

*in submission: please do not quote  
without the author's permission.*

**RUNNING HEAD: Assessing Narratives from Bilingual Children**

**Assessing Narratives from Bilingual Children:**

**Borrowing a System from Gymnastics Judging**

**Barbara Zurer Pearson**

**University of Miami Bilingualism Study Group**

**c/o Department of English**

**Box 248145**

**Coral Gables, FL 33124**

**USA**

**bpearson@umiami.ir.miami.edu**

### Abstract

A two-part scoring system for assessing children's narrative was derived from a gymnastics judging model. The "frog story," from the cross-linguistic research of Berman & Slobin (1994), was used as a compulsory routine, and the principles for extending the metric to other stories (or "optionals") are discussed. The stories are evaluated on the one hand for discursive or narrative elements--the Story Score--and then separately for the more narrowly linguistic elements, or Language Score--verb phrase elaboration, embeddings, morphosyntax, and vocabulary. The two measures, each with a maximum of 48 and a midpoint of 24, representing the average performance of the idealized 2nd grader, were used to discriminate among stories narrated by 2nd and 5th graders, monolinguals in English and bilinguals in English and Spanish. The reliability of the scoring system was tested with a test-retest procedure which showed the measure to be relatively stable across different renditions of the story told by the same child one to two weeks apart, while capable of finding differences between groups, here between children in the two grades.

Assessing Narratives from Bilingual Children:  
Borrowing a System from Gymnastics Judging

Everyone tells stories. For children, telling stories is an important step forward in both thinking and talking. A story creates a frame of reference beyond the here and now. To know what comes next--and what the listener must hear next--requires considerable power of abstraction on the part of the child. Indeed, giving children the opportunity to tell stories is a powerful means to promote their intellectual growth (Hedberg & Westby, 1993). In terms of language skills, much of the grammatical development of the primary school years takes place beyond the level of the sentence, in the way speakers learn to connect and interweave larger stretches of text (Halliday & Hasan, 1976; Karmiloff-Smith, 1987). Thus, the stories children tell reveal much about the general cognitive and linguistic skills essential for literacy and academic success.

Narratives demonstrate children's abilities on many levels at once. But in order to gain access to what stories can tell about development, one must first learn what to listen for in them. Unlike short-answer tests, there is no "answer key." To use a story as a measure of a child's progress, we must train our observations to include both analytic and holistic judgments focused 1) linguistically within and across sentences and 2) conceptually at the larger units of discourse created by the child. Such an approach holds special relevance for evaluating the narratives of bilingual children for whom we hypothesize a greater dissociation of skill level in these two domains, compared to monolingual age-mates.

One key to assessing extended language samples is to find ways to make subjective reactions objective. Subjective reactions, especially by trained experts, are not invalid by

virtue of being subjective. Indeed, in many situations there is no alternative. For a complicated performance like a story or an essay, one can identify in advance elements of an adequate or average performance for a given group and then outline the ways in which a particular performance can surpass or fall short of that standard. The standard is not arrived at from magic or intuition, but rather from observing similar performances by similar groups and analyzing what elements are generally present in "good," "medium," and "poor" renditions of the task, with those anchor-points determined by consensus among experienced evaluators. As in judging gymnastics or figure skating competitions, the consistency of the judging is improved when there is a set sequence, like compulsory routines, but the principles learned with compulsories can be applied to "optionals," that is, more free-form performances.

For a research project on bilingual children's literacy, my colleagues and I have the task of evaluating 360 stories collected from 20 groups of 2nd and 5th graders. In establishing our method, we drew on the model of gymnastics judging, (which I did semi-professionally for 10 years, FIG, 1989). We chose Mayer's (1969) "frog story," Frog. Where Are You? for our "compulsory routine." Ten years of intensive international scrutiny, (summarized in Berman & Slobin, 1994), have given us a clear idea of what constitutes a mature narrative performance based on this wordless picture-book. The literature on narrative analysis has pointed out areas that are the locus of development and has identified a number of elements that can act as "indices" or markers of that development.

While individual researchers have isolated elements to compare across stories-- instances of ambiguous reference (Hemphill, Picardi, & Tager-Flusberg, 1991), or pronominal chains (Hedberg & Westby, 1993), for example--and teachers have a generalized rubric based on story grammar principles (Stein & Glenn, 1979), there is not (to our knowledge) a standard

evaluation procedure that allows separate measurement of the elements of the children's conceptual performance, on the one hand, and their more narrowly-defined linguistic performance on the other. The two tables that follow summarize the approach we took. Table 1 shows how to get a Story Score for a child's narrative; Table 2 outlines the elements of a Language Score.

