers of the target language who had immigrated to the U.S. Wang argues that many teacher training programs are based on assumptions about teacher characteristics which are not true for many populations of LCTL teachers. Thus, special attention to the needs of this population is an important area of research.

Although a focus on the needs and characteristics of LCTL teachers is important in language teacher education, insights about language teachers more generally are also relevant to the current study. Borg (2009) points out that research about teacher beliefs often focuses on the tension between beliefs and real-life practices. While research has focused on inconsistencies between these two areas, the tension between beliefs and practices should not necessarily be interpreted as a flaw. Teachers' actions are often mediated by contextual factors including time constraints, funding, and the pressures and demands of high-stakes testing. Borg argues that, "The

results of such constraints may be teaching that does not reflect the teacher's ideals" (p. 167). Further examination of the tension between teachers' beliefs and practices in classroom language assessment may reveal areas where additional resources and educational opportunities are needed.

A range of research methodologies has been used to measure teacher cognition, including questionnaires, observations, written reflections, interviews, and stimulated recalls (Borg, 2009). While questionnaire data can be useful, Borg cautions against relying on this data as the sole measure of teacher cognition, especially when researching the intersection of teacher belief and practice. Effective research design should include qualitative data analysis that can provide information about the means and scope of change in teacher cognition as well as any tensions between cognition and practice.

Course Design

The present research study was conducted during a seven-week course in language assessment. The course includes six weeks of online components and a two-day live workshop in the middle of the course. This blended learning format allows teachers from across the country to participate in the program and receive extended training and support in language assessment. The course is designed as an introduction to language assessment and focuses on helping participants develop language assessment tasks. Table 1 provides an overview and short description of key course content.

Table 1. Overview of course content

Week	Content	Description
1	Introduction Formative and summative assessment	Participants learn about the role of classroom-based assessment and the role of assessment during and at the end of instruction.
2	Key concepts in language assessment: validity, reliability, practicality, and impact	Through real-life scenarios participants are introduced to important concepts in language assessment.
3	Live workshop: Needs assessment, assessment planning, task development	Participants collaborate with colleagues to establish assessment goals and to develop classroom assessment tasks.
4	Review of live workshop material	Participants review what they learned at the workshop and continue developing assessment tasks.
5	Skills assessment	Participants discuss practical strategies for assessing listening and speaking, reading and writing.
6	Scoring performance assessments and reporting results	Participants discuss developing rubrics, giving feedback on assessments, and reporting results to stakeholders.

The purpose of the course is to provide teachers with foundational knowledge about assessment and to generate classroom assessment tasks that reflect participants' instructional and program goals. In particular, the course focuses on assessing students at the ACTFL Novice and Intermediate levels. Most STARTALK student programs focus on these two proficiency levels. Assessment can be particularly challenging because instructors must develop engaging assessments that measure small gains in language skills.

Methods

An exploratory study was designed to inform the content of the course and to serve as a pilot study for future research. The study included a pre-course questionnaire and a document analysis of assessment tasks developed by participants during the course. The study is part of a larger program of research which adopts an interpretive approach to research design, similar to the method described by Borg (2006). This allows for flexibility as the project progresses in areas such as participant selection and the development of additional research questions. A flexible approach is especially useful in this study for two reasons: First, as Borg (2006) points out, there are no established frameworks for such research; and second, the study was conducted during an assessment course. In order to minimize disruption to the course and the burden on participants, a flexible approach allowed data collection and analysis to coincide with course requirements. Future stages of the study may include additional sources of data such as participant interviews. The exploratory analysis described here provides an empirical basis for establishing next steps for the project.

In addition to questionnaires and document analysis, the study also relies on the perceptions of course instructors. The four authors all served as course instructors and worked regularly with participants in both online and face-to-face formats. Our observations and experiences as course instructors inform the findings and provide an additional source of information in the design of the study.

Research questions

In this exploratory study, we address two general research questions which allow for additional questions to emerge and facilitate the development of future stages of the study:

1. What are the assessment practices and needs of LCTL teachers?
2. How does the assessment literacy of LCTL teachers develop during a training course?

