Abstract Title Page
Not included in page count.

Title: The Influence of Testing Prompt and Condition on Middle School Students' Retell Performance

Author(s):
Deborah K. Reed, Ph.D., The University of Texas at El Paso
Yaacov Petscher, Ph.D., Florida Center for Reading Research

Abstract Body

Limit 5 pages single spaced.

Background / Context:

Description of prior research and its intellectual context.

Retell protocols are commonly used for assessing reading comprehension (Cohen, Krustedt, & May, 2009) and, when included with ORF measures, contribute additional information about students' reading comprehension abilities (Marcotte & Hintze, 2009; Reed, Vaughn, & Petscher, in press). However, there is wide variability in the prompting conditions. Prior research suggests alterations in the wording of a prompt require different comprehension skills and result in different retell responses (Gagne, Bing, & Bing, 1977). In addition, there is disagreement as to whether oral or silent reading is more efficacious for retell performance (e.g., Hale et al., 2007, McCallum, Sharp, Bell, & George, 2004; Miller & Smith, 1990; Prior & Welling, 2001).

Purpose / Objective / Research Question / Focus of Study:

Description of the focus of the research.

The purpose of this study was to improve the utility of retell protocols by more clearly identifying the testing conditions under which students demonstrate better performance. The research questions concern whether the wording of the initial prompt, the inclusion of a follow-up prompt, or the opportunity to silently re-read the passage is related to retell performance. Because retells were scored quantitatively by the percent of pre-determined idea units included, it was hypothesized a prompt asking students to tell everything they remember would be associated with better scores than a prompt asking students to tell what a passage is "mostly about." Furthermore, it was hypothesized that students of various ability levels would have improved retells if provided an opportunity to re-read a passage silently prior to retelling it, rather than delivering their responses immediately following a 1-minute timed ORF component.

Setting:

Description of the research location.

Students were from two middle schools in Texas serving students in grades 6-8. Approximately 94% were tested during their English language arts class.

Population / Participants / Subjects:

Description of the participants in the study: who, how many, key features or characteristics.

The original sample included 589 sixth-, seventh-, and eighth-graders of all ability levels who returned parent permission forms to participate in the study. Seventeen of those students were absent or otherwise unavailable on the days of testing, and 45 students experienced testing abnormalities as explained in the section on fidelity of test administration procedures. This left a final sample of 527 students: 150 in grade 6, 168 in grade 7, and 209 in grade 8. Half of the students received free or reduced-price lunch, and 68% were Hispanic.

Intervention / Program / Practice:
Description of the intervention, program or practice, including details of administration and duration.

Students were randomly assigned to one of four retell testing conditions: 1) one-time oral reading and prompted to tell in their own words what the passage was mostly about; 2) one-time oral reading and prompted to tell everything they remembered; 3) one-time oral reading, prompted to tell everything they remembered, and provided follow-up prompting; 4) oral reading followed by silent rereading, prompted to tell everything they remembered, and provided follow-up prompting. The first condition was intended to capture an alternative way of asking for a retelling that would not specifically preference summarizing or giving a main idea ("what the passage was mostly about") and that would leave open the possibility of paraphrasing ("in your own words"). The latter was done to create more equivalency to the prompt used in the other three conditions, "tell me everything you remember reading in the passage." Similarly, the "tell everything" prompt avoided using the word "retell," so that the comparison between groups could more confidently be based on nuanced variations in prompt wording rather than on student's level of familiarity with a named skilled (i.e., summarizing, identifying a main idea, retelling). The fourth condition introduced silent reading after the oral reading required for the ORF component of the assessment. This was done for two reasons: 1) having a silent reading only condition would require that three additional passages be administered, thus potentially confounding results with measurement artifacts; and 2) combining oral and silent reading would allow for a comparison between conditions 3 and 4 that was not based on an advantage of one method of reading over another, but on the efficacy of allowing more time to reread for meaning.

Research Design:
Description of research design (e.g., qualitative case study, quasi-experimental design, secondary analysis, analytic essay, randomized field trial).

Randomized field trial

Data Collection and Analysis:
Description of the methods for collecting and analyzing data.

