

APPENDIX 11

DYSLEXIA STUDY REPORT



REPORT

Accommodated Reading Assessments for Students with Dyslexia

Submitted to the Texas Education Agency

August 31, 2004

Revised – October 21, 2004

Jack M. Fletcher

Amy Boudousquie

Sharon Kalinowski

Center for Academic and Reading Skills

University of Texas Health Science Center – Houston

Kim Copeland

David J. Francis

Texas Institute for Measurement, Evaluation, and Statistics

University of Houston

Executive Summary

In response to legislation from the 78th Texas Legislature (House Bill 1998), the Center for Academic and Reading Skills (CARS) and the Texas Institute for Measurement, Evaluation, and Statistics (TIMES) conducted a study on behalf of the Texas Education Agency addressing the effects of accommodations on the Grade 3 Texas Assessment of Knowledge and Skills (TAKS) reading assessment. Students in Grade 3 who were identified with dyslexia and a comparison group of average readers in six school districts in the Houston area were randomly assigned to take accommodated and standard administrations of the same version of the TAKS reading assessment. The accommodated administration allowed the student to take the test over two sessions in a small group. In addition, all proper nouns and the stems used to assess comprehension of the passages were read to the student. None of these accommodations invalidated the TAKS as an assessment of reading. The results showed a significant increase in the performance of students with dyslexia who received the accommodated administration compared to students with dyslexia who were administered the TAKS under standard conditions. This improvement led to a seven-fold increase in the odds of passing for students with dyslexia who received accommodations (pass rate of 41%) compared to students with dyslexia who did not receive accommodations (pass rate of 9%). Accommodations had no effects on the level of performance or the pass rate for the comparison students, who were average readers (pass rate about 80% in each group), thus demonstrating that the modifications are not effects on the test that would benefit any student. The impact of the accommodations was not related to level of vocabulary development or to severity of reading disability. The effects of these accommodations may be apparent for older students with dyslexia as well as those with other disabilities and reading problems involving word recognition and should be further evaluated.

Accommodated Reading Assessments for Students with Dyslexia

Introduction

We report a study of the influence of a package of three accommodations for students with dyslexia on their performance on the Texas Assessment of Knowledge and Skills (TAKS) that represents a collaboration of the Texas Education Agency, the Region X Education Service Center, the Center for Academic and Reading Skills (CARS), the Texas Institute for Measurement, Evaluation, and Statistics (TIMES), and the Vaughn Gross Center for Reading and Language Arts. At the request of TEA, CARS and TIMES conducted a study of the impact of an accommodation package on student performance on the Grade 3 TAKS reading assessment. This study was conducted in response to legislation from the 78th Texas Legislature (House Bill 1998) mandating alternative assessments for students with dyslexia. Given that the available literature does not show robust effects of accommodations for students with disabilities and that allowing a student with dyslexia to have an accommodation in which the passages are read to them invalidates the test, this study was undertaken in order to assess whether accommodations that do not invalidate the TAKS reading test will lead to improvements in the performance of students with dyslexia. The researchers administered the TAKS reading assessment in standard and accommodated conditions to students in Grade 3 with dyslexia and to a comparison group of average readers in order to a) establish the impact of accommodations on the TAKS performance of students with dyslexia and b) determine whether any changes in performance attributed to the accommodations are specific to individuals with dyslexia.

Design

The study design was proposed by a working group convened at TEA (see Appendix I) and reviewed by Pearson Educational Measurement (Jon Twing, Michael Yoes) and Ed Miller at TEA. The essence of the design was a comparison of Grade 3 students with dyslexia and a

comparison group of average readers who received either an accommodated administration of the TAKS or completed a standard administration of the TAKS. In addition, we evaluated whether performance varied with overall verbal ability and level of decoding ability. The accommodated and standard administrations of the TAKS were given to Grade 3 students who were average readers in order to establish that the modifications represented a true accommodation as opposed to modifications that would benefit all children. This comparison was important to ensure that the accommodations do not affect the validity of the TAKS. If the students with disabilities improved, but not the average readers, this finding would indicate that the accommodations did not alter the validity of the test.

