

**2011 TELPAS**  
**Listening and Speaking Audit**  
**Report**

## Table of Contents

<b>I. 2011 TELPAS Listening and Speaking Audit .....</b>	<b>3</b>
<b>Background.....</b>	<b>3</b>
<i>2008 Pilot Audit.....</i>	<i>3</i>
<i>2011 Audit.....</i>	<i>3</i>
<b>Method.....</b>	<b>4</b>
<i>Sample.....</i>	<i>4</i>
<i>Documentation Form and Questionnaires .....</i>	<i>4</i>
<i>Audit Evaluation Procedures.....</i>	<i>5</i>
<b>Results.....</b>	<b>5</b>
<i>Listening and Speaking Audit .....</i>	<i>5</i>
<i>Feedback from Teachers on Audit Process and Documentation Form .....</i>	<i>8</i>
<b>II. 2011 District, Campus, and Teacher Procedural Questionnaire for Listening, Speaking and Writing.....</b>	<b>9</b>
<b>DTC Questionnaire Results .....</b>	<b>9</b>
<b>CTC Questionnaire Results.....</b>	<b>10</b>
<b>Teacher Questionnaire Results .....</b>	<b>11</b>
<b>Summary of Listening and Speaking Audit and Audit Questionnaire Analyses.....</b>	<b>12</b>
<b>Appendix A: Teacher Documentation Form .....</b>	<b>13</b>
<b>Appendix B: Teacher Instructions .....</b>	<b>16</b>
<b>Appendix C: Teacher Questionnaire.....</b>	<b>17</b>
<b>Appendix D: Audit Form for Evaluating Teacher Documentation.....</b>	<b>20</b>
<b>Appendix E: Audit Meeting Agendas .....</b>	<b>22</b>
<b>Appendix F: DTC Questionnaire .....</b>	<b>24</b>
<b>Appendix G: CTC Questionnaire.....</b>	<b>26</b>

## 2011 TELPAS Audit

This audit had two major components:

- I. audit of assessments for listening and speaking domains
- II. audit of training and administration procedures for listening, speaking, and writing domains

### I. 2011 TELPAS Listening and Speaking Audit

#### Background

Since the 2004–2005 school year, the Texas Education Agency (TEA) has conducted periodic audits of the Texas English Language Proficiency Assessment System (TELPAS) as a means of collecting reliability and validity evidence for the assessment program. For TELPAS writing, annual audits were conducted from 2005 through 2008 (the next writing audit is planned for the 2012–2013 school year). For TELPAS listening and speaking, an audit method was piloted in 2008, and based on internal discussions and external input the first audit was conducted in 2011.

#### *2008 Pilot Audit*

A small pilot audit was conducted in May 2008 to explore a method of gathering reliability and validity evidence for the TELPAS listening and speaking domains. For this pilot, 43 trained audit raters observed ELLs in participating districts and conducted second blind ratings of the students' listening and speaking proficiency with the goal of verifying the ratings assigned during the spring 2008 statewide TELPAS administration. Audit raters included staff from TEA, the testing contractor (Pearson), school districts, and education service centers. The purpose of the pilot was to examine the audit method and the feasibility of conducting a similar audit on a larger scale.

An important finding was that the time of year of the audit posed logistical challenges for the participants. Audit raters from the school districts and education service centers indicated that they would not volunteer again for an audit in May because of the time-consuming nature of their end-of-year duties. They also indicated that teachers of the students would need to be instructed ahead of time to plan academic lessons that engaged students in more listening and speaking activities. Because of the limited time spent with the students, audit raters sometimes indicated the need to see students in a greater variety of academic listening and speaking interactions to be confident about their verification ratings. Audit raters suggested conducting the audit during the actual TELPAS testing window.

TEA's Texas Technical Advisory Committee (TTAC) of national psychometric experts reviewed the findings and made recommendations to address the findings. More information about this pilot can be found in Chapter 16 of the *Technical Digest for the Academic Year 2007–2008* at <http://www.tea.state.tx.us/student.assessment/techdigest/yr0708/>.

#### *2011 Audit*

Based on the 2008 pilot, plans were initially made to proceed with a larger-scale audit using a similar methodology but to conduct the audit during the spring TELPAS administration window in March rather than in May. Input from bilingual/ESL leaders in the state indicated, however, that TEA would not likely be able to obtain enough volunteers to serve as auditors due to staff priorities and

workloads during the spring statewide assessment season. The TTAC was consulted, and a new audit methodology was chosen.

Instead of focusing on interrater reliability estimates calculated by using a second rating from an audit rater, the 2011 listening and speaking audit was designed to require teachers to submit documentation of their justification for the listening and speaking ratings they assigned to students selected for the audit. Trained reviewers from TEA and Pearson examined the documentation to determine whether the teachers had based their ratings on the appropriate classroom interactions and interpreted the scoring rubrics in the intended manner. This audit procedure, which resulted in a slightly different type of validity evidence from the previous audit, provided information about whether student ratings were aligned with the required scoring rubrics and based on the appropriate types of classroom oral interactions and instructional activities.

This approach provided the state with evidence of the degree to which the TELPAS holistically rated listening and speaking assessments are being appropriately administered in order to evaluate whether revisions to the training and assessment processes are needed.

TEA plans to conduct additional listening and speaking audits targeting different grade clusters over the next few years.