For both scores, we borrowed another useful concept from gymnastics judging: the partitioning of the score into "Composition" and "Execution." In composition, elements and transitions are counted as having a specified value. For a given routine, there must be so many acrobatic and dance elements of medium and high value, at least three elements directly connected, and so forth. By adding the values of all the elements, one arrives at the highest possible score for the particular routine performed--IF all the elements were executed perfectly. Then, one subtracts execution deductions according to the general table--take off so much for a fall, so much for bending the arms in an element or throughout--to arrive at the final score.

What we establish here in our scoring system, then, is the value of the elements and the table of deductions for the frog story. The tables do not list every possible element. Rather we have tried to isolate the most salient features of mature narratives based on this stimulus, Mayer's picture book. The goal was to be specific, but in the interest of practicality, economy was valued over exhaustiveness.

### The Scoring System

---

Put Tables 1 and 2 about here

---

### The Story Score

As indicated by the column titles of Table 1, the child must provide story elements and sequence them from a unified perspective with enough engagement for the listener to want to follow. The child must introduce characters, relate key events (while omitting others), and finally resolve the search for the frog. The "modal story" by the idealized, average 2nd grader tells the story as a chronicle of events: this happened and then this happened and then this happened. The child who does this clearly and simply gets the mid-point Story Score, 6 on each of the 4 dimensions, or 24. It is important to note that neither the standard nor the indices chosen are fixed. They were established to distinguish degrees of distance from the average 7- or 8-year-old's story that we would be likely to find among the stories of children between 7 and 11. If one were trying to distinguish between adult productions, greater elaboration among higher level markers would be necessary; to find more gradations among the less mature stories, one would identify more markers at that level.

For our measuring scale, the child who manages a clear, straight-line narration can still lose "execution" points for failing to introduce one or more of the characters, or for losing track of clear reference when using pronouns to talk about the characters (or not reserving the pronoun exclusively for the main character, as described by Karmiloff-Smith, 1981). More points can be lost by getting side-tracked into details that do not advance the story or by lapsing into mere picture-description. On the other hand, the child who articulates the characters' goals, or gives information about their feelings or states of mind is awarded credit

for extra elements (Labov & Waletzky, 1967; Peterson & McCabe, 1983). Double credit is given for remarking on the characters' misperceptions, that is, conveying a perspective on a story event that is wider than what the character sees. Points are also awarded if the child manages to "make a story"--use appropriate exclamations or direct speech to engage the listener, or use summary statements to tie events together and bring in more than what is seen on a single page.

The columns can be summed or, depending on the purpose, used as individual scores.

### The Language Score

The Language Score, from Table 2, is composed of three subscores: Complex Syntax, Lexicon, and Morphosyntactic Accuracy (or Errors).

Complex Syntax subscore. We first try to establish how rich the child's language is and then focus on specific devices the child uses to convey more nuanced information. Following Silva-Corvalán (1991), we use the verb phrase as an index of syntactic richness. We have chosen different markers for English and Spanish, but the principles are similar. In Spanish, the child gets credit for using the past perfect tense or the subjunctive mood, as both forms follow constraints established outside the clause they occur in. In English, we looked for verb phrase complexity in the use of modals ("would call," "couldn't find") or aspectual marking ("started running," "kept on calling"). The other category of complex verb usage in both English and Spanish involved noun clauses ("He forgot where he was") and adjective clauses with relative pronouns, ("the deer that was carrying him").

The second aspect of the Complex Syntax subscore, labeled "Between Clauses," looks at the child's use of conjunctions and time adverbs. In particular, we tracked how the child expressed causes, intentions, and simultaneous actions. We also gave credit for adverbs that

point to two spots on a time line. For example, using the word "still" means that something started at time 1 and it was continuing at time 2. We considered this a more difficult time reference than "in the morning" or "then." For this submeasure, one short-cut we considered was just to count conjunctions, like "while" for simultaneity of two on-going actions. However, the evaluation needs to be more nuanced: "when" sometimes, but not generally, also means "while," and the concept can also be neatly conveyed with participles and no conjunctions at all, as in "he climbed up on the rock, calling out to his frog." (Of all the elements of our scores, these two columns require the most specialized language knowledge on the part of the rater.)