These questions necessarily relate, albeit indirectly, to issues of teacher cognition, or the complex ways that teachers understand their work (Borg, 2006). As Johnson (2009) describes, “learning to teach is conceptualized as a long-term, complex, development process that is the result of participation in the social practices and contexts associated with teaching and learning” (p. 10). While the current study focuses on teachers’ practices before and during a short-term training course, the research questions are situated within the larger context of how practicing teachers understand and implement assessment.

Data collection

As discussed earlier, data were collected as part of normal requirements for the training course. Before beginning the course, participants completed a short background questionnaire (Appendix A). As in previous years, the questionnaire was used to gauge interest in potential course topics and results were used to tailor the course to participants’ needs and interests. Additionally, the questionnaire included questions pertaining to participants’ beliefs and practices. These data address the first research question. The survey included several Likert-scale questions in which participants indicated their agreement with statements about assessment or their level of knowledge about particular topics.

Performance assessment tasks developed by participants provided a second source of data. During both the live workshop and the online course, participants worked individually or in groups to develop assessment tasks for their language programs. In this course, an assessment task was defined in a broad sense relying on Bachman

and Palmer's (1996) definition of a task as "an activity that involves individuals using language for the purpose of achieving a particular goal or objective in a particular situation" (p. 44). This definition, which was adopted for the course, is more general than some task-based approaches to teaching and testing (e.g., Ellis, 2003). Luoma (2004, p. 32) defines speaking tasks as "activities that involve speakers in using language for the purpose of achieving a particular goal or objective in a particular speaking situation." Examples of speaking tasks include giving oral directions based on a map or participating in an interactive, spoken role play scenario. Thus, performance tasks yield evidence of an examinee's language ability within a specific context. The evidence collected by performance tasks must be scored according to established criteria. Because the purpose of the course was to help participants develop materials for their language programs, the tasks varied in terms of purpose (formative or summative) as well as what was being assessed (i.e. reading, writing, speaking or listening).

Participants

We recruited volunteers for the study from the participants in the training course. A total of 29 course participants volunteered for the study. Because of varying levels of participation in the course, not all of the 29 participants completed both a questionnaire and an assessment task. While future stages of the study will only analyze results of participants supplying both types of data, all available data was included at this exploratory stage to identify potential themes. The fact that the research questions do not draw a direct link between questionnaire data and the results of the task analysis provided further justification for including all available data in analysis.

Of the 29 total participants included in the study, 23 completed the questionnaire and 22 designed an assessment task; 19 participants completed both. Participants represented a variety of LCTLs, and 28 of the 29 participants were native speakers of the languages they taught. Table 2 lists the languages represented in the

study. The numbers in the table do not add up to 29 because some programs targeted more than one language or age group.

Table 2. Participants by language and age taught

Language	N	Age group	N
Arabic	3	Elementary school	9
Chinese (Mandarin)	15	Middle school	13
Dari	1	High school	20
Hindi	4	Post-secondary	4
Persian/Farsi	1		
Swahili	1		
Turkish	3		
Urdu	3		

As Table 2 shows, Chinese (Mandarin) was the most heavily represented language among participants. High school programs were also well represented among participants, though there was more even distribution among target age groups than there was among target languages.

Results

Pre-course questionnaire

As stated earlier, 23 course participants completed a pre-course questionnaire before beginning the online course. The questionnaire (Appendix A) included nine items about participants' previous assessment training as well as their assessment beliefs and practices. The questionnaire included six Likert-scale questions and three open-ended questions. Likert-scale questions were on a four- to seven-point scale. For analysis, text response choices (e.g., "strongly agree") were converted to numeric values. There is no definitive standard for interpreting scale data (McIver & Carmines, 1981). A mean above 3.0 indicates general agreement and a mean of 2.0 or lower indicates general disagreement. Therefore, for this study we interpreted a mean agreement above 2.75 as an indication that most respondents agreed with the statement; below that number means that, on average, respondents disagreed with the statement. Participants could also indicate "I don't know" or "Not applicable."

These responses were not assigned a numeric value and were excluded from the analysis.