## **Method**

### *Sample*

A sample of approximately 500 students in grades 3–5 was included in the audit. The sample was selected to include students of all proficiency levels (beginning, intermediate, advanced, and advanced high) from across the state. The student's proficiency level from the previous year (2010) was used to inform the sample selection, and an assumption was made that the student would be one proficiency level higher in 2011 compared to 2010. Students at the beginning level were sampled by including students who were enrolled in newcomer centers designed to assist students with their transition into the United States. Approximately equal proportions of students were selected across the three grade levels. To minimize involvement at the campus level, no more than three students per campus were sampled, with only one student per grade level. To minimize district involvement, no more than ten campuses per district were selected. Finally, the sample was selected to be representative of the Texas ELL population in terms of the gender distribution of ELLs and the distribution of ELLs across the 20 state regions.

### *Documentation Form and Questionnaires*

An online documentation form (see Appendix A) was designed to elicit the type of information auditors needed from the students' raters in order to evaluate the efficacy of the ratings from the operational assessment. Instructions for how to complete the documentation form (see Appendix B) were provided to each campus selected for the audit. Information was requested about the types of instructional activities students' listening and speaking ratings were based on and justifications for the proficiency level ratings assigned.

After teachers completed the online form, they were required to complete a short audit questionnaire (see Appendix C). Questions 1–3 were designed so teachers could provide feedback

about the audit process and the documentation form, which will help in refining future audits of this type. The rest of the questionnaire required teachers to answer audit questions related to their rater training and the operational TELPAS holistic rating process for all language domains.

### *Audit Evaluation Procedures*

The teacher documentation was reviewed by TEA and Pearson staff composed of experienced and trained TEA and Pearson personnel who had participated in the development of the scoring rubrics and/or had responsibility for the design and approval of annual rater training and calibration materials used to prepare teachers to apply the scoring rubrics. Three meetings were held to establish and finalize an evaluation form for these evaluators to use in examining the submitted documentation.

In the first meeting, evaluators met to apply a draft evaluation form to teacher documentation and to make modifications to enhance the usefulness of the form. The evaluation form was applied to a small sample of teacher documents first, with evaluators working as a group. Following the initial discussion, evaluators used the form with several documents individually and compared their evaluations as a group afterwards. Modifications to the evaluation form were made to support consistent interpretations by the evaluators. Teacher documentation reviewed during the first meeting was used for training purposes only, and was not included in the audit analyses.

In a subsequent meeting, the revised evaluation form was applied to additional teacher documents, and minor modifications were made to finalize the evaluation form. Teacher documentation rated by the group during this meeting was included in the audit analyses. The final evaluation form is included in Appendix D. Agendas for the first two meetings are included in Appendix E.

After the second meeting, each evaluator was given a set of teacher documents to review on their own. Evaluators were asked to flag cases where some aspect of the teacher documentation was difficult to evaluate in order to have a group discussion. A third meeting was held a few weeks later to discuss flagged documentation and arrive at group consensus about how to apply the evaluation form.

To maintain high classification consistency, each teacher document that was not discussed by the group was assigned to two evaluators. In instances where the evaluators did not arrive at the same conclusion, the documentation was evaluated a third time by TELPAS project management staff, and the third evaluations were considered final.

## **Results**

### *Listening and Speaking Audit*

A random sample of the submitted teacher documentation was evaluated because of the time-intensive nature of the training and evaluation process. In all, 279 teacher documents were evaluated. As shown in Table 1, more than 60 documents were reviewed per proficiency level and TELPAS domain. Three documents included information for only one TELPAS domain because the teachers indicated an admission, review, and dismissal (ARD) committee decision had been made not to rate the student with a disability in a particular domain because of the nature of the student's disabling condition.

Table 1. Number of Teacher Documentation Forms Reviewed per Proficiency Level and Language Domain

Proficiency Level	Listening		Speaking	
	Frequency	Percent	Frequency	Percent
Missing*	2	0.72	1	0.36
Beginning (B)	66	23.66	74	26.52
Intermediate (I)	66	23.66	75	26.88
Advanced (A)	70	25.09	67	24.01
Advanced High (H)	75	26.88	62	22.22

\*Students with a disability not rated based on ARD committee decision

A comparison of the TELPAS ratings for the listening and speaking domains is provided in Table 2. Approximately 80% of students received the same rating in both domains. Only one student received a non-adjacent rating across domains, with a rating of beginning in listening and advanced in speaking.

Table 2. Teacher Assigned Student Ratings across TELPAS Domains

Listening Rating	Speaking Rating	Frequency	Percent
Missing*	B	2	0.72
B	Missing*	1	0.36
B	B	63	22.58
B	I	1	0.36
B	A	1	0.36
I	B	9	3.23
I	I	52	18.64
I	A	5	1.79
A	I	22	7.89
A	A	46	16.49
A	H	2	0.72
H	A	15	5.38
H	H	60	21.51
Total		279	100.0

\*Students with a disability not rated based on ARD committee decision

As shown in Table 3, evaluators determined that approximately three-quarters of the teacher documentation forms supported the student ratings for both listening and speaking based on the scoring rubrics (Proficiency Level Descriptors, or PLDs).

Table 3. Evaluation of Teacher Justification of Assigned Student Rating Based on PLDs

Category	Listening		Speaking	
	Frequency	Percent	Frequency	Percent
Supported Rating	205	74.01	213	76.62
Did Not Support Rating	72	25.99	64	23.02
Incomplete Documentation	0	0.00	1	0.36
Total	277	100.00	278	100.00

Table 4 provides an analysis of the types of problems seen in the approximately 25% of cases in which the teacher documentation did not support the student rating (72 instances for listening, 64 instances for speaking). In approximately 10% of these cases, the rating justifications were based on factors primarily outside of the PLDs. About 38% of the justifications that did not support the rating were PLD-based but described the wrong proficiency level. For example, the teacher may have rated a student as intermediate in listening, but the evidence provided actually supported a rating of advanced based on the PLDs. Finally, approximately 60% of the time, the unsupported ratings provided information too general or unclear to substantiate. The percentages held across both listening and speaking domains.