Lexicon subscore. The third column on Table 2 summarizes the child's use of the specific vocabulary suggested by the pictures, but we also tried to give credit for other well-chosen words: "sting" for the bees versus "bite" for the groundhog, "encaramar" or "treparse" rather than "subir." (They all mean "to climb up.") We also gave a penalty in this column when the child referred to a character or object but couldn't find the appropriate word: "animal muy grande" [very large animal] for "deer," "pájaro" [bird] for "owl" or "the place where they kept the frog" for "jar" or "bowl."

Morphosyntactic Error score. The final column tallies errors in word forms: "en la agua" for "el agua," "runned" for "ran," or agreement errors (but allowing for the dropping of final "s" in the Caribbean varieties of Spanish). In this category, judgment was sometimes needed to distinguish if the child was truly changing tenses or was just failing to use the correct form, "and then he fall" (for "fell"). As on the other scales, we used 12 as a target score, but children could get less than zero in these two columns and more than 12 in lexicon.

As with the Story Score, we used the modal 2nd grader as our idealized "midpoint" of



the scale. To get a "base score" of 10 in Complex Syntax, the child just tells what she sees: "this, then this, and then this," mostly in one or two tenses, generally the simple past and the past continuous (or imperfect in Spanish).

Tables 1 and 2 tell what we recommend looking for in children's narratives. One factor we chose not to score was length. Granting that longer versions of the story tended to score higher on story score because they usually gave more information, it is important to consider that some of the most mature stories are among the shortest. One of the ways children pick up points in story organization is by leaving irrelevant material out. In addition, while length showed reasonable correlation to these story measures, it was much less associated with various proficiencies as measured through standardized tests than were the story measures (Pearson & Umbel, 1995).

### Sample Story

A sample of a child's story in English and Spanish and comments from our score sheets are included here for illustration. The child is a 2nd-grade girl in a two-way bilingual program who speaks only Spanish at home. Her Spanish story is quite developed (a 30 on our scale), and the Spanish language score (25.5) and Complex Syntax subscore (18) are relatively high, as well. By contrast, her English language appears very restricted. Despite the very low level of language (a 4.5 Language Score), the English Story Score manages to be quite average (23).

Lines 1-6, Spanish:

(Note: "-" means pause; false starts are in curly brackets; rough phonetic information is in square brackets.)

Había una vez un niño

que tenía - un perrito y una - ranita  
 y entonces [entonce] - él estaba durmiendo un día  
 entonces [entonce] la ranita se le había salido  
 {de una - de u: - } de una cajita que él la tenía.  
 Cuando él se despertó - {estaba - mu:} estaba triste.

Versus lines 1-6, English:

One time a dog - was looking in a jar.  
 And - the kid sleep with the dog -  
 {and -and then the frog - um a- I don't know} the frog just go out the jar  
 and then when is {um - night - uh } day  
 the kid look in the jar  
 and then he do not see the frog.

### Rater's Comments

While there are grammatical errors in the first passage and some obvious lexical gaps, the tenor is comfortable and the use of the past perfect marks it as unusually good. The child introduces the characters with indefinite pronouns and, having used the subordinate clause to introduce the dog and the frog, she makes it clear that the boy is the subject in lines 3 and 6. She loses points in the Spanish for not using a mental verb to indicate the boy's discovery of the frog's escape, but she gains them back by saying that he's sad.

In the English passage, the child gets story points for mentioning specifically that the kid looked in the jar and didn't see the frog, but she loses them for her haphazard introduction of the boy and the frog, using the definite pronoun "the" for their first mention. The problems with verb forms in the second passage show the child is clearly struggling, and she uses no complex verb phrases or special vocabulary. Later in the story, she gets points for one noun clause construction "to see if the frog was there." (Interestingly, "a ver si" was one of the elements that was also counted in her Spanish language score and therefore might have transferred for her.)

Reliability of the measure.