Participants first reported on their previous experience with language assessment. As shown in Table 3, most participants had not taken a university course in language assessment but 15 of the 23 respondents had participated in some sort of training or workshop.

Table 3. Previous assessment experience.

Types of assessment experience	Yes	No
	N (%)	N (%)
College course in language assessment	5 (21.7%)	18 (78.2%)
Workshop in language assessment	15 (65.2%)	8 (34.8%)
Rater training for a test	9 (39.1%)	14 (60.9%)

As Table 3 illustrates, many participants had at least some formal training in assessment. The 5 participants who had taken a college course in language assessment were all Chinese teachers. Of the 23 respondents, 7 had not participated in any type of assessment training prior to the course.

The survey also asked participants to indicate their preparedness for various assessment activities. Participants indicated how prepared they felt for each of the activities listed in Table 4 on a four-point scale ranging from “Very prepared” (4.0) to “Very unprepared” (1.0).

Table 4. Preparedness for assessment activities

Assessment activity	M	SD	N
Making my own assessments	3.05	.65	22
Scoring speaking assessments	2.95	.74	21
Scoring writing assessments	2.82	.80	22
Using standardized tests	3.10	.85	20
Interpreting standardized test results	2.75	.91	20

As Table 4 shows, mean scores fell above the 2.75 mark for agreement for the categories of developing assessments, scoring speaking and writing assessments, and using standardized tests.

However, mean scores fell at the 2.75 mark for interpreting test results, indicating that, at the start of the course, some respondents felt prepared for this activity while others did not. The 5 respondents who had previously taken a college course in language assessment generally reported high levels of preparedness for the various assessment activities.

The survey also asked participants to rate their familiarity with various assessment topics. These topics, listed in Table 5, covered key concepts and terminology from the course. Participants indicated how familiar they were with each topic.

Table 5. Participants' familiarity with assessment concepts

Assessment concepts	I've never heard this before	I've heard this term before but don't know what it means	I've heard this term and know a little bit about it	I could explain this term to my colleagues	N
Formative assessment	12 (52.2%)	6 (26.1%)	3 (13.0%)	2 (8.7%)	23
Summative assessment	9 (39.1%)	7 (30.4%)	5 (21.7%)	2 (8.7%)	23
Reliability	8 (34.8%)	11 (47.8%)	3 (13.0%)	1 (4.3%)	23
Validity	6 (26.1%)	10 (43.5%)	6 (26.1%)	1 (4.3%)	23
Achievement testing	12 (52.2%)	5 (21.7%)	3 (13.0%)	3 (13.0%)	23
Washback	7 (30.4%)	7 (30.4%)	4 (17.4%)	5 (21.7%)	23
Performance assessment	1 (4.3%)	6 (26.1%)	6 (26.1%)	10 (43.5%)	23
Proficiency testing	10 (43.5%)	9 (39.1%)	3 (13.0%)	1 (4.3%)	23
Backward design	10 (43.5%)	7 (30.4%)	4 (17.4%)	2 (8.7%)	23

As Table 5 shows, there were several terms that some participants had not heard before. More than half of participants were not familiar with the terms “formative assessment” and “achievement testing.” Additionally, while other terms were more familiar to participants, few felt confident enough in their knowledge to be able to explain the term to their colleagues. Those participants who had previously taken a college course in language assessment generally reported that they understood the terms and in many cases that they felt they could explain them.

In addition to their familiarity with key assessment concepts, participants also indicated how important it is to assess various skills as well as how frequently they assess these skills. The survey question about importance utilized a six-point scale ranging from “Very important” (6.0) to “Very unimportant” (1.0). In this case, a mean above 3.75 can be interpreted as “important” and 3.75 or below interpreted as “unimportant” on the scale.

