



Research on the Integrated Performance Assessment in an Early Foreign Language Learning Program

Kristin Davin
Loyola University Chicago

Francis J. Troyan
University of Pittsburgh

Richard Donato
University of Pittsburgh

Ashley Hellman
Falk Laboratory School, Pittsburgh, Pennsylvania

Abstract: *This article reports on the implementation of the Integrated Performance Assessment (IPA) in an Early Foreign Language Learning program. The goal of this research was to examine the performance of grade 4 and 5 students of Spanish on the IPA. Performance across the three communicative tasks is described and modifications to IPA procedures based on the needs of the young learner are presented. Comparisons of the performance of monolingual and multilingual students and information collected through a student post-IPA survey are also reported. The researchers argue that in addition to benefits of the IPA to describe student communicative performance, integrate teaching and assessment, and promote standards-based teaching practices, the IPA also has the potential to identify strengths and weaknesses of elementary school foreign language programs.*

Key words: *Spanish, classroom-based assessment, elementary school language assessment, integrated performance assessment, integration of teaching and assessment*

Kristin Davin (PhD, University of Pittsburgh) is Assistant Professor of Foreign Language Education at Loyola University Chicago, Chicago, Illinois.

Francis J. Troyan (MA, University of Pittsburgh) is a doctoral student in Foreign Language Education at the University of Pittsburgh, Pittsburgh, Pennsylvania.

Richard Donato (PhD, University of Delaware) is Associate Professor of Foreign Language Education at the University of Pittsburgh, Pittsburgh, Pennsylvania.

Ashley Hellman (MAT, University of Pittsburgh) is a K–5 Spanish teacher at Falk Laboratory School, Pittsburgh, Pennsylvania.

Introduction

The purpose of this study is to examine the implementation of the Integrated Performance Assessment (IPA) in an Early Foreign Language Learning (EFL) program, to uncover patterns of performance in the three modes of communication, and to identify necessary adaptations when using an IPA at this level of instruction based on student commentary and performance. To date, IPA studies in foreign language education have focused on the middle school, high school, and university levels. One key to the health and longevity of well-articulated foreign language programs in elementary schools is careful documentation of student achievement (Donato & Tucker, 2010) and the assurance that students will continue to develop their language abilities across several years of instruction. For this reason, it is essential to conduct research on current professionally endorsed assessment practices at the elementary school level to ensure that we can describe clearly the outcomes of instruction to students and stakeholders. Additionally, assessments developed for the older learner and adults may not be categorically appropriate for the young learner and must be researched to determine where principled adaptations need to be made (Donato, 1998). Finally, research is needed on the assessment frameworks for young learners that seek to integrate instruction and assessment and reflect the communicative and cultural goals of language education as we currently understand them. To our knowledge, this study is the first to provide comprehensive documentation of the IPA process with novice level learners of Spanish in elementary school.

Review of Relevant Literature

Through a federally funded project, the American Council on the Teaching of Foreign Languages (ACTFL) created the IPA, an assessment prototype for measuring students' progress in meeting specific aspects of the *National Standards for Foreign Language Learning in the 21st Century*

[SFL] (National Standards, 2006). The IPA is a performance-based assessment that consists of an interpretive, interpersonal, and presentational communication task¹ aligned within a single theme in which a rater assesses students' proficiency in each mode of communication. Glisan, Adair-Hauck, Koda, Sandrock, and Swender (2003) developed a prototype of the assessment, conducted widespread field tests in six research sites, and have subsequently published the findings for the profession (Adair-Hauck, Glisan, Koda, Swender, & Sandrock, 2006). In their work, Glisan et al. (2003) set out to (1) design an assessment to measure students' progress in meeting the SFL, (2) research the effectiveness of the IPA as a standards-based assessment instrument, (3) assess the feasibility of implementing the IPA in a typical classroom, and (4) understand to what extent the IPA motivates teachers to change their practice (Adair-Hauck et al., 2006, pp. 363–364).

The IPA project has offered the profession an opportunity to reconceptualize classroom assessment. This performance-based system of measurement offers the classroom teacher an innovative framework for the development of assessments. Equipped with this framework, instructional conversations between and among students and teachers are transformed into ongoing dialogue regarding student performance (Adair-Hauck et al., 2006). Adair-Hauck et al. argue the potential of the IPA for washback effect, defined as “the effect and influence that the introduction of tests has on the educational context” (Shohamy, 2001, p. 46), on student learning and classroom instruction (Alderson & Wall, 1993; Green, 2006; Messick, 1996; Muñoz & Álvarez, 2010; Shohamy, 2001).

In the literature on school reform, Bryk and Gomez (2008) suggest a three-phased process for (1) development of innovation, (2) field trials, and (3) large-scale adoption, or what design innovation and school reform researchers have termed “going to scale” (McDonald et al., 1999; McDonald, Klein, & Riordan, 2009; Roschelle, Tatar, & Kaput, 2008). Since the completion of phases

1 and 2 and the subsequent publication of the IPA manual, *ACTFL Integrated Performance Assessment* (Glisan et al., 2003), two studies have been published on the IPA. In the first, Adair-Hauck et al. (2006) describe the development and field trials of the IPA prototype. Through qualitative data analysis, the researchers describe the effect of the IPA on teacher practice during field trials, which took place in middle school and high school foreign language classrooms. They found that IPA training assisted teachers in modifying and restructuring instructional activities to support standards-based curricula. They suggest further studies to investigate the efficacy of the IPA at other grade levels, as the IPA goes to scale through professional development opportunities provided by ACTFL.

To measure postsecondary students' progress toward meeting the SFL, Glisan, Uribe, and Adair-Hauck (2007) implemented an IPA in a postsecondary Advanced Spanish course. They found that students performed best on the presentational task and worst on the interpretive task. Based on quantitative analysis, they found (1) a positive correlation between years of study in middle school and performance on the interpersonal task of the IPA and (2) a negative correlation between the number of years of high school Spanish and performance across the three tasks. Of the studies on the IPA to date, none have attempted to describe the implementation of the IPA in an elementary school foreign language program.

The Study

Research Questions

This study attempts to answer four research questions: (1) How do elementary students who have studied Spanish for four years perform on an IPA? (2) What patterns of performance in the three modes of communication are revealed in the ratings? (3) What adaptations, if any, are needed in the use and scoring of IPAs for the young learner? and (4) What are students' perceptions of the IPA and attitudes toward preparation for the IPA?

Context

To answer these questions, a 16-week study of an EFLL program was conducted at a laboratory school in Western Pennsylvania. This K–8 private school enrolls approximately 300 students, and having just completed an expansion project, plans to expand enrollment by about 10% each year for the next five years. The population of this school resembles that of the surrounding areas in that approximately 10% of students are African American, Asian, Hispanic, and Indian minority students. Begun in the fall of 2006, the Spanish program makes use of a curriculum taught by an itinerant teacher during 15 minutes of daily instruction. The curriculum at this school is composed of semester-long thematic units (e.g., *A Trip to Perú*, *All About Me*) designed to meet all of the SFFLs (National Standards, 2006). Details on the instructional approach will be provided in the context of interpreting the findings. As we will argue, student performance on the IPA has the potential to describe the strengths and areas in need of improvement of foreign language programs.