The validity of the measure needs to be assessed externally, as we have done, by reference to constructs of development in the targeted domains during this age range. However, it is clear that even a measure deemed to be valid will have no utility if it is measuring aspects of the children's performance that are not representative of their performance more generally. If we are looking at the use of modal auxiliaries in the story, for example, as a marker of verb phrase development, we need to be assured that the child who uses modal auxiliaries in one day's performance of the story will likely use them in another rendition, and conversely that if someone fails to use modals, that individual would also fail to use them in another similar circumstance close in time. And the same assurance is necessary for the other points measured. Therefore, to determine whether the aspects of the performance that we chose as markers of maturity were in fact reliably present in the stories, we devised the following procedure for test-retest reliability.

Twenty-four monolingual English-speaking children in our control groups, twelve 2nd-graders and twelve 5th, were asked to tell the story a second time during another test session, 1 to 2 weeks after the first. The two sets of stories were put in a single batch with no marks to indicate the child's grade or whether it was a first or second telling. All stories were coded "blind" by two raters (as we always do). When the scores from the 24 time-1 tellings were compared to the 24 time-2 tellings with a paired-samples t-test, the  $t$  value was less than 1 (n.s.) for both story and language scores. Mean differences on the measures were very small, always less than .6 between tellings (cf. standard deviations were around 6 points); and the correlations between times of telling were .77 for the story scores, .76 for the language scores, and .8 for the two summed together.

While the means and correlations show stability, this is not to say that individual scores did not change from one time to the next. Indeed, while the mean for the group did not change, the range of the absolute value of the time1-time2 discrepancy was 0 to 7 for the language score and 0 to 13 for the story scores, with means of the absolute differences of 4 and 3, respectively. Sixteen of 24 children remained within 6 points in total score, but equal numbers of stories for the others got better or got worse. This pattern can be seen clearly by comparing the length of the stories: 12 children stayed within 5 clauses of the original story length (average story length equaled 48); 5 of the longer stories got an average of 20 clauses shorter, and 7 stories got longer by about 10 clauses. Another factor of interest was the pattern of relation between the two scores, whether the child's language score was higher than her or his story score (as it generally is for the monolinguals): only 2 of the 24 children changed in this respect. In both cases, the difference was not very large, but it straddled zero. Finally, when we checked three "big-ticket" items in the story scores--the mention of a mental verb for the discovery of the missing frog (+2), mention of the boy's misperception of the deer's antlers as sticks, or similar misperception (+4), and the inclusion of more than one comment on the "internal" state of the characters--we found that between 83 and 92% of the children were consistent across tellings. Of the 3, 4, and 2 children, for each element respectively, who were not consistent, there was, as with the story length, an equal tendency to embellish the story more or elaborate it less in the second telling.

It would appear, then, that the measure is relatively stable from a test-retest point of view, but there is no fixed criterion to evaluate whether a magnitude of the test-retest coefficient of around .8 will be adequate to allow discrimination of true differences, ones we are interested in ascertaining between groups. For these 24 children, the mean differences

on the two story and language scores, grouped by grade (not for the whole group as with the means reported above) are shown in Table 3. One can see that grade effect for the story scores was about 4 times the difference observed between tellings, a pattern repeated in the language score as well. In order to test the relative size of potential group differences vis-a-vis the size of the probable error associated with the reliability of the scoring, we did an analysis of variance of story scores to compare the within-subjects effect of "time" to a between-subjects effect, here grade (the groups of 12 each were balanced with respect to socioeconomic status, the other independent variable in the larger study's design). The  $F$ -statistic for time and for the time by grade interaction were both less than 0.6, in comparison to the  $F$  of 2.61 for grade,  $p = .1$ . While this is higher than most researchers accept for a  $p$  value, it seems likely that with larger groups, a significant difference using this measure could be found. Indeed, in preliminary analyses involving the different language groups (monolingual versus bilingual, only Spanish in the home versus Spanish and English in the home), larger differences on the dependent variables were observed, but we cannot test them in this manner since we do not have the two stories in the same language for them.

### Discussion

What, then, are we learning with this framework for analyzing stories? Thus far, 199 stories have been examined on a pilot basis. The comments that follow here are based on our preliminary analyses.