Table 6. Importance and frequency of assessment practices

	Importance			Frequency N (%)				
	M	SD	N	Daily	Weekly	Monthly	<Monthly	Never
Reading	5.48	0.73	23	8 (34.8%)	10 (43.5%)	3 (13.0%)	2 (8.7%)	0 (0%)
Writing	5.48	0.59	23	8 (34.8%)	10 (43.5%)	3 (13.0%)	2 (8.7%)	0 (0%)
Listening	5.91	0.29	23	18 (78.3%)	4 (14.4%)	0 (0%)	1 (4.3%)	0 (0%)
Speaking	5.91	0.29	23	17 (73.9%)	5 (21.7%)	0 (0%)	1 (4.3%)	0 (0%)
Culture	5.17	0.83	23	1 (4.3%)	13 (56.5%)	5 (21.7%)	3 (13.0%)	1 (4.3%)
Grammar	4.74	1.10	23	5 (21.7%)	12 (52.2%)	2 (8.7%)	3 (13.0%)	0 (0%)
Vocabulary	5.48	0.73	23	13 (56.5%)	8 (34.8%)	1 (4.3%)	1 (4.3%)	0 (0%)

As Table 6 shows, respondents reported feeling that all skills except grammar are important or very important to assess. The mean for grammar is the lowest, falling between “important” and “somewhat important.” Speaking and listening had the highest scores for importance, and almost all respondents rated these as “Very Important” to assess.

Overall, respondents also reported assessing each skill frequently, either daily or weekly. Culture was the least frequently assessed area. Most participants reported assessing culture weekly but eight participants reported assessing it once a month or less and one participant reported never assessing culture.

The final survey question asked participants to rate their confidence in different assessment practices. For each of the statements listed in Table 7, participants indicated their agreement on a four-point scale ranging from “strongly agree” (4.0) to “strongly disagree” (1.0).

Table 7. Confidence on assessment practices

Statement	M	SD	N
I feel confident in the assessments I use with my students.	2.95	0.76	20
My students understand why they are being assessed.	3.00	0.59	20
I know how to provide feedback to students.	3.10	0.55	20
I use the results of assessments to help me plan instruction.	3.15	0.75	20

As Table 7 shows, the mean for all statements fell above the 2.75 mark, indicating that participants tended to agree with these statements.

Participants also responded to two open-ended questions about their assessment practices. First, participants were asked to indicate their strengths related to classroom assessment. Nineteen respondents answered this question. Some responses focused on specific assessments or practices such as creating speaking assessments. In several cases, respondents discussed how they integrated assessment and instruction. For example, one participant wrote that assessment could help in “understanding the difficulties that inhibit [student] learning” while another participant wrote about using assessments to “[see] if the students are learning the material being taught.”

In a second open-ended question, participants were asked to describe what they would like to improve about the way they assess their students. Nineteen respondents answered this survey question. Several respondents discussed improving their existing practices by

increasing the accuracy of their assessments or developing better rubrics. In several cases, respondents mentioned students. One respondent wanted to engage students, another mentioned making assessments more interesting, and a third comment discussed providing better feedback to students. Overall, the comments reflected an interest in linking teaching and assessment and in engaging students in the assessment process.

Analysis of assessment tasks

The second part of the study involved an analysis of assessment tasks developed during the course. As stated earlier, these tasks were developed and revised during the course as an assignment. Each task was developed using a task template (Appendix B). Using this template, participants described key characteristics of their tasks. Participants were instructed to develop a task to use in their summer language programs and could work individually or in pairs. Each participant received comments and feedback from course instructors and could turn in revised tasks as part of the course. A substantial part of the live workshop was dedicated to developing these tasks. Participants who only completed the online sections of the course had access to all materials from the live workshop and developed assessment tasks as part of the online course requirements. All tasks reviewed for this study were completed by participants who attended the live workshop. Though not addressed in this study, this factor may prove to be a salient issue for future phases of the study as the face-to-face workshop provides more opportunities for reflection and feedback during the task development process.

A total of 22 tasks were analyzed. The first and second authors of this paper reviewed each task and compiled a set of themes that reflected task contents. The researchers met before, during and after the coding process in order to discuss findings and reach consensus on the themes. Although there were no explicit disagreements between the two researchers, each tended to notice different themes. Possible interpretations were discussed and agreed upon, and notes from the research meetings were used in subsequent phases of the analysis of assessment tasks and in reporting the results. Furthermore, analysis was reviewed and commented on multiple times by other members of the research team. The analysis of

assessment tasks addresses the second research question, which looks at the ways that teachers' knowledge developed throughout the course. In addition, the task analysis indicates areas where additional professional development may be needed. At this stage of analysis, this more general approach was deemed most appropriate. In future stages, a more specific coding scheme may be used to capture more detail about tasks.