The Passage Reading Fluency (ORF and retell components) and Word Reading Fluency subtests of the Texas Middle School Fluency Assessment Assessment ([TMSFA], Texas Education Agency, University of Houston, & The University of Texas System, 2008) were administered in May at the End of Year testing point. The transcribed retell responses were scored by the principal investigator and two research assistants who had participated in the development of the scoring guides. Each response was compared to a set of pre-determined idea units for the passage. The scoring guides include allowable alternatives for wording the idea units, and scores are based on the percent of idea units accurately recalled out of the number possible for the word count attained. That is, students read to different places in the passage during the ORF portion of the assessment and, therefore, were eligible to recall a different number of idea units based on the total number of words read. The scoring guides provide the minimum word count needed to have read the information contained in each idea unit(see Appendix A for a sample scoring guide).

All data were manually entered into a database by the two research assistants involved in the development of the retell scoring guides. This provided another opportunity to check the accuracy of the scores recorded on the examiner documents. The research assistants also typed all student retells into the database before scoring them a second time to calculate inter-rater reliability. When entering the retell, assistants verified the number of words in the retell as recorded by the original examiner. Twenty-five percent of the retell responses were double scored for the purposes of determining inter-rater reliability among the three individuals applying the scoring guides. Using the database with the typed retells, an individual other than the original scorer applied the scoring guide to the student response a second time. Observed agreement on scores for individual passage retells ranged from a low of 82% (sixth-grade passage 1) to a high of 95% (seventh-grade passage 2). However, the average of the scores obtained on the three passages is the main unit of analysis for the TMSFA. Therefore, agreement was also calculated by determining the correlation of the average retell scores obtained by the two raters. The relationship, as measured by the intraclass correlation (i.e., .976), was strong and suggested agreement would not have occurred by chance.

Means on each outcome measure were calculated for the full sample. In addition, means for the average retell score and the length of retell utterance were calculated for each testing condition. These were analyzed using one-way analysis of variance and Cohen's *d* (Cohen, 1988). The Benjamini-Hochberg correction (Benjamini & Hochberg, 1995) was utilized to control for the false-discovery rate, rather than the family-wise error rate.

Findings / Results:
Description of the main findings with specific details.

Because attrition in randomized experimental designs may create differences among the groups, it was important to assess the extent to which attrition contributed to a bias of estimated treatment effects. This was examined through a comparison of the differential proportion of missing data between the treatment and control groups with the overall proportion of missing data. Using guidelines outlined by What Works Clearinghouse (2008), the relationship between total missing and differential missing data for each outcome was determined to be at an acceptable level of bias under conservative assumptions. Students retold, on average, only about 31% of the idea units they read in the passage. Their responses were about 33 to 38 words in length. The ORF portion of the Passage Reading Fluency subtest was strongly correlated (r = .861) to the Word Reading Fluency subtest, an ORF assessment of words in isolation. However, retell scores bore little relationship to ORF scores. With the exception of the weak correlation between retell scores on passage 2 with the Passage Reading Fluency equated score, only the lengths of utterances in retelling were significantly (r = .185 to .300) and mostly weakly related to the ORF components. This suggests retell is measuring different reading skills.

Both the type of prompt and the use of follow-up prompting were significantly related to the percentage of pre-determined idea units retold. A change to either the wording or the addition of follow-up prompting produced approximately moderate effect sizes (d = .44 – .62). The combination of changes produced strong effect sizes (d = .96 – 1.05). However, the addition of silent reading did not significantly improve performance. Similar results were found when comparing the mean lengths of utterance. That is, students said significantly more words when

prompted to "tell everything" and when encouraged to continue with follow-up prompting, but length of utterance was not significantly different when provided an opportunity to silently reread the passage before retelling on passages 2 and 3. On passage 1, the easiest of the three passages read, student retells were modestly though significantly longer when allowed to silently reread.

Conclusions:

Description of conclusions, recommendations, and limitations based on findings.

When scored quantitatively, an initial prompt that asks students to tell everything they remember results in better responses than telling what the passage was "mostly about." Moreover, encouraging students to continue retelling by asking if they remembered anything else, also significantly increases students' scores. Offering addition time to reread a passage silently does not significantly improve retell as compared to retelling immediately after an ORF assessment.

The results were similar when considering the number of words students included in their retells. The length of utterance was significantly longer when asked to tell everything and was longer still when encouraged to keep retelling. It seems students considered the initial prompt a literal indicator of what they were to do: either give a succinct "gist" of what the passage was "mostly about" or give as much information as possible. Yet even with the "tell everything" prompt, students did not provide all the information they actually recalled unless they were repeatedly prompted to continue retelling.