The separation of the assessment into a story score and a language score appears especially important for bilingual students. Beyond its usefulness for the objectives of our project, it is important for teachers of bilingual students and those learning English as a new language to focus on the two elements of a narrative as separate aspects within the

developmental process. This separation enables teachers to distinguish the child's language resources in a given language from the information she can convey in that language. Too often surface features of language--wrong word endings, faulty subject-verb agreement, even limited vocabulary--mask the child's competence in the narrative/conceptual domain. This is especially important for students in the process of learning a new language. In our preliminary results, the monolingual children's language scores were almost always higher than their story scores, while the opposite was often true for the bilinguals. Their story scores were higher than their language scores one-third of the time. It would be crucial, therefore, to credit the bilingual child's conceptual ability by removing any "language effect" from it as much as possible.

Our results also underline the need to assess bilinguals in both of their languages. Not surprisingly, for children born in the US and educated in the US school system as our subjects have been, two-thirds of these bilingual 2nd and 5th graders had story scores as high as or higher in English than in Spanish. That still leaves one-third of the bilinguals who showed their conceptual capabilities best when tested in Spanish.

Another advantage of this multi-faceted system of analysis is the perspective it puts on the role of vocabulary in a complex language performance. Very often vocabulary will be correlated with intelligence quotients and so vocabulary measures, which are relatively easy to obtain, are used to "stand in" for the measures they have been correlated with. It is also quite common to use vocabulary measures as indicators of general language competence. What we have found in preliminary analyses is that vocabulary is not a good index of narrative skills. In the first 111 stories, there was only a moderate correlation between the lexicon measure obtained from the sample and the Story Score (Pearson & Umbel, 1995). That is, a

child can control reference and first mentions well and include important story elements and information about the characters' internal states without using the specific vocabulary of the story ("jar," "deer," "log" etc.) By contrast the Complex Syntax subscore had a higher correlation with the Story Score in both English and Spanish.

In addition, it appears from our preliminary analyses that the Complex Syntax and Story Scores are more similar across languages than the vocabulary and morphology scores (Pearson & Umbel, 1995). That is, children whose abilities are very unequal in the two languages seem able to carry over the elements of complex syntax and perspective taking from their stronger language to their weaker languages, whereas there is little transfer of vocabulary or inflections.

### Conclusion

The system described here for analyzing and assessing stories is still under development. But it has already shown itself to be reliable across different raters and multiple tellings of the story. As in gymnastics judging, the application of an analytical system for assigning value to a complex performance has yet another advantage. By using the table of values consistently, little by little we learn to make the same judgments more automatically. When experienced gymnastics judges evaluate a routine holistically, their scores generally agree with scores derived by a novice's careful application of the judging protocol. Just as the children grow through their opportunities to tell stories, we train our assessment abilities by listening to their stories within a careful and specific framework such as this one.

## References

- Berman, R. A. and Slobin, D. I. 1994. Relating events in narrative: A crosslinguistic developmental study. Hillsdale, NJ: Lawrence Erlbaum.
- FIG, Women's Technical Committee. 1989. Code of points: Artistic gymnastic for women. Switzerland: International Federation of Gymnastics (FIG).
- Halliday, M.A.K. and Hasan, R. 1976. Cohesion in English. London: Longman.
- Hedberg, N. L. and Westby, C. E. 1993. Analyzing storytelling skills: Theory to practice. Tucson, AZ: Communication Skill Builders, Inc.
- Hemphill, L., Picardi, N. and Tager-Flusberg, H. 1991. Narrative as an index of communicative competence in mildly mentally retarded children. Applied Psycholinguistics, 12, 263-279.
- Karmiloff-Smith, A. 1981. The grammatical marking of thematic structure in the development of language production. In Deutsch, W., editor, The child's construction of language (pp. 123-147). London: Academic Press.
- Karmiloff-Smith, A. 1987. Comparing language and cognition. In Hickman, M., editor, Social and functional approaches to language and thought (pp. 185-202). London: Academic Press.
- Labov, W. and Waletzky, J. 1967. Narrative analysis: Oral versions of personal experience. In Helm, J., editor, Essays on the verbal and visual arts, (pp. 12-44). Seattle: University of Washington Press.
- Mayer, M. 1969. Frog, Where Are You? New York: Dial Books.
- Pearson, B. Z. and Umbel, V. M. 1995, November. The relationship between narrative and linguistic skills in English and Spanish among Hispanic school children in Miami.



Paper presented at the 20th Annual Boston University Conference on Language Development, Boston, MA.

Peterson, C. and McCabe, A. 1983. Developmental psycholinguistics: Three ways of looking at a child's narrative. New York: Plenum Press.