Table 4. Teacher Documentation that Did Not Support the Student Rating

Category	Listening		Speaking	
	Frequency	Percent	Frequency	Percent
Used Justification not Consistent with PLDs	7	9.72	6	9.38
Used PLDs but Assigned Incorrect Rating	27	37.50	24	37.50
Information Provided too General/Unclear	43	59.72	38	59.38

Note: Percentages do not sum to 100% because evaluators were asked to mark all that apply.

When the information provided in Table 3 is broken down into proficiency levels in Tables 5 and 6, a clear pattern emerges: the beginning rating (B) is judged as accurate over 85% of the time, the advanced high rating (H) is judged as accurate almost 95% of the time, but the judged accuracy of intermediate and advanced ratings is lower (51-67% accurate). This may be because rating, justifying, or evaluating documentation for these levels is more challenging and/or because there are two adjacent ratings for each of these levels (students at each of these levels may be bordering on the level below or the level above) whereas there is only one adjacent rating for the beginning and advanced high ratings.

Table 5. Evaluation of Teacher Justifications by Listening Proficiency Level

Substantiated Rating?		Listening Rating			
		B	I	A	H
Yes	Frequency	57	42	36	70
	Percent	86.36	63.64	51.43	93.33
No	Frequency	9	24	34	5

Percent      13.64      36.36      48.57      6.67

Table 6. Evaluation of Teacher Justifications by Speaking Proficiency Level

Substantiated Rating?		Speaking Rating			
		B	I	A	H
Yes	Frequency	64	50	40	59
	Percent	86.49	66.67	60.61	95.16
No	Frequency	10	25	26	3
	Percent	13.51	33.33	39.39	4.84

In all but one case, the teacher who provided the documentation appeared to be the student’s teacher. About 94% of teachers provided evidence of observations and interactions with the student they rated in a variety of social and academic contexts (see Table 7). Of those teachers who did not provide evidence of social and academic contexts (“No” in Table 7), about 87% failed to provide descriptions of observations in a variety of academic contexts, and 69% failed to provide descriptions of observations in a variety of social contexts (see Table 8).

Table 7. Teachers’ Documentation of Student Observations

Provided Evidence of Observations in Social and Academic Contexts?	Frequency	Percent
Yes	263	94.27
No	15	5.38
Incomplete Documentation	1	0.36
Total	279	100.00

Table 8. Missing Information from Teachers’ Documentation of Student Observations

	Frequency	Percent
No Variety of Academic Contexts	13	86.67
No variety of Social Contexts	9	69.23

*Note.* There were a total of 15 teachers who lacked evidence of student observations. Six of these teachers lacked evidence in both academic and social contexts.

#### *Feedback from Teachers on Audit Process and Documentation Form*

After teachers completed the online documentation form, a short questionnaire was administered. Questions 1–3 asked teachers for information about the audit process itself and documentation form used, which will be used in refining future audits of this type. The rest of the questionnaire asked the teachers about their experiences with TELPAS holistic rater training and the operational TELPAS holistic rating process for all language domains. The questionnaire is provided in Appendix C.



The teacher questionnaire, which was completed by 536 raters, provided the following information related to the audit process and documentation form. Approximately 60% of raters were able to finish the audit documentation form within 1 hour, and almost 90% finished within 2 hours. However, teacher comments suggested that there was confusion about reporting the time spent on the audit. Some teachers thought they should include all the time they spent observing students in order to arrive at a rating. Others were not clear that “rating documentation” referred to the electronic audit form. More than 99% of teachers indicated that the audit documentation form was very or generally clear. Most teachers found the audit process to be straightforward.

Suggestions for improvement included:

- Providing a “save” function and a “back” button in the online system so that the teacher doesn’t have to restart the process if interrupted
- Include a spell-check function
- Provide teachers with a hard copy of the audit form, or a mechanism for printing out the document
- Include more background questions about the students, for example, years in US

## **II. 2011 District, Campus, and Teacher Procedural Questionnaire for Listening, Speaking and Writing**

District and campus testing coordinators and the specified teachers on campuses selected for the audit were asked to fill out a questionnaire as part of an audit of the training and administration procedures they followed for the listening, speaking, and writing domains. The domain of reading is not included because, rather than being holistically rated, it is assessed with a multiple-choice assessment for grades 2–12. The questionnaire was used to collect information about the holistic training materials and process used by the district or the campus for all holistically assessed domains.

Questionnaires for teachers, district testing coordinators, and campus testing coordinators are provided in Appendices C, F, and G respectively.

### **DTC Questionnaire Results**

There were 229 district testing coordinators (DTCs) who responded to the DTC questionnaire, of which 34 were also campus testing coordinators (CTCs). The number of raters per district varied. Approximately 56% of districts had fewer than 100 raters, while roughly 44% had more than 100 raters.

Of 190 DTCs that had raters who did not successfully calibrate by the end of calibration set 2, all but 6 (97%) were given supplemental support by a trained supplemental support provider, which is required before such raters complete the third and final rater calibration set. Based on DTC comments, it appears that most raters who were not given supplemental support were not needed to serve as raters in the live administration because the district had sufficient numbers of raters who had successfully calibrated by the end of set 2.