Additionally, this analysis relies heavily on the course instructors' experiences working with participants to develop the tasks. The researchers responsible for the task analysis also served as instructors in the course and were responsible for providing feedback on the tasks and working with participants to improve their tasks. While the task analysis was conducted separately from the course, our experiences as instructors helped inform our understanding of the tasks and our interpretation of how participants applied what they learned in the course. We discuss the themes in terms of strengths and challenges for participants.

Strengths

The analysis of assessment tasks reflected a number of strengths and areas where participants had developed knowledge during the course. Namely, participants were largely successful in developing contextualized tasks that reflected authentic communicative situations. The use of authentic tasks and materials was a major area of emphasis both in the course and in STARTALK programs generally. The tasks reflected participants' growing understanding that assessment tasks should reflect real-world language situations. For example, in one speaking task developed in the course, students were instructed to discuss their food preferences in the context of a study abroad scenario. The task was motivating and included an authentic situation that students might encounter in the real world. Most, but not all, of the tasks developed in the course used authentic contexts like this. In both the live and online portions of the course, this issue had been a challenge for many participants. Initial ideas for tasks often lacked motivation for the language performance or a connection to authentic contexts. The issue of authenticity in language education can be categorized as a value that is realized in and through certain practices. While the participants in

the course may need to acquire additional skills such as how to select and modify authentic materials, the task analysis indicates a shift in focus toward authentic language assessment.

The tasks also reflected participants' knowledge and experiences as language teachers. Overall, the tasks showed a strong connection to classroom practices and instructional goals. Moreover, participants seemed to demonstrate their understanding of assessment concepts most meaningfully when relating them directly to instruction. Many tasks began with participants describing the relationship between the assessment task and the lesson or instruction that preceded it. Additionally, tasks were generally grade-level appropriate and grounded in thematic instructional units. Because the assessment workshop was developed to support existing STARTALK programs, participants were encouraged to use their lesson plans as a starting point for their assessment tasks. This connection to lesson planning seemed especially effective because it embedded the assessment tasks in the participants' actual classroom activities. Participants also demonstrated best classroom practices in the task development by incorporating paired tasks and group presentations. These tasks often incorporated the target culture by situating the task in the context of study abroad, travel in the target culture, holidays, and other cultural practices.

Challenges

Overall, the task analysis reflected the value of connecting assessment and instruction and showed evidence of participants' emerging understanding of the importance of authentic communicative situations. However, the task analysis also indicated several challenges and areas where assessment concepts seemed unclear to participants.

One recurrent issue across tasks was the appropriateness of the task for the target proficiency level. While most of the tasks were intended for students at the ACTFL Novice or Intermediate proficiency levels, the language demands of tasks often exceeded expected performance at that level. For example, several tasks at the Novice level required students to produce extended oral discourse. Students at the Novice level can produce words and short, familiar phrases (ACTFL, 2012). Additional evidence of the mismatch

between the task and the proficiency level was seen in the rubrics which some participants included with their tasks. In one example, the participants listed “no linguistic mistakes” as part of the performance criteria for a Novice level task. However, at the Novice level, mistakes can be expected (ACTFL, 2012). In both the tasks and the rubrics, many course participants struggled to match the task to the proficiency level. This may indicate several challenges. First, participants may not have a complete understanding of what students are capable of doing at the Novice level. Second, this issue may point to differences between assessment and instructional tasks. In instruction, students often receive scaffolding and support, and thus may be able to produce more language. However, in an assessment task, and in summative assessment in particular, the focus should be students’ independent language abilities. While course participants seemed to have an understanding of the connections between assessment and instruction, the task analysis indicated that they may not be able to draw distinctions between these activities when it comes to evaluating what students can do independent of instructional support.