For the present study, data were collected for 48 students. Students who had fewer than four years studying Spanish or who were absent for one of the three IPA tasks were omitted from the analysis. Once these omissions were made, complete data sets for 30 students remained and were analyzed. Twenty-one of these students were fifth graders and nine were fourth graders. All students in the data set had studied Spanish for four years.

Description of the IPA

The IPA was situated in a semester-long unit on *El Mar Caribe* (The Caribbean Sea). In this unit, students studied animals that live in the ocean, the levels of the ocean in which they live, beach accessories, and beach activities. The activities throughout this unit reflected IPA tasks in the interpretive, interpersonal, and presentational modes. For example, students watched a Spanish video without English or Spanish

TABLE 1

Schedule of IPA

Tuesday	IPA Overview and Review of Interpretive Rubric
Wednesday	Interpretive Task (Caillou video and comprehension task)
Thursday	Feedback on Interpretive Task (Review of rubrics; Discussion of errors)
Friday	Introduction to Interpersonal Task and Rubric
Monday	Interpersonal Task with 1st class (1 hour long) (Paired conversations about differences in pictures)
Tuesday	Interpersonal Task with 2nd class (1 hour long) (Paired conversations about differences in pictures)
Wednesday	Feedback on Interpersonal Task (Review of rubrics; Discussion of errors)
Thursday	Presentational Task (Preparation of magazine articles)
Friday	Feedback on Presentational Task (Review of rubrics; Discussion of errors)

subtitles about the beach and completed comprehension questions, participated with a partner in an information gap activity, and presented descriptions of ocean life. These activities exposed students to the IPA process and rubrics during group feedback sessions with the two classes. After four months of instruction centered around this theme, students participated in an IPA lasting nine days. Table 1 outlines the schedule for the IPA.

Interpretive Task and Feedback

According to the IPA manual, appropriate novice-level texts for the interpretive task should be “strongly supported by contexts, usually visual, with content of a frequent everyday nature” (Glisan et al., 2003, p. 28). Students began the IPA with a contextualized situational prompt that stated, “You are a freelance reporter for *Viajar*, a Spanish language magazine for travelers. Your boss wants you to write an article on the Caribbean. So, you begin by researching life at the beach.” A two-minute portion of a video entitled “Caillou se va a la playa”

available on *YouTube.com* was selected. In this video, a child named Caillou spends a day at the beach with his family observing wildlife, swimming, and playing in the sand. A comprehension task (see Appendix A) was created according to the interpretive task template in the IPA manual (Glisan et al., 2003; Appendix C). Because the IPA template is based on an interpretive reading task and this study implemented an interpretive listening task, some modifications were made to the contents of the template. The interpretive task for novice learners contains three sections. In the first section, 10 key words from the video are listed in English (e.g., vacation, beach, picnic) and students are asked to write the word in Spanish. Although only three of these key words were cognates, students had been exposed to all 10 words during the unit or at some other time in their study of Spanish.² In the second section, *Important Ideas*, eight phrases were listed (e.g., A song is sung about the sea). Students were asked to identify which phrases were true and which were false. In the third section, students were asked to write the main

idea of the video. Although not present in the IPA novice-level interpretive template, the researchers added a fourth section to probe students who might be approaching the intermediate-low level. In this section, students were asked to infer the meaning of three phrases in the video that they had not previously studied (e.g., ¡Qué asco!, ¡Mira!).

On the following day, students received their scored interpretive tasks and were given an opportunity to watch the video again. During this time, their teacher carefully led the students through the video in order to help them find the correct answers and understand where misinterpretations might have occurred. The class reviewed the correct answers and discussed the meaning of the scores on the rubrics. As stated above, the students were familiar with the scoring rubrics as they had been used throughout the semester.

Interpersonal Task and Feedback

For the interpersonal task, students were paired with a classmate who had a similar proficiency level. The reason for this system of pairing was to prevent an imbalance in conversation that might silence less proficient students. A student teacher supervised the class as the Spanish teacher and one of the university researchers assessed the interpersonal communication of pairs of students. Students were given five minutes to converse about the prompt below in a quiet room. All conversations were videotaped for analysis.

For the interpersonal task, the IPA Manual suggests that “each of the two speakers comes to the task with information that the other may not have, thereby creating a real need for students to provide and obtain information through the active negotiation of meaning” (Glisan et al., 2003, p. 20). To achieve the information gap described above, each student in the dyad was given one of two pictures of Caillou at the beach. The pictures were accompanied by the following directions:

You and your partner work for a huge magazine company that has offices in Pittsburgh and San Juan, Puerto Rico. The June issue of the magazine has to go to press today, but there is one big problem. The printing machines messed up and changed the images some, so you have to fix this. One of you lives and works in Pittsburgh, the other in Puerto Rico. You must ask your partner questions to determine the differences between the two pictures. Your boss is listening in on this conference call and speaks only Spanish. Try to be as descriptive as you can and use as much detail as possible. The more complete sentences you can use the better. (See Appendix B for the entire task sheet.)

Objects in the two pictures differed slightly in color, number, and location. For example, picture A depicted three palm trees while picture B depicted only two. Looking at only their picture, students were instructed to converse to discover all of the differences in the two pictures.

Prior research describing novice level performance in EFL programs indicates that young L2 speakers need L1 support during oral assessment to demonstrate more fully their speaking proficiency (Igarashi, Wudthayagorn, Donato, & Tucker, 2002). Anticipating the need of such assistance during the interpersonal task, four prompts were developed and provided by the assessment facilitator, if and when necessary. The following prompts were given orally as needed by the assessment facilitator during the interpersonal task: (1) Talk about any differences that you see; (2) Look for differences in color; (3) Sometimes objects are added or omitted; and (4) Sometimes the location of objects is changed.

After the interpersonal task was completed for each class, the teacher returned scored rubrics to each student, provided feedback to the class as a whole, and discussed with the class the meaning of the scores. The teacher then showed the

students examples of their utterances from the interpersonal conversations, such as complete utterances, well-formed questions, and utterances that were a mixture of Spanish and English. This activity mirrored an interpersonal activity that occurred earlier in the semester in which the students practiced scoring conversations with rubrics.

Presentational Task and Feedback

For the presentational task, students were asked to write a magazine article in which they described the Caribbean. They were given the following writing prompt:

Your boss at the magazine needs your help yet again. He has asked you to write several paragraphs describing the Caribbean in which you talk about the weather, activities you can do there, how to get around, and types of wildlife. Because this is for publication, please be sure that your writing is neat and can be read.