Within each group, assignment to testing condition (i.e., accommodated or not accommodated) was random. The overall design was a two groups (dyslexic versus average readers) by two conditions (accommodated versus standard administrations) design that permitted testing of three specific questions: 1) Do students with dyslexia perform better on the assessment with an accommodated administration than on the standard assessment with a standard administration? This is a comparison of the two groups with dyslexia. 2) Is the advantage conferred by the accommodated administration greater for children with dyslexia than for children without dyslexia? This is the test of the interaction of group and type of TAKS administration. 3) Does the advantage conferred by the accommodated revision for students with dyslexia vary by virtue of the student's vocabulary level and word decoding ability? This was tested by including vocabulary and word decoding as covariates in the model and evaluated as interactions of the covariate with accommodation status.

Nature of the Accommodations

A working group convened by the Texas Education Agency assessed the literature on accommodations and worked with the Dyslexia Statewide Focus Group to develop the

accommodations and design the study. Appendix I includes the roster of the TEA- convened working group. On April 25, 2003, a meeting was held with the Dyslexia Statewide Focus Group to review the study, with additional recommendations incorporated into the study. The critical consideration for students with dyslexia is to minimize the impact of their difficulties with word decoding. It was determined that three accommodations might potentially reduce the impact of decoding problems on an accommodated administration without invalidating the TAKS as an assessment of reading: 1) extending the testing session into two blocks to complete the TAKS; 2) reading of proper nouns to students; and 3) reading of comprehension stems to students. The rationale behind these accommodations is as follows: 1) When decoding words is difficult, students with reading difficulties commonly experience fatigue and diminished interest during the task. Consequently, *extending the testing session into two blocks* may be beneficial to counteract effects of fatigue and loss of interest due to word decoding difficulties. 2) Given the decoding difficulties specifically associated with dyslexia, it is important not to require children with dyslexia to decode proper nouns that are often idiosyncratic and do not contribute to understanding the meaning of a passage. Therefore, *reading proper nouns* may reduce reliance on decoding skills without invalidating the test. 3) The TAKS reading test is a comprehension assessment that determines how well the student understands the passage that he or she reads. The determination of comprehension is not dependent on the student's ability to read the comprehension questions. Therefore, no "read aloud" accommodation can be provided while the student actually reads the TAKS passages. However, the *comprehension stems can be read* to the student, and the student can provide a written response in the TAKS booklet in the standard manner, thus reducing demands on word decoding skills in students with dyslexia.

Although it would be interesting to try and determine which of these accommodations may or may not be most effective, much of the literature does not suggest that accommodations have any major effects on student performance. Therefore, the working group and the Dyslexia

Statewide Focus Group recommended that all three accommodations be provided in an effort to maximize the power of an accommodated administration of the TAKS.

Procedures

The study was conducted with a practice form of the Grade 3 TAKS developed in field testing. No modifications were made to the TAKS booklets; only instructions that examiners provided to students were modified. For the accommodated administration, the instructions were modified by TEA so that the TAKS could be administered in two sessions. In the first session, the student completed the practice story and the first two stories in the test; in the second session, the third story was completed. Booklets and scoring sheets were provided and scored by Pearson Educational Measurement, TEA's contractor for TAKS. A data file was provided with appropriate student identifiers (PEIMS numbers). These data were merged with demographic data and other performance data collected by CARS.

The students with dyslexia were third graders who were identified and served in programs specifically designed for students with dyslexia. Districts that maintained programs specifically for students with dyslexia were asked to participate in the study. These districts were in the Houston area and included the Crosby Independent School District, Deer Park Independent School District, Galena Park Independent School District, Goose Creek Consolidated Independent School District, Pasadena Independent School District, and the Spring Branch Independent School District. In these 6 districts, a total of 48 schools and 113 teachers participated.

Based on a power analysis designed to detect small effects of accommodations, the sampling plan was to recruit 100 students with dyslexia and 100 average decoders. Of the 100 students with dyslexia who completed the screening, 9 did not complete the study because the parents declined permission ($n = 2$), illness precluded participation ($n = 1$), the student transferred out of the district between screening and the TAKS assessment ($n = 4$), or the school declined

permission because it was felt the student was too severely impaired to take the test ($n = 2$). The latter issue was raised much more frequently, but we were usually successful in suggesting that students at all levels of reading ability should be included and that precautions would be taken to ensure that the student did not find the experience humiliating. As the 91 remaining students with dyslexia were matched with an average decoder from the same classroom, we only needed 91 comparison students.