Another observation from the task analysis was that tasks tended to focus on formative assessment. Many of the tasks were informal classroom assessment tasks and often included group or pair work. While this may only be indicative of the types of tasks participants chose to develop in the course rather than their actual assessment practices, our experience as course instructors indicates that summative assessment seems to be more challenging for participants than formative assessment. This issue also relates to the connection between instruction and assessment. Formative assessment is necessarily embedded in instruction and may often be difficult to distinguish from it; many tasks developed in the course could serve as instructional tasks. Summative assessment is separated from instruction and focuses on what students can do independently, but course participants seemed to struggle in designing tasks that required students to perform independently. Some of our observations of and discussions with participants showed that they may feel that mistakes in student performances reflect failure or problems in their teaching. This may be addressed by helping teachers develop a deeper understanding of proficiency levels and

language development. For example, grammatical errors in the speech of a student at the ACTFL Intermediate level may be a sign of linguistic creativity and emerging proficiency.

One final challenge reflected in the tasks was the lack of opportunities for spontaneous and independent speech. Again, this may only indicate that participants chose to develop tasks that focused on the presentational communication; however, our observations from working with participants as well as some features of the tasks themselves indicate that participants had trouble designing tasks that elicited spontaneous speech. In many cases, participants identified the target communicative mode as Interpersonal Speaking (according to the ACTFL definition) when the task instead assessed presentational communication. Group or paired tasks often included time for participants to prepare for the situation before presenting or performing in front of the class; few tasks elicited spontaneous speech from students.

Discussion and Conclusions

The purpose of this study was to address two exploratory research questions. First, through analyzing survey data, we examined assessment beliefs, practices, and needs of LCTL teachers. Next, through an analysis of assessment tasks, we examined how participants' assessment knowledge developed during an assessment training course. This revealed areas of growth as well as challenges faced by participants. The results of the survey show that many of the LCTL teachers in the course had not taken a university-level assessment course although many had participated in some sort of assessment workshop or rater training session. There were also a number of participants who had not received any other types of prior assessment training. As Wang (2009) points out, many LCTL teachers may lack the formal background in pedagogy that other teachers have. A close analysis of survey data did suggest that the small number of course participants who had taken a college course in language assessment tended to feel more confident in their assessment practices and were more comfortable with assessment concepts at the start of the course.

This has several implications for the assessment literacy of this group. Because of limited opportunities for pre-service

assessment courses, LCTL teachers may not have formal knowledge of technical concepts related to assessment. As with this course, assessment training may often be conducted as in-service rather than pre-service training. In this case, assessment concepts can be tied to the practical knowledge that LCTL teachers have about instruction. In this study, the analysis of assessment tasks showed that teachers were able to connect instruction and assessment in meaningful ways. Future assessment literacy efforts may further focus on these connections and help LCTL teachers build technical concepts that are firmly rooted in their practical knowledge. Assessment training efforts for this population may also need to focus on building teachers' formal knowledge around concepts such as formative and summative assessment.

The survey indicated that many participants were not familiar with the assessment terms used in the course. Assessment may be seen as difficult or technical because it includes many unfamiliar terms. Assessment literacy efforts should be focused on making this terminology accessible and meaningful to teachers. In many cases, the concepts may be familiar but the terminology may be new. Once again, this is an opportunity for teacher educators to draw connections between technical knowledge and practice.

The second research question in this study asks how the assessment training course described in this study impacts participants' assessment knowledge and practices. Analysis of the assessment tasks created by course participants shows that the course's face-to-face workshop component is effective in increasing participants' knowledge of assessment concepts and applying these to assessment tasks that are strategically embedded in instruction. Participants demonstrated increased awareness of the types of assessment tasks that were appropriate for their programs' age group, language, and proficiency level, although participants did struggle to develop tasks that matched the targeted proficiency level. Future iterations of the course may focus more on appropriate proficiency level expectations and language elicitation. Speaking and listening may be a particular area of focus as well. The survey indicated that respondents were very consistent in rating the assessment of these skills as "Very Important," a result which most likely reflects the communicative focus of many STARTALK programs. Future

research may focus more particularly on how teachers approach and implement oral proficiency assessment as well as the reasons for its perceived importance.