Students worked on this task independently without notes or dictionary. Although not required, some students added illustrations to accompany their writing. Following the presentational task, the Spanish teacher returned students' magazine articles and rubrics. Feedback on common errors in their written magazine articles was given to the students in the group setting.

Methodology

The Collaboration

To pilot the IPA with elementary school Spanish language learners, a university/school partnership was formed. The two principal researchers in the study, Davin and Troyan, met weekly throughout the spring term of 2010 with the K–5 Spanish teacher, Ms. Hellmann, a first-year teacher and a recent graduate of a MAT program in foreign language education. As part of her preservice coursework, Ms. Hellmann was trained in foreign language assess-

ment design and implementation, which included the creation of a unit of study linked to the IPA. Davin was the instructor of the testing and assessment course for which the teacher created the IPA. Davin also designed and implemented the Spanish program at the Spanish teacher's current school in 2006 where she taught K–5 Spanish for three years. Troyan, also a doctoral student in foreign language education, designed and implemented a standards-based high school curriculum connected to the IPA and currently teaches testing and assessment to preservice teachers. Donato is a specialist in foreign language education and has considerable experience in research on EFL programs and language learning in young children. Through collaborative meetings with the teacher, the interpretive, interpersonal, and presentational tasks were designed and concerns of the teacher were addressed.

Data Collection

Data collection occurred throughout the course of one semester in the form of field notes, artifact collection, interviews, journaling, focus groups, surveys, and assessment scores. Throughout 12 weeks during the spring of 2010, two researchers alternated observing the two combined grade 4 and 5 sections. During observations, field notes were written focusing on the teacher's instructional activities and assessments.

During the implementation of the IPA, student performance on each task was documented. For the interpretive and presentational tasks, photocopies were made of students' completed performance and their rubric. For the interpersonal task, each dyad was video-recorded and photocopies were made of each rubric. Scoring for each task was done using the novice level rubrics from the ACTFL IPA guide. To ensure rater reliability, the researchers and teacher first rated five samples together for each task. They then double rated all of the students' performances. If there was a discrepancy for a particular rating, the

researchers and teacher collaborated and agreed upon a rating.

Once the data from each of the IPA tasks were placed into the spreadsheet, the number of students who received *exceeds expectations*, *meets expectations*, or *does not meet expectations* for each task was counted. Students for which Spanish was their L3 or L4 were identified and their scores on the tasks were compared to the scores of students for which Spanish was their L2. These results were aligned with survey data collected from the students that asked about their language background and their perceptions of the IPA (see the section on student responses to post-IPA questionnaire).

In addition to data collected from students, data were also collected from the Spanish teacher. The researchers interviewed the teacher before observations began in February of 2010. During this interview, the teacher described her current assessment practices, her prior experience with the IPA, her feelings about implementing an IPA with young learners, and challenges that she thought she might encounter. A second interview was conducted in June of 2010 after the IPA had taken place and once the semester had ended. During this interview, the teacher reflected on the IPA, spoke about difficulties she experienced with the IPA, and described changes to her teaching practices brought about through her experience with the IPA. In addition to the interviews, the teacher responded to journal prompts weekly from February to May of 2010. The researchers provided these prompts based on discussions that had taken place or instruction observed in her classes. Sample journal prompts included:

Regarding our work on Monday, March 22, could you reflect on the rubric that we created and the decisions that were made together during the meeting? Also, do you think that the rubric will have an impact on the assignment?

and

Today your students have completed the interpersonal task of the IPA. Could you reflect on this experience? Did things go the way that you had anticipated? What did you expect? What surprised you? Were there any issues that came up that you didn't expect?

Data from the two interviews and 11 journal entries were coded and analyzed similarly. The findings from this analysis are beyond the scope of this publication and will be reported in a future publication.

Data Analysis

Overall Performance

Descriptive statistics were used to analyze students' performance across the three IPA tasks, as shown in Table 2. As Table 2 indicates, students attained the highest ratings on the interpersonal task with over half of the students receiving ratings in either the *exceeds expectations* or *meets expectations* categories. Seventeen out of 30 students *met* or *exceeded expectations* for interpersonal communication, whereas only 13 students *did not meet expectations*. Five of these 17 students *exceeded expectations* on the interpersonal task. The lowest ratings for performance occurred on the interpretive task. On the interpretive task, only 8 students out of 30 *met expectations* whereas 22 out of 30 students *did not meet expectations*. On the presentational task, 12 out of 30 students *met* or *exceeded expectations*, whereas 18 students did not meet expectations.

As previously discussed, three probing questions were added to the interpretive task to document the performance of students who might be approaching the intermediate-low level of proficiency. According to the IPA Manual (Glisan et al., 2003), novice level students are capable only of literal comprehension and cannot infer the meaning of language from context, an assertion that we argue needs further investigation. To assess whether any of these students could interpret beyond the literal meaning of the text, a section was

TABLE 2

Performance Across Modes of Communication, *n* = 30

Task	Exceeds Expectations	Meets Expectations	Does Not Meet Expectations
Interpretive	N/A*	8	22
		27%	73%
Interpersonal	5	12	13
	17%	40%	43%
Presentational	4	8	18
	13%	27%	60%

*Based on the interpretive novice level rubric, students cannot exceed.

TABLE 3

Correlation of Student Opinion of Difficulty and Task Performance

IPA Task	Student Opinion of Difficulty
Interpretive	-.54**
Interpersonal	.03
Presentational	-.66**

n = 24, ***p* < .01.

Student Performance and Perception of Task Difficulty

Following the descriptive analysis, statistical correlations were established between student performance on each task and their opinion of task difficulty. A numerical value was assigned to each student's rating on the three IPA tasks: *exceeds expectations* = 3, *meets expectations* = 2, *does not meet expectations* = 1. After completing the IPA, students responded to a questionnaire, in which they rated the difficulty of each task on a 1–5 scale, with 1 as *very easy* and 5 as *very difficult*. Spearman's rho correlations were computed and analyzed to determine whether a relationship existed between students' perceptions of task difficulty and their actual task performance. Summaries of the correlations are presented below.

Table 3 summarizes the correlations of students' performance on each task with their evaluation of the task difficulty. Significant negative correlations were found between students' opinion of task difficulty and their ratings on the interpretive and presentational tasks ($r = -.54, p < .01$ and $r = -.66, p < .01$, respectively). The negative correlations indicate that as the students' perceptions of task difficulty increased, task performance decreased. The importance of this finding is that students

added to the interpretive task to determine if students could infer the meaning of three unfamiliar phrases in new contexts.³ Based on these interpretive questions, 6 out of 30 students were able to infer the meaning of two of the three phrases, whereas six additional students were able to infer the meaning of one of the three phrases. No students were able to infer the meaning of all three phrases. This finding suggests that some elementary school students at novice levels of proficiency are capable of inferring the meaning of unfamiliar phrases and that assessment of the interpretive abilities of these students should not categorically be excluded from the IPA.