To identify participants for this study, students who had previously been identified with dyslexia were administered the Letter-Word Identification and Word Attack subtests of the Woodcock-Johnson III Tests of Achievement (WJ). These tests are well-established measures of word recognition for real words and pseudowords, and form the Basic Reading composite, which has reliability (coefficient alpha) of .97. As significant difficulty in word reading is the primary marker of dyslexia, the administration of the WJ subtests was undertaken to ensure that the child had been appropriately identified as dyslexic and that the child remained impaired in reading, a necessary criterion because some students show significant improvement in reading skills after they begin programs for dyslexia. Students were accepted into the group with dyslexia only if they read below the 26th percentile on the Basic Reading skills cluster score of the WJ. Students in the comparison group were nominated by teachers as average readers from the same classes as the students with dyslexia. These students had WJ reading scores that were above the 39th percentile, but below the 76th percentile.

Table 1 presents demographic variables by group and condition. This table shows that the groups were comparable in age and demographic characteristics, with good representation of gender and different ethnicities. About half of each group qualified for a free/reduced lunch. The groups with dyslexia were slightly older (about 6 months) because some students had previously been retained. Most of these students were served through Section 504; less than 27% ($N = 24$) also identified for special education. We observed that districts were reluctant to

identify students in dyslexia programs if they were also identified for special education, expressing concern that the TAKS was not appropriate for students in special education and that these students should be exempted or take the SDAA.

Students within the dyslexic and comparison groups were randomly assigned to the accommodated and standard versions of the TAKS. To ensure that group size did not influence the results, all students within a particular school and condition were tested in small groups by a trained examiner supervised by CARS in a room provided by the school. The group sizes ranged from 2-8 students and always included students of comparable numbers from the dyslexia and comparison groups, thus controlling for group size. As Table 2 shows, most schools (N = 24) had just two students in the study. For both groups, the administration directions were read exactly as they appear. The CARS administrator followed guidelines developed for oral administration of the TAKS Grade 7 Mathematics test, reading only the proper nouns, comprehension stems, and possible answers. Students also had booklets containing the proper nouns and used these to read and follow along with the examiner. Students in both groups responded by writing their responses in the TAKS booklet, as oral responses would be disruptive in a small group setting. Although we attempted to blind examiners to group identity, students often had to be tested in their dyslexia classes. To ensure that the administrators did not inadvertently assist students with dyslexia, administration was highly scripted and the examiners were closely supervised.

In addition to the administration of the TAKS, each student received an individual assessment of oral vocabulary using the Woodcock Language Proficiency Battery-Revised Picture Vocabulary subtest. This well-standardized vocabulary assessment requires students to name pictures and has reliability (coefficient alpha) of .85. It was used in order to provide an index of the child's overall development of language. The purpose of this assessment was to determine whether accommodations differentially affected reading performance depending on

the child's level of language ability. Some individuals with dyslexia have specific reading disability, while others have dyslexia as part of a more pervasive disturbance of oral language. By administering an assessment of oral vocabulary, it was possible to determine if accommodations are differentially more effective in individuals who vary in their oral language development.

Statistical Analysis

From a statistical viewpoint, the *first question* involves the magnitude of the effect of accommodations for students with dyslexia. This question, framed in this way, presumes that the effect of accommodations is different for dyslexic and non-dyslexic students, i.e., that group and assessment type will interact. Thus, the *second question* involves the group by accommodation interaction. The *third question* is essentially the effect of the covariates (vocabulary and word reading ability) on group and accommodation status. Vocabulary was tested in terms of an interaction with group. Word reading is highly correlated with group, so the interaction was tested only within the group with dyslexia.

TIMES used SAS PROC MIXED to test for differences between groups. A mixed model is one that has both fixed effects and random effects. A fixed effect is one that is estimated to be the same for all members of a group, while a random effect permits members of a group to have varying values. In our sample, students can be simultaneously classified as belonging to classrooms, schools, and districts. This nesting of students is important because some of the variance in students' scores could be attributed to the placement of students in the same classroom, or the same school, or the same district. Current statistical models can take into account the dependence of observations at multiple levels of such a hierarchy. Because our sample was composed of only 1-6 students per teacher and only 2-12 students per school, we clustered schools within districts and calculated variance due to dependence of observations in

the same schools and in the same districts. Thus, the model ignored clustering at the teacher level but included clustering at the school and district level in order to obtain more accurate standard errors for estimates of the effect of accommodation. By using a mixed model instead of a standard regression model, we could allow different values for estimates at the school level, taking into account the fact that some schools are clustered together in the same district. Had clustering been ignored in the analysis, it is likely that standard errors would have been underestimated, resulting in an inaccurate (i.e., too liberal) test of the effects of accommodation.