All but 3 of the 229 DTCs reported that their district implemented procedures to support the validity and reliability of the TELPAS rating process, as required by state policy in the annual *District and Campus Coordinator Manual*. From those recommended in the manual, the procedures implemented included —

- having raters collaborate with other teachers of the students in determining the students' ratings (77%),
- having raters collaborate with each other in determining the ratings of students near the border between proficiency levels (69%),
- having a district-determined selection of writing collections rated a second time and, if the ratings differed, resolution ratings were provided and used (46%), and
- having all writing collections rated a second time and, if the ratings differed, resolution ratings were provided and used (27%).

DTCs indicated the following practices regarding selection and implementation of procedures to support the validity and reliability of the TELPAS rating process:

- district selected the procedures for campuses to follow (55%)
- district generally selected procedures, but allowed variability at the request of individual campuses (32%)
- campuses were directed to submit plans to the district for approval or campuses established procedures independently and no district approval was required (10%)
- district required each campus to keep documentation of the procedures at the campus (67%)
- district required each campus to submit documentation describing the procedures followed (40%)

Nearly 70% of DTCs reported that uncalibrated individuals did not serve as 2011 raters in their district. Another 18% reported that only 1 or 2 uncalibrated individuals served as raters in their district. One DTC indicated that more than 20 uncalibrated individuals were raters in their district in 2011. DTCs indicated that uncalibrated individuals were paired with a calibrated rater so that the reliability and validity of the scoring process would be maintained, as is permitted by state policy.

### **CTC Questionnaire Results**

Of the 432 CTCs who responded to the CTC questionnaire, 35 (8%) also served as a DTC in 2011. Nearly all (431 of 432) CTCs reported using the *2011 TELPAS Manual for Raters and Test Administrators* in training their raters. The *2011 District and Campus Coordinator Manual* was used by 82% of CTCs, and the PowerPoint training slides developed by TEA were used by 74% of CTCs to provide rater training. Over a quarter of CTCs reported using other training materials. Many of these materials are locally developed presentations, guides, and tips.

Over 99% of CTCs reported having fewer than 51 individuals serving as TELPAS raters in 2011. Almost two-thirds of campuses had all raters calibrated by the end of calibration set 2; of the 115 remaining campuses, 103 (90%) had only 1 or 2 teachers left to finish calibrating. To help the few remaining teachers prepare for calibration set 3, 121 out of 124 (98%) CTCs reported using Model 1 in the *2011 District and Campus Coordinator Manual* (pgs. 275-276) with individual or group sessions. Model 2 was used by 16 of 124 (13%) CTCs that provided supplemental training. Supplemental

training was most often provided on the CTC's campus (54 out of 127 or 43%) or elsewhere in the district (68 out of 127 or 54%). CTCs reported that a trained supplemental support provider was able to provide supplemental support to all raters who were required to receive additional training. Only 11 CTCs reported using uncalibrated individuals as raters in 2011 and, in all cases, there were only 1 or 2 uncalibrated individuals. Based on CTC comments, it appears that the uncalibrated raters were paired with one or more trained calibrated raters to provide TELPAS ratings, as is permitted by state policy.

Procedures identified by CTCs used to support the validity and reliability of the TELPAS rating process included:

- all language domains raters collaborated with other teachers of the students in determining the students' ratings (74%).
- raters collaborated with each other in determining the ratings of students near the border between proficiency levels (72%).
- a district-determined selection of writing collections was rated a second time, and if the ratings differed, resolution ratings were provided and used (33%).
- all writing collections were rated a second time, and if the ratings differed, resolution ratings were provided and used (29%).

### **Teacher Questionnaire Results**

The teacher questionnaire, which was submitted by 536 raters after they completed the audit documentation form, included a number of questions that can be used to summarize characteristics of the teachers audited and evaluate the degree to which raters followed TELPAS training and administration procedures.

- About 84% of the respondents were language arts teachers, 63% were mathematics teachers, 62% were science teachers, and 70% were social studies teachers. About 4% indicated that they were not the teachers of record of these foundation subjects.
- During the 2011 TELPAS administration, 50% of teachers rated between 11 and 25 students, which is about the same number of students in a single classroom. Only 10% of teachers rated 26 or more students.
- Approximately 99% of teachers reported having enough information about the English language abilities of their students to make judgments in each language domain of listening, speaking, and writing.
- Less than 15% of the raters in the study were in their first year of rating TELPAS. Of those with more than one year of experience, around 51% had 5 or more years of rating experience. The majority (72%) of new raters attended state-recommended professional development training in the fall, which is geared toward improving ELL instruction throughout the year and laying the foundation for in-depth rater training in the spring.
- TELPAS general administration procedures training, which precedes holistic rating training, was typically held in 2 hours or less (69%). However, 11% of teachers reported sessions 4 or more hours long.
- With regard to spring 2011 holistic rating training, more than a third of the audited teachers, regardless of whether they were new or returning raters, were required to take the online basic training course designed for new raters. The online basic training course was completed in 3 or

fewer hours by 68% of teachers. Almost 9% reported needing more than 5 hours to complete the training. However, nearly all (99%) of these teachers reported that the course was beneficial in preparing them to rate students. The two teachers who indicated that the training was not helpful reported having already covered the material in previous years.