TABLE 4

Interaction Between Pairs of IPA Task Scores

	Interpretive	Interpersonal	Presentational
Interpretive	–	–	–
Interpersonal	(.19)	–	–
Presentational	(.15)	(.68)**	–

n = 30, ***p* < .01.

appear to accurately assess their abilities on the interpretive and presentational tasks. Students' perceptions of difficulty align with their task performance. Perceived task difficulty on the interpretive and the presentational tasks is reflected in their final scores. There is no correlation with students' perceptions of difficulty and performance in the interpersonal task. The results of this analysis suggest that students are aware of their strengths and weaknesses relative to two of the IPA tasks and allude to the IPA's potential to raise students' metacognitive ability to assess their own learning and set goals for addressing areas that need improvement. Later in this article, we will provide corroborating evidence of this potential through a descriptive analysis of students' narrative survey responses.

Integrated Nature of the IPA

The integrated nature of the IPA tasks assumes a relationship between one task and the next. In other words, the learning that is to occur in the interpretive task becomes the basis for the face-to-face interaction during the interpersonal task. In turn, the information gleaned from the first two tasks is used to construct the presentational task. Seeking to explore this task relationship statistically, Spearman's rho correlations were performed on each possible pairing of task scores.

As Table 4 indicates, the relationships between the interpretive and presentational task scores (*r* = .15, *p* = .42) and the

interpretive and interpersonal task scores (*r* = .19, *p* = .31) were not significant. However, a statistically significant correlation was found between the scores on the interpersonal and presentational tasks of this IPA (*r* = .68, *p* < .001). This finding suggests consistency between performance on the interpersonal and presentational tasks. Furthermore, it indicates a discontinuity between the interpretive task and the other two tasks of this IPA. The observed relationship suggests the need to explore the nature of the direction of the relationship between the two constructs, and the relationship between the communicative modes of the IPA, in general. The absence of a relationship between the interpretive task and the other two tasks is intriguing but perhaps not surprising. Assuming that a learner's productive abilities in the foreign language are parallel to those of their reading and listening comprehension abilities would indeed be a misguided assumption. While the three modes of communication are used in concert in authentic communication, this does not mean that one's ability across the modes is the same (E. Glisan, personal communication, November 8, 2010). Additionally, the lack of a relationship between the interpretive task and the other two tasks, coupled with the low overall performance of the students on the interpretive task, are similar to the findings of Glisan et al. (2007), who also found the weakest performance on the interpretive task in their study of students in a postsecondary Advanced Spanish course.

TABLE 5

Mean Scores by Language Group

Language Group	Interpretive	Interpersonal	Presentational
Second Language Speakers ($n = 22$)	1.32	1.64	1.4
Third Language Speakers ($n = 7$)	1.14	2	1.57
Fourth Language Speakers ($n = 1$)	2	3	3

1 = does not meet expectations; 2 = meets expectations; 3 = exceeds expectations.

This finding points to an overall trend in the classroom involved in this study and perhaps, in general, across classroom settings

Student Language Background

The student survey also provided data regarding the number of languages known by the students. Language background data were then compared to student performance on each of the three IPA tasks. Our survey indicated that 22 of the 30 students were currently learning Spanish as a second language. We referred to them as *Second Language Speakers*. Seven of the 30 students were *Third Language Speakers*; they spoke Spanish, English, and another language at home. One student was a *Fourth Language Speaker*. She spoke Spanish, English, and two additional languages (French and Italian) at home. For each group, the mean score for each task was calculated with 3 being the highest rating and 1 being the lowest.

Table 5 depicts the mean performances on the three tasks by language group. The most salient trend in these data is the students' performance on the interpersonal task, which steadily improves across the three groups as the number of languages spoken increases. The scores on the presentational task also show increases across the groups. The interpretive task performance was not consistent across groups, however. The data for the interpersonal and presentational tasks suggest that as number of languages spoken increases, so does

the performance across these tasks of the IPA. Interpretive communication was unaffected by students' language background, further suggesting that performance on the interpretive mode of communication is not analogous to other modes of communication and may be affected by factors and develop in ways that are not equivalent to the interpersonal and presentational communicative modes. This point will be addressed further in the Discussion section.

Summary of Findings

In summary, students attained the highest levels of performance on the interpersonal task. In contrast, over two thirds (22) of the 30 students in the study *did not meet expectations* on the interpretive (listening) task. Twelve of the 30 students *met expectations* on the presentational task. Analysis of students' opinions of their performance on the three tasks in comparison with their actual performance indicated strong student awareness of the difficulty of the interpretive and presentational tasks and accounted for their final ratings. An analysis of the relationship between the possible pairs of scores across the three tasks of the IPA showed no relationship between the interpretive task scores and the other two task scores, whereas a significant relationship between the interpersonal and presentational task scores was found. The statistical difference is supported by the students' low performance on the interpretive task and the challenges (as noted by the teacher) in providing

students with authentic interpretive tasks, an issue we will revisit in the Discussion section. Finally, the results of an analysis of student performance on the three tasks of the IPA by language group suggest that the more languages students spoke, the better they performed overall on this novice-level Spanish IPA. The next section will present students' post-IPA narrative comments that reveal student opinions of the IPA, support the finding that students are aware of their difficulties, and show a relationship between students' orientation to language learning and their performance.

Student Responses to Post-IPA Questionnaire

Following the IPA, students were asked to complete a five-item questionnaire that asked them to describe (1) what best helped them prepare for the IPA; (2) what helped them to be successful across all three tasks; (3) the difficulties they experienced during the assessment; (4) their level of enjoyment when participating in the IPA; and (5) any suggestions for improving IPA procedures. Thirty-seven students responded to the questionnaire and their responses were grouped by theme for each question. Interestingly, student responses were highly consistent and indicated an awareness of the nature of the IPA, the need for preparation, and the academic content-focus of the assessment in contrast to a focus on primarily grammatical accuracy. We also tried to determine differences in responses between students who performed well (i.e., *meets* and *exceeds expectations*) on the IPA compared to those students who did not meet expected levels of performance. This analysis, however, did not consistently show differences.

Question one asked the students to describe which types of classroom activities helped them to prepare for the IPA. Two themes emerged in their responses: (1) instruction that aligned with the content and language demands of the IPA, and (2) review and practice. Of the 37 students

who responded to the question, 15 students commented that the teaching and learning that took place during the thematic unit, which preceded the IPA, helped them to prepare for the IPA. Twelve students stated that review and practice were also helpful. Student comments related to instruction were prefaced by the phrases "we learned" and "we were taught" followed by specific references to the academic content and vocabulary of the thematic unit necessary for discussing the topic of the Caribbean. Representative comments include "we learned about the Mar Caribe," "we learned words for animals, activities, transportation, and clothing," and "we learned motions to help us remember new words." As one student stated emphatically, "definitely the [thematic] unit on Mar Caribe."