An analysis of covariance model was fit that predicted TAKS standard scores from accommodation status, dyslexia status, vocabulary, and their interaction. As indicated above, the mixed model allowed that mean TAKS standard scores varied across students within schools, and schools within districts. Prior to the analysis, vocabulary scores were centered at the grand mean in order to allow for more straightforward interpretation of the model parameters. Specifically, centering the vocabulary measure at the grand mean forces the differences among groups to be estimated at the grand mean value of the covariate. Finally, we tested additional models to see if the covariate interacted with Group, with Accommodations, or with Group and Accommodations together. The effect of basic reading skills was addressed similarly but just within the group with dyslexia.

The mixed model analysis examined performance on the TAKS as a continuous scaled score. However, the TAKS also has established scale score cut-points that define state-mandated proficiency levels. Thus, on the basis of each student's TAKS continuous score, it was possible to determine whether the student had met the state's accountability standard. After converting each student's continuous score on the TAKS to Pass/Fail on the basis of the established cut-point, we conducted a log-linear analysis to evaluate the effects of dyslexia status, accommodations, and their interaction on TAKS pass rates. In this analysis, we modeled the

natural logarithm of the odds of passing the test as a function of Vocabulary, and Accommodation (Yes, No), Group (Dyslexic, Comparison), and the interaction of Accommodation and Group.

Results

Table 3 provides means, standard deviations, and ranges by group and administration condition for the Grade 3 TAKS reading test that are not adjusted for the vocabulary covariate. As expected, the groups with dyslexia had word reading scores that were at the 12-13th percentile and well below the comparison groups, who were slightly above average. Although students were not selected on the basis of their vocabulary development, the scores are comparable across groups and in the average range.

The mixed model analysis showed that the Group (Dyslexic, Comparison) X Accommodations (Yes, No) interaction was statistically significant, $F(1, 155) = 12.04, p = .0007$. In addition, regardless of the presence of dyslexia, higher vocabulary scores significantly predicted higher TAKS performance ($b = 4.17, SE = 0.76, F(1, 173) = 29.87, p < .0001$) but did not significantly interact with Group or Accommodation. The results clearly showed that students with dyslexia who received accommodations had significantly higher TAKS reading scale scores (least squares $M = 2059; SE = 19.1$) than those with dyslexia who did not receive accommodations (least squares $M = 1919; SE = 18.3$), $t(176) = 5.43, p < .0001$. In contrast, there was no difference between the performance of comparison students without dyslexia who received accommodations (least squares $M = 2179; SE = 19.2$) and those who took standard (least squares $M = 2162; SE = 18.3$), $t(176) = 0.68, p = .50$ administrations of the Grade 3 TAKS reading test. Please note that the least squares means reflect groups' expected performance estimated at the grand mean of vocabulary (that is, expected performance in the

group whose vocabulary is equal to the grand mean). These adjusted means will differ from the observed means in table 3 to the extent that groups differ in vocabulary and vocabulary is related to TAKS performance.

Although the models were fit allowing for variability in mean performance at the school level within districts, variance between schools was small relative to variability in students within schools ($\sigma^2_{\text{schools}} = 1,052, p < .078$; $\sigma^2_{\text{students within schools}} = 14,092$; ICC = .07). Fitting the model under various alternative covariance structures did not alter conclusions. In each case, accommodations improved performance for children with dyslexia and had no effect on performance for students without dyslexia. Thus, statistical conclusions about the effects of accommodation were unaffected by assumptions about the random effects structure of the data.

When models were fit that included the WJ Basic Reading skills cluster, the interaction between Accommodations and Group was not significant, $F(1, 171) = .15, p = .70$. However, such an effect is to be expected as group membership was defined on the basis of the Basic Reading skills cluster, leaving Group membership and Basic Reading skills cluster redundant with one another (see Table 3; point biserial correlation of group and Basic Reading = .89). Running the model within the group with dyslexia showed significant effects of accommodations, $F(1, 87) = 4.61, p < .04$, but no interaction of WJ Basic Reading and accommodations, $F(1, 85) < 1$. Thus, severity of reading disability is not related to the effects of the accommodations.