- Of the teachers in the audit, 80% reported successfully calibrating on the first attempt, 95% successfully calibrated on the first or second attempt, and 99% successfully calibrated by the third attempt. Nearly all (98%) teachers reported that the calibration activities were beneficial in preparing them to rate students.
- Teachers that required supplemental support reported being given additional individual or group training, reviewing the *TELPAS Manual for Raters and Test Administrators*, reviewing writing collections, discussing the rubrics, and reviewing practice students from the online training courses. Only four teachers reported not being successful after the third calibration set. These teachers mentioned receiving additional training, having a designated teacher support person, ESL and bilingual teacher support, and support from the school counselor during the TELPAS assessment window.

### **Summary of Listening and Speaking Audit and Audit Questionnaire Analyses**

The 2011 Listening and Speaking Audit process provided evidence of the types of activities teachers were using to evaluate students and how well teachers could support their student ratings. The 2011 audit approach avoided some of the timing concerns, logistical challenges, and time-consuming teacher involvement associated with the 2008 pilot. Although all audit processes used to provide validity evidence for the TELPAS listening and speaking ratings are intensive, a benefit of the new process was that it was not overly time-consuming for teachers, requiring only 1-2 hours of their time. However, the new process was not without limitations. First, teacher documentation was assumed to be accurate and evaluated accordingly. There was not a way to verify the accuracy of what a teacher reported about a student or observational activity. Second, despite instructions on how to fill out the documentation and examples of the level of specificity required, some teachers provided information that was too general or unclear to use in substantiating a student's rating.

However, results indicate that most district testing coordinators, campus testing coordinators, and teachers understand and appropriately participate in the TELPAS training sessions, online calibration, and TELPAS administration. In addition, over 75% of teachers submitted adequate justification of the ratings assigned to their students. This number may increase if future instructions can be improved to elicit more specific and clear documentation from teachers. These results provide validity evidence for TELPAS by indicating that most Texas teachers appear to be well-trained to rate the listening and speaking ability of ELL students, and that they are appropriately applying the scoring rubrics to rate students.

## Appendix A: Teacher Documentation Form

### Listening and Speaking Audit Rating Documentation Form Rating Information

Provide the information requested below.

Place the mouse cursor over the Sample Response boxes in each section if you would like to see example responses.

CONTACT INFORMATION			
District	A.W. Brown-Fellowship Leadership Academy	Campus	A.W. Brown-Fellowship North Campus
Student's Grade	4	Collaboration?	Yes
Rated—Listening	Yes	Rater's Name	a a
Rated—Speaking	Yes	Role	Bilingual Education Teacher

  

LISTENING	
What did you rate the student in LISTENING?	<input type="radio"/> Beginning <input type="radio"/> Intermediate <input type="radio"/> Advanced <input type="radio"/> Advanced High
Is this student in the early or late stages of this proficiency level?	<input type="radio"/> Early <input type="radio"/> Late <input type="radio"/> Not Particularly Early or Late <input type="radio"/> Unsure
LISTENING ACTIVITIES AND INTERACTIONS	
Use the spaces below to describe the academic activities and social interactions during which you observed the student's listening proficiency. Include activities from teachers you collaborated with, as applicable.	
Academic (Instructional) Activities  Character Count <div style="text-align: center; font-weight: bold;">0</div> (Max. 3000 characters)  <div style="background-color: #90EE90; border: 1px solid black; padding: 2px; text-align: center; font-size: small;">Sample Response</div>	<div style="background-color: #FFFF00; border: 1px solid black; height: 60px; width: 100%;"></div>
Social (Informal) Interactions  Character Count <div style="text-align: center; font-weight: bold;">0</div> (Max. 3000 characters)  <div style="background-color: #90EE90; border: 1px solid black; padding: 2px; text-align: center; font-size: small;">Sample Response</div>	<div style="background-color: #FFFF00; border: 1px solid black; height: 60px; width: 100%;"></div>
OBSERVATIONS AND RATING JUSTIFICATION	
Describe your observations of the student's listening proficiency (and observations of others you may have collaborated with) based on the activities and interactions described above. Explain how the observations support your rating.	
Observations and Rating Justification—Listening  Character Count <div style="text-align: center; font-weight: bold;">0</div> (Max. 3000 characters)  <div style="background-color: #90EE90; border: 1px solid black; padding: 2px; text-align: center; font-size: small;">Sample Response</div>	<div style="background-color: #FFFF00; border: 1px solid black; height: 60px; width: 100%;"></div>

### Listening Academic (Instructional) Activities Sample Response:

- In the past weeks, I've been observing the student's listening proficiency in these kinds of ways:
- I conference with her regularly about independent reading assignments and pay attention to her ability to understand the questions I ask about the stories. I also observe how well she is understanding her classmates during group work about the stories.
  - I've engaged her frequently in answering my questions about the science unit we're doing on sun and weather patterns.
  - During whole-class discussion of social studies lessons, I've keyed in on whether her contributions show she is understanding her classmates' responses and how much she needs me to accommodate my speech.
  - I watch her following verbal instructions for classroom assignments.
  - During small-group math tutoring sessions, I pay attention to how well she is understanding language vs. math content.

## Listening Social (Informal) Interactions Sample Response:

I've routinely interacted with this student or watched her interact with other students in these kinds of ways:

- Conversations about family or after-school activities
- Following instructions about routine classroom procedures
- Comprehension of social language used by friends and classmates during recess and free time
- Engagement in class discussion about recent field trip we took

## Observations and Rating Justification – Listening Sample Response:

During the reading conferences, the student needs little time to process what I'm saying before responding to my questions. In the science units we've been covering, she understands the language I use to teach the main concepts, and I seldom need to rephrase my language or use verbal cues for her. She has shown little difficulty following whole-group science and social studies discussions, although here and there she may ask me to repeat something if I'm talking too fast. When she interacts with friends, she follows conversations easily. She responds quickly and needs help to understand everyday conversations about family or after-school activities only rarely, such as when someone uses a very uncommon expression or word. During math tutoring, difficulties she has with oral comprehension are with concepts. Such difficulties could occur with any student and are not specific to the learning of academic English.