Silva-Corvalán, C. 1991. Spanish language attrition in a contact situation with English. In Seliger, H. and Vago, R. M., editors, First language attrition (pp. 151-171). Cambridge: Cambridge University Press.

Stein, N. L. and Glenn, C. G. 1979. An analysis of story comprehension in elementary school children. In Freedle, R. O., editor, New directions in discourse processing (pp. 53-120). Norwood, NJ: Ablex.

# FROG STORY CODING SHEET

Coded Subject No. \_\_\_\_\_ Lang/Order \_\_\_\_\_ Coders \_\_\_\_\_ No. of Clauses \_\_\_\_\_

Table 1. STORY SCORE

Each category is worth 12 points. In each column, choose the "highest" behavior that describes the story being coded, deducting for behaviors NOT shown, according to the scale in that column.

<u>Frog Story Elements</u>	<u>Sequence</u>	<u>Perspective</u>	<u>Engagement</u>
Specific Mentions:	0 Picture Description	-2 Uses "here" "there" "now" to refer her own ref. frame (not story's)	
+2 Discovers missing frog (uses mental verb) (looks only, +1)	1	Poor "first mention" (uses "the" or pronoun right away)	-2 vague confused (in parts)
+2 Initiates search	2	-1 Main characters	disfluencies
+1 Finds frog	3 Isolated Events	-1 Other characters	-2 grave
+1 Takes frog (home)	4	Lapses in reference (reader asks "who?")	
	5 Sequential Events (some, not all)	-1 1 defective ref.	
	-1 for picture descr.	-2 2 "	
	-1 for irrelevant details	-4 5 or more	
(6) average	6 FACTUAL STORY	6 NEUTRAL OBSERVER	6 MATTER OF FACT
+1 Articulates goal (each, up to 2)	7 Elaborated episodes	(7 Ascribes Intention) or	attempts to be lively or engaging
+1 Articulates lack of success	9 Hierarchical struct. (begin/middle/end) (highlighting)	8 Gives Internal State info (affective) 1 or 2	+1 ex. use of "refrain"
+4 notes character's misperception (branch/antler or other)	12 Retrospective or prospective Summary +2 for summary stat. +3 for 2 or more	9 3 or more	+1 appropriate exclamations
		12 Mentions "ironic" perspective*	+1 extensive direct speech
			12 figures of speech
			_____ /12
			Total: _____ /48

\_\_\_\_\_ /12

\_\_\_\_\_ /12

Table 2. LANGUAGE SCORE

Complex Syntax		<u>Morphosyntactic Accuracy</u>	
<u>Verb Phrases</u>	<u>Between Clauses</u>	<u>Lexicon</u>	<u>Errors</u>
Mixed Time Sequence -1 for one (grave) -2 for 2 or more		Circle if present (by name)  frog  jar/bottle  bees (wasps)  beehive  rodent-type  owl  trunk (Eng)  deer  antlers  pond/lake  log  cliff  other "specific" vocabab:  other vocab mistakes:	-1 each "type" (not "token") articles  prounouns  prepositions  verb forms  conjunctions  word order  other:  Total (subtract from 12) _____/12
6 NATURAL FLOW OF EVENTS	4 NO CONJUNCTIONS or MOSTLY "AND THEN"		
Eng: Modals or aspectual auxiliaries +1 each (to 3)	Causal connections (or intentions) +1 1 or 2 +2 3 or more		
Eng or Spa: Complex Ss Relative Clause Noun Clause +1 (for 2, up to 3)	Simultaneous actions ("while"/"when"-simul) +2 1 +3 2 or more		
Spa: Perfect Tenses			
+2 for one +3 for 2 or more	Retrospective reference "still"/ "already" "todavia"		
Spa: Subjunctive  +2 for one +3 for 2 or more	+1 each, up to 3		
	_____/12		
Complex Language		Total: 0-12	Total (all 4 language subscores) :
Total: _____/24			_____/48

Table 3. Story and Language Scores by Grade and Time of Telling,  
Means and standard deviations.

	N	(Time 1)	(Time 2)
<hr/>			
Story Score			
Grade 2	12	27.9 (6.6)	26.8 (5.6)
5	12	30.5 (5.4)	31.5 (6.2)
<hr/>			
Language Score			
Grade 2		34.1 (7.8)	34.0 (5.2)
5		37.2 (5.5)	38.6 (6.0)