The mixed model analysis demonstrates that performance on the Grade 3 TAKS was significantly higher for children with dyslexia who took the TAKS under this “bundled” accommodations condition. However, this analysis did not address whether students with dyslexia were more likely to pass the TAKS if the test was taken under accommodated condition

or if these particular accommodations had a differential effect on pass rates for children with and without dyslexia. It is possible that average TAKS scale scores could be affected with no effects on the pass rates. Thus, to determine if accommodations differentially affected pass rates for children with and without dyslexia, we conducted a logistic regression analysis using vocabulary as the covariate, along with Accommodations, Group, and their interaction. This analysis predicted the log odds of passing, i.e., the natural logarithm of %passing divided by %failing as a function of vocabulary, Accommodations, Group, and their interaction.

The 140.4 point increase in the Grade 3 TAKS reading assessment scale score for children with dyslexia translated into a significant increase in the TAKS pass rate, with 18 of the 44 (41%) students with dyslexia who received accommodations achieving a passing score (three at a commended level) compared to 4 of the 47 students with dyslexia (9%) who received the standard administration of the Grade 3 TAKS reading assessment. This statistically significant difference in pass rates translates into a roughly seven-fold increase in the odds of passing the test when children with dyslexia are given appropriate accommodations ($.41/.59 = .695$ with accommodations vs. $.09/.91 = .099$ without accommodations; Wald Chi-square = 12.82, $p < .0005$) after controlling for vocabulary. Accommodations had no effect on the pass rate of students in the comparison group, which ranged from 77% for those receiving accommodations to 83% for those not receiving accommodations.

Conclusions

This study clearly establishes that the package of accommodations resulted in significant improvement in the performance of students identified with dyslexia on the Grade 3 TAKS reading assessment. The significant increase in the TAKS scale scores associated with accommodations translated to a significant difference in the pass rate for those with dyslexia who

received accommodations versus those with dyslexia who did not receive accommodations. The improvement could not be attributed to the student's level of vocabulary development or the severity of reading decoding difficulties. There was no effect of accommodations on the performance of the comparison group of average readers, suggesting that the accommodations were specific to the students with dyslexia and that the accommodations do not invalidate the TAKS.

Recommendations

1. These results are limited to students identified with dyslexia and served in districts with well-established dyslexia programs. Although we would expect that these benefits of accommodations would be apparent at other grade levels, the results of this study cannot be extrapolated to other grades, particularly as the nature of the TAKS reading assessment may change over time. High priority should be given to determining whether these results would extrapolate to other grade levels, including the state high school exit exam, and should be evaluated in a separate study. Studies of student characteristics associated with TAKS performance are strongly encouraged in order to plan effective instruction and suggest additional accommodations that might help those students with disabilities.
2. Students served in special education who have word reading difficulties should benefit from this accommodation package to the extent that students are similar to those evaluated in this study. Additional research should evaluate the effect of accommodations on students identified for special education who have reading difficulties. A limitation of this study reflects the tendency for districts not to refer students for the study if they were identified with dyslexia and also placed in special education. Those served in special education were felt to have reading problems of

sufficient severity to exempt them from TAKS. Although such an exclusion may be appropriate for students who are essentially nonreaders, the lack of relation of word recognition scores and TAKS outcomes in this study does not support this belief/practice, which should be further evaluated.

3. The approach taken in the current study is also viable for studying the impact of accommodations of varying types for students who are limited English proficient. In addition to the accommodations considered in the current study, specific accommodations that should be considered include the following: (a) providing access to background knowledge that is assumed for monolingual students; (b) providing access to dictionaries with information on word meanings which might be assumed for monolinguals; (c) providing support for common names and colloquial expressions which may be unfamiliar to English language learners but common knowledge to monolinguals. Whether such accommodations invalidate the test for English language learners will require careful study and consideration of the kinds of inferences that test scores are intended to support.
4. The study cannot establish an exact threshold for levels of decoding ability at which the accommodations would be effective because students with scores between the 26th and 40th percentile were excluded. Future evaluation should include this range of performance to try and establish a possible threshold.
5. The performance of students who participated in this study on previous administered Grade 3 TAKS/SDAA assessments should be compared.
6. Appendix II includes specific recommendations to consider in administering an accommodated version of the TAKS. These recommendations include the need to monitor use of test-taking strategies especially by students with disabilities for whom self-monitoring is often difficult. In addition, specific questions about the nature and

extent of possible accommodations for specific items are included that may help refine these accommodations.