SPEAKING	
What did you rate the student in Speaking?	<input type="radio"/> Beginning <input type="radio"/> Intermediate <input type="radio"/> Advanced <input type="radio"/> Advanced High
Is this student in the early or late stages of this proficiency level?	<input type="radio"/> Early <input type="radio"/> Late <input type="radio"/> Not Particularly Early or Late <input type="radio"/> Unsure
SPEAKING ACTIVITIES AND INTERACTIONS	
Use the spaces below to describe the academic activities and social interactions during which you observed the student's speaking proficiency. Include activities from teachers you collaborated with, as applicable.	
Academic (Instructional) Activities  Character Count 0 (Max. 3000 characters) <a href="#">Sample Response</a>	
Social (Informal) Interactions  Character Count 0 (Max. 3000 characters) <a href="#">Sample Response</a>	
OBSERVATIONS AND RATING JUSTIFICATION	
Describe your observations of the student's speaking proficiency (and observations of others you may have collaborated with) based on the activities and interactions described above. Explain how the observations support your rating.	
Observations and Rating Justification—Speaking  Character Count 0 (Max. 3000 characters) <a href="#">Sample Response</a>	
<input type="button" value="Go to Next Step"/>	

### **Speaking Academic (Instructional) Activities Sample Response:**

Examples of speaking observations and interactions:

- I made sure to engage this student in a number of class discussions about Martin Luther King, Jr. I also gave her opportunities to speak in extended ways by asking her to summarize key points of the discussions and to recount what she learned from a video about his life.
- During a science lesson, I watched her speak as she worked with a partner to practice for and then give an oral presentation about relationships in a food chain.
- During math tutoring sessions after school, I routinely ask her to explain her problem-solving strategies.
- During guided reading, I listen carefully to her ability to discuss and react to the stories.
- When she works on assignments and doesn't understand something, I observe how well she uses English to ask for help and explain what she doesn't understand.

### **Speaking Social (Informal) Interactions Sample Response:**

Here are some ways I've observed her speaking proficiency:

- I routinely interact with her and watch her talk to friends during recess and while waiting for the bus.
- During a class discussion after a recent field trip, I watched her talk about what she enjoyed and gave her opportunities to go into detail and add to what her classmates said.
- I find opportunities to converse with her throughout the day to see her ability to talk at length about everyday events and topics of interest. For example, as I was preparing to rate her for TELPAS, she shared with me all the things that happened at her recent birthday party.

### **Observations and Rating Justification – Speaking Sample Response:**

- During science and social studies discussions such as those listed above, she typically speaks pretty comfortably after the topics and academic vocabulary become familiar to her. She uses common present and common past tenses without hesitation but sometimes makes mistakes
- When she's asked to speak in extended ways, she's able to connect ideas and sometimes tries using complex sentences. She has some difficulty narrating and describing in detail, and she occasionally hesitates as she searches for the right words and exactly how to explain things in English. Recounting what she learned from the Martin Luther King video, which presents a lot of details and new information, was more challenging for her than the oral presentation on the food chain, which allowed her to use academic vocabulary she'd been learning in class. She is usually able to speak in detail about familiar math content and problem-solving strategies during math tutoring sessions.
- She is generally comfortable speaking to her friends, me, and other teachers during casual conversations. She sometimes mispronounces words or makes linguistic errors, but her friends and teachers don't usually have trouble understanding her.

## Appendix B: Teacher Instructions

### March 1, 2011 TELPAS Audit Notification Letter—Attachment

#### RATER INSTRUCTIONS

#### Spring 2011 TELPAS Audit

##### Overview

Your campus has been selected for an audit that the Texas Education Agency (TEA) is conducting to evaluate TELPAS training and administration processes and the grades 3–5 listening and speaking proficiency ratings assigned during the spring 2011 administration. The information gathered will be used to plan for future TELPAS administrations and provide assessment validity and reliability data in fulfillment of federal requirements.

On each selected campus, the ratings of 1–3 students will be audited. You are the TELPAS rater of one or more of the selected students. See the enclosed student name(s).

##### Audit Instructions for Raters

After completing the TELPAS rating process, access the Web-based audit form at the link below, and follow the onscreen instructions to—

1. document how you determined the listening and speaking proficiency ratings of the selected student, and
2. complete a questionnaire related to your experiences in fulfilling spring 2011 TELPAS training and administration requirements for all language domains.

The Web-based form will be available at the link below by **March 21, 2011**. It must be completed and submitted no later than **April 18, 2011**.

[www.txetests.com/TELPASAUDIT](http://www.txetests.com/TELPASAUDIT)

Do not fill out the audit form until you have completed the rating process. Check with your campus testing coordinator to determine when the TELPAS Student Rating Roster with your recorded ratings will be signed by your campus principal and considered final.

Do not complete the audit form for students not selected for this audit.

**Your online documentation form is required to be completed and submitted by April 18, 2011.**

For questions about how to use the Web-based form, contact the TEA Student Assessment Division at 512-463-9536. For other questions about the audit process, contact your campus coordinator. Thank you for your assistance in ensuring the integrity of this statewide assessment program.