Students also stated that "we did little tests and went over answers," "reviewed connector words for writing," and "reviewed what would happen during the IPA." One student who *met expectations* on all three tasks provided a response that went beyond classroom tasks and practice and reflected her awareness of the learning process itself. In her words, "we needed to pay *attention* because it involved learning Spanish." Conversely, the comments of the six students who *did not meet expectations* on any of the three IPA tasks were typically vague, truncated, and underspecified. For example, "we learned more Spanish," "we learned clothes, weather, *and much more*," and "practice video, found differences, paragraph." As the representative comments indicate, the majority of the students perceived the connections between the thematic unit, which preceded the IPA, and their feelings of preparedness.

Question two, on the topic of what the students found most helpful to their performance on the IPA, did not, for the most part, yield information that differed significantly from the first question. Students continued to mention the learning of key vocabulary, the use of gestures, and motions to assist memory of the new words, and practice sessions on the topic of the IPA. When

reviewing the comments of students who *met* or *exceeded expectations* on three tasks, an interesting divergence emerged from the typical “practice and learn responses” given to *question one* and *question two*. These students reported specific classroom activities that linked directly to the IPA tasks, such as having opportunities to talk and interact with one another in Spanish; reviewing the IPA rubric; being asked to sound out important key words and guess their meanings (“[it] was really fun”); discussing and explaining the theme, reasons, and concepts embedded in the unit on the Caribbean; and seeing and using key vocabulary words in sentences rather than in isolation. These comments indicate that some elementary school students are capable of perceiving the connection of their daily classroom activity with expected performances on the IPA and did not view the IPA as being radically different from what they typically do in class. If assessment and teaching are to be seamless, these students clearly saw the close relationship of how they learn Spanish to how their progress is monitored and assessed in the IPA, which may have accounted for their strong performance.

In *questions three* and *four*, students were asked to comment on their difficulties during the IPA and the aspects of the experience that they enjoyed. For these two questions, we analyzed only the responses of students who completed all three IPA tasks. Additionally, no major differences in the quantity or quality of responses were found when comparing students who received ratings of *meets* or *exceeds expectations* to those who *did not meet expectations* on at least two of the three tasks.

The theme of the most frequent response to the question of difficulty was gaps in language knowledge. Approximately half of the students ($n = 12$) commented on their lack of vocabulary knowledge during the IPA. As one student stated, “it was hard finding the exact words for the things I [was] trying to say in Spanish.” Other difficulties reported were evenly distributed among student responses. These comments

included the speed and length of the listening comprehension material ($n = 4$), difficulties in combining words into sentences and paragraphs ($n = 3$), nervousness ($n = 3$), working without assistance ($n = 2$), not knowing how to participate in the conversation ($n = 1$), and the paired speaking task ($n = 1$). Clearly, students experienced some difficulty on the IPA and could articulate the sources of their struggles, primarily lack of adequate vocabulary resources for the tasks, difficulties in listening comprehension, and sentence combining skills. What is surprising is that, although students perceived gaps in their language knowledge and were challenged by the tasks, they remained quite positive about the assessment and did not express frustration, as the responses to the next question will show.

When asked to describe what they enjoyed about the IPA, a positive picture emerged that is consistent with the theory of assessment on which the IPA has been developed. Student comments also indicate that the IPA is perceived in ways that differ from traditional testing that emphasizes student deficiencies rather than achievements. Seven of the students reported that the IPA helped them learn new vocabulary and content and made them aware of how much Spanish they actually knew. One student stated, “I knew some words and the ones I did not know, I now know.” Another remarked, “I learned that I knew more than I thought I did.” Along with the learning potential of the IPA, seven of the students’ comments described the experience as enjoyable using words such as “creative, interactive, and fun.” Other students enjoyed talking to their classmates during the interpersonal task, illustrating their written work, and being able to receive feedback on their performance during the assessment. Curiously, three of the 16 children who *did not meet expectations* on two or three of the IPA tasks stated that what they enjoyed the most was the challenge (“The IPA was sort of challenging and I learned more”).

Finally, when asked to comment on suggestions for improvement, some student

responses took the form of realizations of the type of instruction needed to become proficient in the various modes of communication. Their comments were direct and focused and included “more class time learning words that are not just objects and actions,” “reading and [stating] main ideas of sentences and paragraphs,” and “more time [in class] explaining, like, in between words, like conjunctions; basically how to say more detailed sentences.” Thus, one positive outcome of the IPA is that the students become aware of what they need to know and be able to do to use Spanish in purposeful and meaningful ways. Likewise, the teacher also becomes aware of areas in the curriculum that need to be strengthened and is potentially well positioned to address these language issues in class with students who clearly perceive the purpose of the instruction.

In summary, students understood the close alignment between the thematic unit and the IPA and the seamlessness of teaching, learning, and assessment. Following their IPA experience, students reflected on how they had learned and perceived the increased need for learning vocabulary beyond nouns and verbs, sentence- and paragraph-building skills, reading for main ideas, and having frequent opportunities to speak to each other. Students could also comment on the content focus of the assessment and the learning that occurred during the assessment itself. Students enjoyed the interactive and social nature of the assessment, which is not surprising given that most classroom assessments are noncooperative. What is most compelling about the students’ responses, however, is that, from the experience of participating in an IPA, new language goals and purposes for learning emerge. One outcome of the IPA, therefore, seems to be a growing metacognitive awareness of the students’ own process of foreign language learning. This awareness of the learning process is expressed by the students in their desire for vocabulary resources and textual cohesive devices, rules of participation in interactions, selective

listening skills, and ongoing self-assessment of what they know and do not know.

Discussion

Based on the findings of this research, the IPA proved to be an effective assessment of proficiency in the three modes of communication with young Spanish language learners. Additionally, the IPA also provided valuable insight into the strengths and weaknesses of the EFLL program at this school. Four salient issues are discussed in this section: (1) reasons for the results on the interpersonal task; (2) reasons for the results on the interpretive task; (3) the need for student-friendly rubrics; and (4) patterns of performance based on number of languages spoken.

Two reasons can be given for why students performed best on the interpersonal task of the IPA. First, because the itinerant Spanish teacher teaches approximately 200 students for only 15 minutes daily, the majority of class time is focused on oral communication, typical of many elementary school foreign language programs (Adair-Hauck, 1992; Knop & Sandroch, 1994; Met, 1994, 2000). Given the constraints of time, systematic instruction in interpretive and presentational communication through meaningful activities is not possible on a regular basis. Thus, the default mode that often takes precedence during short class periods is interpersonal communication. Second, several useful adaptations were made during the interpersonal task to support the spoken interaction of the young learners. Unlike typical IPAs, the assessment facilitator provided pictures for the interpersonal task and prescribed prompts to guide the production of the students. In order not to invalidate the IPA process, students completed a picture difference task during the unit of study preceding the assessment.