Table 1. Age, gender, ethnicity, and free lunch status by group and condition

	Group/Condition			
	Dyslexic		Comparison	
	Accommodated	Standard	Accommodated	Standard
N	44	47	44	47
Age (months)				
M	115.3	115.9	110.6	111.2
SD	6.4	7.4	4.5	4.9
Gender				
N (%) Female	21 (48)	22 (47)	20 (46)	22 (47)
Ethnicity				
White	17 (39)	22 (47)	15 (34)	19 (40)
African-American	10 (23)	6 (13)	10 (23)	7 (15)
Hispanic	16 (37)	18 (38)	17 (39)	21 (45)
Asian/Other	1 (2)	1 (2)	2 (5)	0 (0)
Free Lunch ^a				
N	43	47	35	38
Free & Reduced				
N (%)	23 (54)	23 (49)	17 (49)	20 (53)

^aData on free lunch status was missing on 19 students

Table 2. Number of students within schools

Schools	Students
24	2
1	3
8	4
1	5
11	6
2	8
1	12

Table 3. Unadjusted means and standard deviations for the WJIII Basic Reading standard score, WLPB-R Picture Vocabulary standard score, and Grade 3 TAKS Reading scale score by Group and Condition.

Variable	Group/Condition			
	Dyslexic		Comparison	
	Accommodated	Standard	Accommodated	Standard
N	44	47	44	47
Basic Reading* Skills Cluster				
M	83.6	81.8	102.2	102.7
SD	6.0	6.7	3.0	4.1
Range	62-89	55-90	98-109	96-110
Picture Vocabulary*				
M	92.0	94.6	95.4	95.0
SD	12.1	14.2	12.2	10.3
Range	66-132	61-138	71-134	75-122
TAKS Reading				
M	2055.3	1921.7	2184.8	2166.7
SD	162.5	132.3	122.0	116.0
Range	1838-2588	1745-2309	1944-2460	1911-2588

Notes. WJ = Woodcock-Johnson; WLPB-R = Woodcock Language Proficiency Battery-Revised; TAKS = Texas Assessment of Knowledge and Skills. *Mean (SD) = 100 (15).

Tish Denney

David Francis

Katherine Immel

Sandy Maddox

Chris Schatschneider

Ann Smisko

Sharon Vaughn

Victoria Young

Michael Yoes

Appendix II: Additional Accommodations to Consider

- Before beginning the test, remind students to use TAKS strategies they have been taught by their teachers.
- Before beginning the test, briefly review TAKS strategies with students.
- When the examiner reads proper nouns aloud, ask students to point to each word with their pencil erasers. Therefore, examiners can monitor students to ensure that they are viewing the correct word as it is pronounced.
- Special consideration needs to be given to the following items on the Third Grade TAKS Reading Field Test and similar items on other forms of TAKS Reading:
 1. *Item 10 of Story Number One*: There is a diagram students view and read. The comprehension question follows and then the multiple-choice answers. The diagram is comprised of two interlocking circles. The first circle includes information about one

character in the story. The interlocking circle includes information about both characters with a blank space. One attribute is listed and the other is blank. The student must decide which multiple-choice answers belong in the blank.

- QUESTION: Should examiners read information in the diagram to students? If not, the average reader will have an unfair advantage over the dyslexic reader. This may also cause fatigue for the dyslexic reader.

- QUESTION: If it is decided examiners should read information in the diagram aloud to students, how should it be read? This is not scripted for the examiner.

2. *Item 14 of Story Number Two*: There is a table with one for students to view and read. (This would be more straightforward for examiners to read to students than Item 10.)

- QUESTION: Should examiners read information in the table to students before reading the comprehension questions and multiple-choice answers to them?

3. *Item 21 of Story Number Two*: There is a chart for students to read and study.

It contains four boxes in a sequence going down the page with text printed in three of them. The third box is empty. The student must decide which multiple-choice answer belongs in the blank box.

- QUESTION: Should examiners read information in the boxes to students? If so, when they get to the blank box, what should be said, "There is an empty box and the last one says..."?

4. *Item 24 of Story Number Two*: There is a picture of a notebook with a summary printed on it for students to read.

- QUESTION: Should examiners read the summary aloud to students before reading the comprehension question and multiple-answers?

5. *Item 27 of Story Number Three:* There is a two-sided chart of information for students to study and read before the multiple choice question is read aloud to them. On the right side of the chart there are three statements. On the left side of the chart there are two statements. There is a blank space for the third. Students must decide which multiple-choice answer belongs in that space.

- QUESTION: Should examiners read the information in the chart to students?