## Appendix C: Teacher Questionnaire

1. **How long did it take you to complete the listening and/or speaking rating documentation?**

- Less than 30 minutes
- 30 minutes – 1 hour
- 1 – 2 hours
- 2 – 3 hours
- Longer than 3 hours

2. **Was the rating documentation form worded clearly?**

- Very clear
- Generally clear
- Not clear

**If parts were unclear, please explain.**

3. **If you have suggestions for improving this audit process, please explain.**

4. **What grade(s) do you currently teach? (Mark all that apply)**

- K
- Grade 1
- Grade 2
- Grade 3
- Grade 4
- Grade 5
- Grade 6
- Grade 7
- Grade 8
- Grade 9
- Grade 10
- Grade 11
- Grade 12
- None of the above

**If you have responded "None of the above", please explain.**

5. **In which foundation subjects are you a teacher of record for the student selected for this audit? (Mark all that apply)**

- Language Arts (reading and writing)
- Mathematics
- Science
- Social Studies
- None of the above

6. Approximately how many students were you assigned to rate for the spring 2011 TELPAS administration?
- 5 or fewer
  - 6 to 10
  - 11 to 25
  - 26 to 50
  - More than 50
7. Did you feel that you had enough information about the English language abilities of your students to make judgments about their proficiency levels in each domain – LISTENING?
- Yes
  - No
8. Did you feel that you had enough information about the English language abilities of your students to make judgments about their proficiency levels in each domain – SPEAKING?
- Yes
  - No
9. Did you feel that you had enough information about the English language abilities of your students to make judgments about their proficiency levels in each domain – WRITING?
- Yes
  - No
- If you responded No for any domain, please explain.
10. Is this your first year to be a TELPAS rater?
- Yes
  - No
11. If this is not your first year, how many years have you been rating students?
- 2
  - 3
  - 4
  - 5 or more
12. If this is your first year, did you attend professional development training in the fall to help familiarize you with the TELPAS assessment process?
- Yes
  - No

**13. How long was the spring TELPAS administration procedures training session your district held related to holistic rating training and assessment requirements?**

- Less than 30 minutes
- 30 minutes to 1 hour
- 1 to 2 hours
- 2 to 3 hours
- More than 4 hours
- No administration procedures training

**If you did not attend a session, please explain.**

**14. Were you required to take this spring's online basic training course designed for new raters?**

- Yes
- No

**15. About how many hours did it take for you to complete the course?**

- Less than 1 hour
- 1 to 2 hours
- 2 to 3 hours
- 3 to 4 hours
- 4 to 5 hours
- More than 5 hours

**16. Was the course beneficial in preparing you to rate your assigned students?**

- Yes
- No

**If you responded No, please explain.**

**17. What was your calibration experience?**

- Success Set 1
- Success Set 2
- Success Set 3
- Not successful

**18. Were the calibration activities beneficial in preparing you to rate your assigned students?**

- Yes
- No

**19. Describe the supplemental support you were provided to help prepare you for Set 3.**

**20. Describe the support you received during the TELPAS assessment window as you rated your assigned students.**

## Appendix D: Audit Form for Evaluating Teacher Documentation

Audit ID #: \_\_\_\_\_

Reviewer: \_\_\_\_\_

Flag for group review: \_\_\_\_\_

**Teacher's Listening Rating:** \_\_\_\_\_

**Teacher's Speaking Rating:** \_\_\_\_\_

### Rater's Adherence to Rating Rubrics (PLDs)

1. Does PLD-related documentation support the rating assigned?

**Listening:**    Yes    No    Incomplete documentation  
   provided

If No: (select from the following reasons)

A - Rating supported by factors **primarily** outside PLDs such that there is not enough rubric-related documentation to substantiate the rating

B - Rating based on PLDs but documentation supports different proficiency level (e.g., rating assigned is ADV but support points overwhelmingly to INT)

C - Supporting information too general or unclear overall to substantiate assigned level

**Speaking:**    Yes    No    Incomplete documentation  
   provided

If No: (select from the following reasons)

A - Rating supported by factors **primarily** outside PLDs such that there is not enough rubric-related documentation to substantiate the rating

B - Rating based on PLDs but documentation supports different proficiency level (e.g., rating assigned is ADV but support points overwhelmingly to INT)

C - Supporting information too general or unclear overall to substantiate assigned level



## Appendix E: Audit Meeting Agendas



### Listening and Speaking Audit Agenda

August 11, 2011

8:30 AM – 4:30 PM

Congress Conference Room

#### I. Overview

*A. Description of process and review of materials*

#### II. Group Review

*A. Discussion of first 10 sets of teacher materials*

*B. Modifications to rating form*

*C. Discussion and ratings of 25–50 sets of teacher materials*

#### III. Lunch

#### IV. Individual Review

*A. Individual ratings of 10 sets of teacher materials*

*B. Analysis of individual ratings*

*C. Discussion of individual results as a group*

#### V. Group Review (Continued)

*A. Discussion and ratings of remaining sets of teacher materials*

#### VI. Closing

*A. Debrief*

*B. Provide raters with materials for individual ratings and debrief*



**Listening and Speaking Audit Agenda  
August 18, 2011  
1:00 PM – 4:30 PM  
ATR Training Room**

**I. Process Clarification**

- A. Decisions and guidelines for reviewing materials***
- B. Modifications to rating form***

**II. Group Review**

- A. Review and discussion of first 10 sets of teacher materials***

**III. Individual Review**

- A. Individual ratings of 10 sets of teacher materials***
- B. Analysis of individual ratings***
- C. Discussion of individual results as a group***

**IV. Group Review (Continued)**

- A. Discussion and ratings of remaining sets of teacher materials***

**V. Closing**

- A. Debrief***
- B. Provide raters with materials for individual ratings and debrief***

## Appendix F: DTC Questionnaire

1. In addition to my role as a district testing coordinator, I also served as a campus testing coordinator for the 2011 TELPAS administration.