Our second point requiring discussion, and similar to the findings of Glisan et al. (2007), is that students performed least well on the interpretive task, which suggests a

need for closer examination of this communication task. Two possible explanations for this low performance are (1) a lack of exposure to authentic spoken and written texts; and (2) the presence of a large amount of unfamiliar vocabulary. First, it is possible that students were unaccustomed to authentic texts and therefore scored poorly on the interpretive task. Students' reactions that the characters in the video spoke too quickly may reveal their lack of opportunities to listen to native speakers using Spanish in nonsimplified ways and at a typical rate of speed. Previous research has shown that speech rate, accent, and opportunities for interaction are significant factors for successful listening comprehension (Carrier, 1999; Garcia & Asención, 2001; Gass & Varonis, 1994; Goh, 1999; Leeser, 2004; Pica, Doughty, & Young, 1986; Pica, Young, & Doughty, 1987; Polio & Gass, 1998; Rodrigo, 2004; Tyler, 2001) and that listening comprehension improves with increased exposure to authentic speech (Herron & Seay, 1991). However, making meaning of complex authentic listening passages requires opportunities for students to develop the necessary listening strategies (Bacon, 1992; Shrum & Glisan, 2010). In elementary language programs, such as the one described here, development of the necessary interpretive listening/viewing strategies is often neglected. In this context, the teacher stated that two factors had limited her opportunities to engage students in listening comprehension activities—limited class time and lack of easy access to technology. Although her school has ample technological resources, assembling these materials in the available 15 minutes required reducing the precious few minutes she has for instruction. As many involved in early language programs already know, the constraints of scheduling and time limitations create conditions that often preclude balanced attention to all aspects of language learning (Mitsui, Haxhi, Donato, & Tucker, 2009).

A second possible explanation for the low scores on the interpretive task is that this video used a larger percentage of unfa-

miliar vocabulary than the interpersonal or presentational tasks. Because the interpersonal and presentational tasks were student generated and the interpretive task was produced for Spanish-speaking children and outside the students' locus of control, it is reasonable to assume that students would struggle more with this task. Additionally, because the language used by peers in the interpersonal task was based on previous classroom instruction, it was most likely more familiar to students than the language of the video. The question arises as to how students would have performed had the text of the interpretive task used a greater number of familiar vocabulary items. The resolution lies in finding a thematically related authentic text with a sufficient amount of vocabulary that reflects the unit of study. Future research might examine the appropriateness of using authentic video materials with novice-level learners for assessment, and what preparation they need to be successful on these tasks. Authentic video materials in which speech is delivered at a rapid rate could very well discourage young learners. This discussion does not suggest that authentic texts should not be used in the IPA with young learners. Rather, it suggests that the young learners' need frequent opportunities to interpret carefully selected authentic texts with appropriate supportive feedback from the teacher (e.g., dynamically assessed performance) (Davin, 2011; Poehner, 2008).⁴ In this way, we may improve young learners' interpretative abilities and be able to incorporate these materials into the IPA to ensure motivation and student success on this phase of the assessment.

Third, throughout the design and implementation of the IPA and the corresponding unit of study, it became apparent that some adaptations to the rubrics were necessary for young learners. The design of the IPA, its rubrics, and its feedback loop promote the communication of effective feedback to students (Adair-Hauck et al., 2006; Glisan et al., 2003). Tunstall and Gipps (1996) argue that descriptive feedback to young

learners must be highly descriptive, contextualized, and comprehensible to have an impact on student learning. The use of student-friendly rubrics during this unit allowed the students and the teacher to engage in the process of “constructing achievement” and “constructing the way forward” (Tunstall & Gipps, 1996, pp. 399–401) by increasing the comprehensibility of the feedback provided to the young learners during the IPA. We maintain that the language of the IPA rubrics was far too complex and specialized for young learners to understand. We revised, therefore, the language of the rubrics to communicate expectations to the young learners in clear and accessible ways. The contents of the rubrics were matched with the rubrics provided in the IPA manual and used to score the tasks (see Appendix C). These rubrics, which were inspired by the Consortium for Assessing Performance Standards (CAPS) rubrics,⁵ were also used throughout the instructional unit preceding the IPA and during the assessment process.

In the presentational task, students were asked to describe the weather, activities, transportation, and wildlife in the Caribbean. Each category of the rubric was aligned with the novice level rubric provided in the IPA manual, with language modifications for increased comprehensibility. For example, instead of labeling one category *Language Function* and using the *exceeds expectations* descriptor “*Creates with language, able to express own meaning a basic way,*” on the student-friendly rubric, the category was posed as the question “*Can I describe the Caribbean in Spanish?*” and the descriptor for *exceeds expectations* states “*I described the weather, the things to do, how to get around and types of wildlife in the Caribbean. I also added additional details and created with the language on my own.*” These questions and descriptors, modeled after the rubrics created by the Consortium for Assessing Performance Standards, allowed students to self-assess their work using language they could easily understand.

Fourth, as presented in Table 5, we found compelling evidence that students who spoke languages in addition to English in the home had an advantage on particular tasks of the IPA. The advantage was most evident in the interpersonal task. While this finding is nonconclusive due to the small sample size, it is an area that should be probed in future research. Previous research indicates that learning a third or fourth language is easier than learning a second language. Research has also indicated that when students are learning an additional foreign language, they bring to the learning task well-developed strategies and greater metalinguistic awareness (Sapienza, Donato, & Tucker, 2006).

Conclusion

In summary, the IPA was a successful classroom-based performance assessment of fourth and fifth graders’ Spanish proficiency in this program. We recommend this assessment in similar programs in which the goal is communicative competence and instruction is provided across all three modes of communication. In order for the IPA to be useful, this assessment should mirror a teacher’s instruction and should be part of a backward planning design process. Students who are not familiar with communication across the three modes of communication would most likely struggle with this assessment.

In the present study, students performed the best on the interpersonal task with 57% of students exceeding or meeting the standard. On the presentational task, 40% of students exceeded or met the standard. Students performed the worst on the interpretive task in which only 27% of students met the standard. In surveys, students indicated that they struggled with the interpretive task because they were unaccustomed to listening to the speech of native speakers of Spanish. In an interview with the teacher, she corroborated this statement by indicating that constraints of time and lack of easy access to the technology hindered instruction in

listening comprehension and interpretation in her classroom.

While the IPA is an effective method of assessing students' proficiency in a foreign language, it is also an effective way to assess a teacher's practice. This IPA offered a window into the foreign language program at this school and revealed its strengths and areas for improvement and additional attention. Findings presented above reveal that more attention in the curriculum is needed for developing stronger listening comprehension ability of native speaker communication and increased opportunities for meaningful practice of the students' written Spanish. These findings are not surprising considering that the teacher has only 15 minutes a day with each class and teaches approximately 200 students.

As discussed above, we found that in EFLL programs, such as the one in which we worked, the constraints of time and teacher resources create an imbalance in instruction on the modes of communication. One way of addressing this imbalance is to work toward the integration of foreign language into the content and goals of the school's academic curriculum and an increase in instructional time. Additional time for Spanish instruction could be justified on the grounds that the language program provides support for other academic areas and contributes to the students' intellectual development in the context of meaningful and purposeful language study of relevant and age-appropriate academic topics. In this way, the instructional inequalities between foreign language instruction and other academic areas would be addressed. We speculate that if the program were content-based and if the students had more time on task, the results of this study might be quite different.