An important result of the assessment training course was an increased awareness of and appreciation for assessment as a tool for guiding and improving language instruction. In the last module's assignment, participants were asked to reflect on what they had learned in the course and how they would apply this to their teaching. While participants identified various basic assessment concepts they had learned, participants primarily focused on the increased range of knowledge and agency they gained as a language instructor with the basic assessment knowledge to develop and use assessments effectively and in ways that support instruction.

In conclusion, this study resulted in some suggestive findings about the ways that LCTL teachers perceive assessment and the ways in which their perceptions relate to practice. This exploratory study is limited in that it relies on self-reported survey data and document analysis. More in-depth, qualitative data are needed to more fully understand the nature of LCTL teachers' perceptions and to answer questions about how beliefs and practices interact and change through professional development. Because many LCTL teachers receive in-service rather than pre-service training in language assessment, assessment concepts must be anchored in, but also differentiated from, instructional practices.

References

- American Council on the Teaching of Foreign Languages. (2012). *ACTFL Proficiency Guidelines--Speaking*. Retrieved March 15, 2012 from <http://www.actfl.org/files/public/GuidelinesSpeak.pdf>
- Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice: Designing and developing useful language tests*. New York: Oxford University Press.
- Borg, S. (2006). The distinctive characteristics of foreign language teachers. *Language Teaching Research*, 10(1), 3–31.
- Borg, S. (2009). *Teacher cognition and language education: Research and practice*. New York: Continuum.
- Boyles, P. (2005). Assessment literacy. In M. Rosenbusch (Ed.), *National Assessment Summit Papers* (pp. 18–23). Ames, Iowa.
- Davies, A. (2008). Textbook trends in teaching language testing. *Language Testing*, 25(3), 327–347.
- Davison, C. (2004). The contradictory culture of teacher-based assessment: ESL teacher assessment practices in Australian and Hong Kong secondary schools. *Language Testing*, 21(3), 305–334.
- Ellis, R. (2003). *Task-based language learning and teaching* (1st Ed.). New York: Oxford University Press.
- Inbar-Lourie, O. (2008). Constructing a language assessment knowledge base: A focus on language assessment courses. *Language Testing*, 25, 328–402.
- Johnson, K. E. (2009). *Second language teacher education: a sociocultural perspective*. New York: Routledge.
- Luoma, S. (2004). *Assessing speaking*. New York: Cambridge University Press.
- Malone, M. (2008). Training in language assessment. In E. Shohamy & N. Hornberger (Eds.), *Encyclopedia of Language and Education* (2nd ed., Vol. 7: Language Testing and Assessment, pp. 429–462). New York: Springer.
- McIver, J. P., & Carmines, E. G. (1981). *Unidimensional Scaling*. Thousand Oaks, CA: Sage.
- Rea-Dickins, P. (2001). Mirror, mirror on the wall: identifying processes of classroom assessment. *Language Testing*, 18(4), 429–462.

Shepard, L. A. (2000). The role of assessment in a learning culture. *Teaching and Learning*, 229–253.

Stiggins, R. (1999). Assessment, student confidence, and school success. *Phi Delta Kappan*, 81(3), 91-98.

Stoyhoff, S., & Chapelle, C.A. (2005). *ESOL tests and testing: A resource for teachers and program administrators*. Alexandria, VA: TESOL.

Taylor, L. (2009). Developing assessment literacy. *Annual Review of Applied Linguistics*, 29(1), 21–36.

Wang, S. (2009). Preparing and supporting teachers of less commonly taught languages. *Modern Language Journal*, 93(2), 282–287.

Appendix A: Pre-Course Questionnaire Items

Questionnaire administered online via SurveyMonkey

Have you participated in any of the following? (Yes/No)

College course in language assessment

Workshop in language assessment

Rater training for a test (e.g., ACTFL OPI)

If you've received any other training related to language assessment, please describe it in the box below.