- Yes
- No

2. Approximately how many total individuals served as raters in your district in spring 2011? (Include both new and returning raters.)

- 1 to 10
- 11 to 25
- 26 to 50
- 51 to 100
- 101 to 200
- 201 to 300
- More than 300

3. Were all raters in your district who did not successfully calibrate by the end of set 2 given supplemental support by a trained supplemental support provider?

- Yes
- No
- Not applicable. All raters successfully calibrated by the end of set 2

If you selected No, please explain.

4. Did your district implement procedures to support the validity and reliability of the TELPAS rating process, as required by page 243 of the 2011 District and Campus Coordinator Manual?

- Yes
- No

If you selected Yes, proceed with questions 5–7. If you selected No, please explain why your district did not follow the state-required procedures. Then answer N/A (not applicable) for questions 5–7.

5. What procedures were implemented? (Mark all that apply.)

- All language domains, K-12: Raters collaborated with other teachers of the students in determining the students' ratings
- All language domains, K-12: Raters collaborated with each other in determining the ratings of students near the border between proficiency levels
- Writing, grades 2-12: A district determined selection of writing collections were rated a second time, and if the ratings differed, resolution ratings were provided and used
- Writing, grades 2-12: All writing collections were rated a second time, and if the ratings differed, resolution ratings were provided and used
- N/A
- Other

If you selected Other, please explain. Exclude supplemental support provider training and rating support for any uncalibrated individuals who served as raters.



6. Did the district establish district-wide procedures, or did campuses select the procedures to follow?

- The district selected the procedures for campuses to follow
- The district generally selected the procedures but allowed variability at the request of individual campuses
- Campuses were directed to submit plans to the district for approval
- The campuses established procedures independently (no district approval required)
- N/A
- Other

If you selected Other, please explain.

7. What steps did the district take to ensure that campuses implemented the validity and reliability procedures? (Mark all that apply.)

- The district required each campus to submit documentation describing the procedures followed
- The district required each campus to keep documentation of the procedures at the campus
- N/A
- Other

If you selected Other, please explain.

8. Approximately how many uncalibrated individuals (individuals who completed all 3 calibration sets unsuccessfully) served as raters in your district in spring 2011?

- Zero
- 1 to 2
- 3 to 5
- 6 to 10
- 11 to 15
- 16 to 20
- More than 20

If other than Zero, describe the process your district followed to monitor that campuses provided support to these raters in all language domains during the operational assessment. See page 243 of the 2011 District and Campus Coordinator Manual for more information.

## Appendix G: CTC Questionnaire

1. In addition to my role as campus testing coordinator for TELPAS, I also served as district testing coordinator for the 2011 TELPAS administration.

- Yes
- No

2. What materials were used in providing TELPAS administration procedures training to raters? (Mark all that apply.)

- 2011 TELPAS Manual for Raters and Test Administrator
- 2011 District and Campus Coordinator Manual
- PowerPoint Training Slides Developed by TEA
- Other

If you selected Other, please describe.

3. Approximately how many total individuals served as raters on your campus in spring 2011? (Include both new and returning raters.)

- 1 to 10
- 11 to 25
- 26 to 50
- 51 to 100
- More than 100

4. Approximately how many individuals who served as raters on your campus did not calibrate by the end of calibration set 2?

- 0
- 1-2
- 3-5
- More than 5

If you selected other than Zero, proceed with questions 5–7. If you selected Zero, answer N/A (not applicable) for questions 5–7.

5. What supplemental training model was used to help the individual(s) prepare for set 3? Model descriptions are found on pages 275–276 of the 2011 District and Campus Coordinator Manual. (Mark all that apply.)

- Model 1, individual session, page 275
- Model 1, group session, page 275
- Model 2, page 276
- N/A
- Other

If you selected Other, please describe.

6. Where was supplemental training provided for these raters?

- On our campus
- Elsewhere in our district
- At the education service center
- N/A
- Other

If you selected Other, please describe.

7. Were any raters unable to be provided supplemental support by a trained supplemental support provider?

- Yes
- No
- N/A

If you selected Yes, please explain the circumstances and what your campus did to guide the rater in preparation for set 3.

8. What procedures were implemented on your campus to support the validity and reliability of the TELPAS rating process? See page 283 of the 2011 District and Campus Coordinator Manual for more information. (Mark all that apply.)

- All language domains, K-12: Raters collaborated with other teachers of the students in determining the students' ratings
- All language domains, K-12: Raters collaborated with each other in determining the ratings of students near the border between proficiency levels
- Writing, grades 2-12: A district determined selection of writing collections were rated a second time, and if the ratings differed, resolution ratings were provided and used
- Writing, grades 2-12: All writing collections were rated a second time, and if the ratings differed, resolution ratings were provided and used
- Other

If you selected Other, please explain. Exclude supplemental support provider training and rating support for any uncalibrated individuals who served as raters.

9. Approximately how many uncalibrated individuals (individuals who completed all 3 calibration sets unsuccessfully) served as raters on your campus in spring 2011?

- 0
- 1-2
- 3-5
- More than 5

If other than Zero, describe the support your campus provided these raters as they assessed their assigned students. See page 283 of the 2011 District and Campus Coordinator Manual for more information. Address the support provided for each language domain.