Based on our findings, we have identified three implications for future research. First, a longitudinal study is needed to track language learners' performance on the IPA as they continue to study the language. We are currently preparing an IPA to be implemented with these same students in their middle school Spanish class. Second, we

found that the IPA had a powerful washback effect on this Spanish teacher's practice. More research is needed to document this change and to determine the nature of the washback. That is, does using the IPA produce a positive washback that truly serves as a guideline for standards-based instruction or conversely does it simply encourage "teaching to the test" (Poehner & Lantolf, 2010, p. 316)? Finally, this research points to the need for a detailed look at the feedback provided during the IPA and how this helps learners to improve language performance. Our hope is that the present study has created the impetus for further research of the IPA at this level of instruction and investigations of the consequences of using IPAs on teachers, the curriculum, and the status of foreign language programs in the school.

Acknowledgments

We wish to thank the Spanish teacher and students described in this study who welcomed us into their classroom and embraced our work. We thank Professor Eileen Glisan of Indiana University of Pennsylvania who encouraged our work on this study and provided numerous insights on this project. We also acknowledge Dr. Alan Juffs from the Department of Linguistics at the University of Pittsburgh for his assistance with the statistical analysis. Finally, we thank the three anonymous reviewers for their useful suggestions.

Notes

1. For the purpose of this article, we refer to these three communication tasks as the interpretive task, the interpersonal task, and the presentational task, respectively.
2. In future IPAs, we need to ascertain that students are indeed understanding the materials rather than simply identifying and guessing the meaning of words based on previous exposure. Little guidance is provided in the IPA manual on the use of listening texts for assessing the interpretive mode of communication. The use of

listening texts with young learners is in need of further investigation.

3. Typically, intermediate-level interpretive tasks also ask students to make concept inferences, such as inferring the author's intent.
4. Shrum and Glisan (2010) outline an interactive model of interpretive communication that could be explored in future research with early language learners.
5. The CAPS Project, which ran from 2003–2006, was funded by a Foreign Language Assistance Program grant and led to the development of a set of Thematically Oriented Assessments that use student-friendly rubrics. The assessments and rubrics can be found at <http://flenj.org/CAPS/?page=parent>

References

- Adair-Hauck, B. (1992). Foreign languages in the middle schools: A curricular challenge. *Pennsylvania Language Forum*, 64, 12–18.
- Adair-Hauck, B., Glisan, E. W., Koda, K., Swender, E. B., & Sandrock, P. (2006). The integrated performance assessment (IPA): Connecting assessment to instruction and learning. *Foreign Language Annals*, 39, 359–382.
- Alderson, J. C., & Wall, D. (1993). Does washback exist? *Applied Linguistics*, 14, 115–129.
- Bacon, S. M. (1992). The relationship between gender, comprehension, processing strategies, and cognitive and affective response in foreign language listening. *Modern Language Journal*, 76, 160–178.
- Bryk, A. S., & Gomez, L. M. (2008). Ruminations on reinventing an R&D capacity for educational improvement. In F. M. Hess (Ed.), *The future of educational entrepreneurship: Possibilities for school reform*. Cambridge, MA: Harvard Education Press.
- Carrier, K. (1999). The social environment of second language listening: Does status play a role in comprehension! *Modern Language Journal*, 83, 65–79.
- Davin, K. (2011). *Group dynamic assessment in the early foreign language learning classroom: Tracking movement through the zone of proximal development*. Unpublished doctoral dissertation, University of Pittsburgh, Pittsburgh, PA.
- Donato, R. (1998). Assessing the foreign language abilities of the early language learner. In M. Met (Ed.), *Critical issues in early second language learning* (pp. 169–197). Glenview, IL: Scott Foresman-Addison Wesley.
- Donato, R., & Tucker, G. R. (2010). *A tale of two schools: Developing sustainable early foreign language programs*. Clevedon, UK: Multilingual Matters.
- Garcia, P., & Asención, Y. (2001). Interlanguage development of Spanish learners: Comprehension, production and interaction. *The Canadian Modern Language Review*, 57, 377–401.
- Gass, S., & Varonis, E. (1994). Input, interaction and second language production. *Studies in Second Language Acquisition*, 16, 283–302.
- Glisan, E. W., Adair-Hauck, B., Koda, K., Sandrock, S. P., & Swender, E. (2003). *ACTFL integrated performance assessment*. Yonkers, NY: ACTFL.
- Glisan, E. W., Uribe, D., & Adair-Hauck, B. (2007). Research on integrated performance assessment at the post-secondary level: Student performance across the modes of communication. *The Canadian Modern Language Review*, 64, 39–68.
- Goh, C. (1999). How much do learners know about the factors that influence their listening comprehension? *Hong Kong Journal of Applied Linguistics*, 4, 17–41.
- Green, A. (2006). Watching for washback: Observing the influence of the International English Language Testing System Academic Writing Test in the classroom. *Language Assessment Quarterly*, 3, 333–368.
- Herron, C. A., & Seay, I. (1991). The effect of authentic oral texts on student listening comprehension in the foreign language classroom. *Foreign Language Annals*, 24, 487–495.
- Igarashi, K., Wudthayagorn, J., Donato, R., & Tucker, G. R. (2002). What does a novice look like? Describing the grammar and discourse of young learners of Japanese. *The Canadian Modern Language Review*, 58, 526–554.
- Knop, C. K., & Sandrock, P. (1994). Counterpoint to the exploratory approach. *Foreign Language Annals*, 27, 74–76.
- Leeser, M. (2004). The effects of topic familiarity, mode and pausing on second language learners' comprehension and focus on form. *Studies in Second Language Acquisition*, 26, 587–615.
- McDonald, J. P., Hatch, T., Kirby, E., Ames, N., Haynes, N. M., & Joyner, E. T. (1999). *School*

reform behind the scenes. New York: Teacher's College Press.

McDonald, J. P., Klein, E. J., & Riordan, M. (2009). *Going to scale with new school designs: Reinventing high school.* New York: Teacher's College Press.

Messick, S. (1996). Validity and washback in language testing. *Language Testing*, 13, 241–256.

Met, M. (1994). Current foreign language practices in middle schools. *Foreign Language Annals*, 27, 43–58.

Met, M. (2000). Instruction: Linking curriculum and assessment to the standards. In G. Guntermann (Ed.), *Teaching Spanish with the five C's* (pp. 49–69). Fort Worth, TX: Harcourt College.

Mitsui, A., Haxhi, J., Donato, R., & Tucker, G. R. (2009). Introducing writing activities into the Japanese-as-a-foreign-language (JFL) classroom: A teacher-researcher collaboration. *NECTFL Review*, 64, 5–17.