How well prepared do you feel for each of the following activities:					
	Very prepared	Somewhat prepared	Somewhat unprepared	Very unprepared	Don't know/not sure
Making my own assessments					
Scoring speaking assessments					
Scoring writing assessments					

Using standardized tests					
Interpreting the results of standardized tests					

The list below includes some terms related to language assessment. Some terms might be familiar and others might be new to you. For each term, select the response which best describes how well you know the term.

	I've never heard this before	I've heard this term before but don't know what it means	I've heard this term and know a little bit about it	I could explain this term to my colleagues
Formative assessment				
Reliability				
Portfolios				
Validity				
Summative assessment				
Achievement testing				
Washback				
Performance assessment				
Proficiency testing				
Backward design				

Based on your experience, how important do you think it is to assess students in each of the following areas?

	Very important	Important	Somewhat Important	Somewhat Unimportant	Unimportant	Very unimportant
Reading						
Writing						
Listening						
Speaking						
Culture						
Grammar						
Vocabulary						

Next, we'd like to know how often you currently assess your students in each of the following skill areas. For each skill, indicate about how often you assess your students.

<http://www.z2systems.com/neoncrm/March-2013-NEON-Group-Training-Session-41334>

	Daily	Weekly	Every two weeks	Every month	Every few months	Once a semester	Never
Reading							
Writing							
Listening							
Speaking							
Culture							
Grammar							
Vocabulary							

Please indicate how much you agree or disagree with each of the statements listed below.						
	Strongly agree	Agree	Disagree	Strongly disagree	Don't know/ not sure	Not applicable
I feel confident in the assessments I use with my students.						
My students understand why they are being assessed.						
I know how to provide feedback to students.						
I use the results of assessments to help me plan instruction.						

Think about how you currently assess your students. What do you think your strengths are related to classroom assessment?

What do you wish you could improve about the way you currently assess your students?

Appendix B: Task Development Worksheet

Task Development Worksheet

Instructions: This worksheet will help you develop assessment tasks. For each task, complete the matrix below.

Name of task	
Communicative mode(s) assessed	
Target proficiency level	
What grade/age levels is the task written for?	

Background and context

A good task should be contextualized. The task instructions should include a thorough description of the background and context. Providing context also supports task authenticity. You should consider the following.

- What information will you provide to give the student a foundation for completing the task? Describe the scenario in which the student is completing the task.
- What materials will you provide to support task completion? Describe any input (written, oral, or graphic) that will be included in the task.

Use the space below to write the background and context.

Instructions to students

In this section, you will describe what instructions will be given to students. Good instructions should be explicit, clear and tell the students what is expected of a response. You should consider the following.

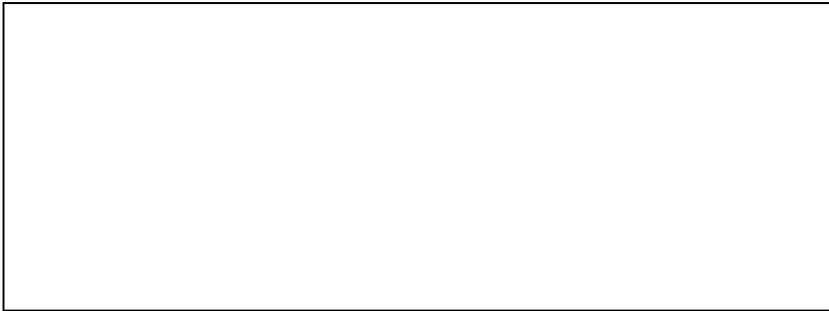
- What will students have to **do**? For example, will they
 read a text in the target language and respond to questions in English?
 listen to a conversation and circle the picture that corresponds to the main topic?
 write the name of a food next to a picture of it?
 ask a simple question about a schedule?
- To what question or statements will students respond?

Use the space below to write the instructions the students will receive. You may also want to attach additional materials such as worksheets or graphics that students will receive as part of the task.

Expected Response

A good task includes clear expectations. Having clear criteria is important to the reliability of the assessment. In addition, these criteria should be shared with students so they have a clear idea of what is expected on the task and can benefit from feedback on their performance.

Use the space below to describe the criteria of an expected response. You may also want to write an example response.



© Center for Applied Linguistics