Muñoz, A. P., & Álvarez, M. E. (2010). Washback of an oral assessment system in the EFL classroom. *Language Testing*, 27, 33–49.

National Standards in Foreign Language Education Project. (2006). *National standards for foreign language learning in the 21st century* (2nd ed.). Lawrence, KS: Allen.

Pica, T., Doughty, C., & Young, R. (1986). Making input comprehensible: Do interactional modifications help? *ITL Review of Applied Linguistics*, 72, 1–25.

Pica, T., Young, R., & Doughty, C. (1987). The impact of interaction on comprehension. *TESOL Quarterly*, 21, 737–758.

Poehner, M. E. (2008). *Dynamic assessment: A Vygotskian approach to understanding and promoting second language development.* Berlin: Springer.

Poehner, M. E., & Lantolf, J. P. (2010). Vygotsky's teaching-assessment dialectic and L2

education: The case for dynamic assessment. *Mind, Culture, and Activity*, 17, 312–330.

Polio, C., & Gass, S. (1998). The role of interaction in native speaker comprehension of nonnative speaker speech. *Modern Language Journal*, 82, 308–319.

Rodrigo, V. (2004). Assessing the impact of narrow listening: Students' perceptions and performance. *Dimension*, 2004, 53–65.

Roschelle, J., Tatar, D. G., & Kaput, J. J. (2008). Getting to scale with innovations that restructure deeply how students learn mathematics. In E. A. Kelly & R. Lesh (Eds.), *Handbook of design research in mathematics, science and technology education: Innovations in science, technology, engineering, and mathematics learning and teaching* (pp. 369–395). Mahwah, NJ: Erlbaum.

Sapienza, B., Donato, R., & Tucker, G. R. (2006, October). Learning a second language: A district-wide foreign language program reaches the middle school. *The Language Educator*, 1, 24–27.

Shohamy, E. (2001). *The power of tests: A critical perspective on the uses of language tests.* London: Pearson Education.

Shrum, J. L., & Glisan, E. W. (2010). *The teacher's handbook: Contextualized language instruction* (4th ed.). Boston, MA: Heinle Cengage Learning.

Tunstall, P., & Gipps, C. (1996). Teacher feedback to young children in formative assessment: A typology. *British Educational Research Journal*, 22, 389–404.

Tyler, M. (2001). Resource consumption as a function of topic knowledge in nonnative and native comprehension. *Language Learning*, 51, 257–280.

Submitted December 18, 2010

Accepted July 20, 2011

APPENDIX A

Interpretive Task

I. Key Word Recognition

Find the equivalent of the following English words in the video. Write your answers in Spanish.

vacation _____

sea _____

beach _____

water _____

- hand _____
- it's okay _____
- picnic _____
- cold _____
- sand _____
- strange _____

II. **Important Ideas.** For each of the following, circle the letter of each detail that happens in the video. **They're worded as True/False.**

- A. A song is sung about the sea.
- B. Caillou hears the sound of the sea while in the car.
- C. The reason Caillou complains about the water.
- D. Caillou gets mustard on his sandwich and is not happy.
- E. The crab is described as "beautiful."
- F. After the first wave hits the sand castle, Caillou says to his parents, "Let's go!!"
- G. Caillou plans to go back to the beach.
- H. Caillou's mom says they will come back to the beach tomorrow.

III. **Main Idea(s):** From what you see in the video, provide the main ideas(s) of the article *Caillou se va a la playa* in English:

IV. **Words in Context**

- ¡Qué asco! _____
- y el otro y el otro y el otro _____
- ¡Mira! _____

APPENDIX B

Interpersonal Prompt

This is so exciting! Today you get to show off in front of our camera. So for the next few minutes, the three of us are just going to talk in Spanish.

You and your partner work for a huge magazine company that has offices in Pittsburgh and San Juan, Puerto Rico. The June issue of the magazine has to go to press today, but there is one big problem. The printing machines messed up and changed the images some, so you have to fix this. One of you lives and works in Pittsburgh, the other in Puerto Rico. You must use Spanish to figure out the differences between the two pictures. Your boss is listening in on the call and only understands Spanish. Try to be as descriptive as you can and use as much detail as possible. The more complete sentences you can use the better.

Before you begin your description

- Pick up the phone
- Greet your partner
- Make some small talk about how they are doing, the weather, etc.

Oooh, it's time for the magazine to go to press. Thank you both for helping with what could have been a huge problem!

APPENDIX C

Student-Friendly Rubric for the Interpretive Task

	Meets Expectations	Does Not Meet Expectations
Can I recognize vocabulary heard in the video?	I accurately identify most vocabulary words that I have already been taught.	I cannot accurately identify many vocabulary words that I have already been taught.
Can I understand the Caillou video?	My answers show that I understood the important idea I heard in the Caillou video.	My answers show that I understood little of what I heard.
Can I identify the main idea in the video?	I can identify the main idea.	I cannot identify the main idea.

Student-Friendly Rubric for the Interpersonal Task

	Exceeds Expectations	Meets Expectations	Does Not Meet Expectations
Can I talk to my partner and describe the differences between the pictures?	During the conversation, I express the differences in the picture and sometimes create my own meaning.	During the conversation, I mostly use memorized language to talk about the differences in the picture.	During the conversation, I use only memorized language and lists of words.
Can I speak complete sentences?	I speak mainly in complete sentences. Sometimes I connect sentences.	I speak using some complete sentences and some memorized language.	I speak using words, phrases, chunks of language, and lists.
How well do I communicate in Spanish?	I can ask and answer simple questions. I clarify by asking and answering questions.	I mainly respond to questions asked by my partner. I clarify by using different words.	I respond only to questions that we practice a lot. I repeat words or use English to clarify.
Am I understood when I speak in Spanish?	I am understood without difficulty.	I am understood with occasional difficulty.	I am understood with much difficulty.
How accurate am I?	I am mostly correct when speaking in simple sentences and make errors when creating.	I am mostly correct with memorized language.	I am correct only at the word level.

Student-Friendly Rubric for Presentational Task

	Exceeds Expectations	Meets Expectations	Does Not Meet Expectations
Can I describe my animal in Spanish?	I described the weather, the things to do, how to get around and types of wildlife in the Caribbean. I also added additional details and created with the language on my own.	I described one or two of the items required. I used mainly memorized language.	I used only memorized language and did not complete the task.
Can I use complete sentences?	I wrote in complete sentences.	I had two complete sentences and some lists.	I used only lists.
How detailed is my vocabulary?	Vocabulary is sufficient to provide information and limited explanation.	Vocabulary conveys basic information.	Vocabulary is limited and/or repetitive.
Can the audience understand me?	The reader understands me without difficulty.	The reader has some difficulty understanding me.	The reader does not clearly understand me.
How accurate am I?	I am mostly correct when producing simple sentences and make errors when creating with the language.	I am mostly correct with memorized language.	I am correct only at the word level.