These are important results given the great amount of variance in the prompts used across retell instruments (Reed, manuscript under review; Reed & Vaughn, manuscript under review). Attempts to synthesize or interpret retell scores from different studies will be problematic if there is not a movement toward greater consistency in the wording and delivery of prompts. An equally compelling finding was the lack of significant differences in the scores of students provided identically worded prompts and follow-up prompting, but a difference in the method of reading prior to retelling. Because retell is being included with ORF progress monitoring instruments (Good & Kaminski, 2010), adding just one minute onto the testing time for each passage can impact the feasibility of using frequently administered measures. The lack of strong correlations between retell and the ORF components of the TMSFA suggest retell is measuring some other skill than what is being captured by ORF (Marcotte & Hintze, 2009; Reed, Vaughn, & Petsher, in press). Therefore, retell has the potential to better inform reading instruction if administered under conditions that preserve instructional time and optimize student performance.

Appendices
Not included in page count.

Retell Scoring Guide

Laura
(page 1 of 3)

Each row of the table represents an idea unit and is worth 1 point.

# of Words Read	Recalled Idea Unit	Maximum Possible Idea Units
7	A girl / A lady / Laura — was a — writer / author — *or* — wrote children's books	1
24	She / Laura — was born — in a log house / in Wisconsin / in 1867	2
37	She / Laura — was close to her older sister / had an older sister named Mary — *or* — was the second daughter of — Charles / Caroline	3
65	They / Her family — had to move — so her / wherever their — dad / father — could — find / get — work / a job — because her dad didn't have	4
72	They / Her family — moved — after Laura was born / to Missouri	5
82	They / Her family — moved — to start a farm / to Kansas / where her sister Carrie was born	6
105	They / Her family — moved — back to Wisconsin / to be near family	7
122	They / Her family — moved — in 1874 / to Walnut Grove / to Minnesota — *or* — Her dad/father / Charles — wanted / longed — to farm	8
151	The girls / The kids / Laura and her sisters — had many adventures	9
176	Grasshoppers / Bugs — ate / destroyed — their crops / their wheat crop	10

Appendix A. References

References are to be in APA version 6 format.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B (Methodological), 57*, 289-300.

Cohen, J. (1988). *Statistical power analysis of the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cohen, L., Krustedt, R. L., & May, M. (2009). Fluency, text structure, and retelling: A complex relationship. *Reading Horizons, 49*(2), 101-124.

Gagne, E. D., Bing, S. B., & Bing, J. R. (1977). Combined effect of goal organization and test expectations on organization in free recall following learning from text. *Journal of Educational Psychology, 69*(4), 428-431.

Good, R. H., & Kaminski, R. A. (2010). *Dynamic Indicators of Basic Early Literacy Skills* (6th ed.). Eugene, OR: Dynamic Measurement Group, Inc. Available: http://www.dibels.org.

Hale, A. D., Skinner, C. H., Williams, J., Hawkins, R., Neddenriep, C. E., & Dizer, J. (2007). Comparing comprehension following silent and aloud reading across elementary and secondary students: Implication for curriculum-based measurement. *The Behavior Analyst Today, 8*(1), 9-23.

Marcotte, A. M., & Hintze, J. M. (2009). Incremental and predictive utility of formative assessment methods of reading comprehension. *Journal of School Psychology, 47*, 315-335.

McCallum, R. S., Sharp, S., Bell, S. M., & George, T. (2004). Silent versus oral reading comprehension and efficiency. *Psychology in the Schools, 41*(2), 2, 41-246.

Miller, S.D., & Smith, D.E.P. (1990). Relations among oral reading, silent reading and listening comprehension of students at differing competency levels. *Reading Research and Instruction, 29*, 73–84.

Prior, S.M., &Welling, K.A. (2001). Read in your head: A Vygotskian analysis of the transition from oral to silent reading. *Reading Psychology, 22*, 1–15.

Reed, D. K. (manuscript under review). The technical adequacy of retell instruments. *Educational Assessment*.

Reed, D. K., & Vaughn, S. (manuscript under review). Retell as an indicator of reading comprehension. *Scientific Studies of Reading*.

Reed, D. K., Vaughn, S., & Petscher, Y. (in press). The contribution of retell to the model of adolescent reading competence. *Journal of Learning Disabilities*.

Texas Education Agency, University of Houston, & The University of Texas System. (2008). *Texas middle school fluency assessment*. Austin, TX: Author.

What Works Clearinghouse (2008). *Procedures and standards handbook: Version 2.0*. Washington, D.C.: Author.

Appendix B. Tables and Figures

Not included in page count.

Table 1
Descriptive Statistics of Full Sample

	N	Mean	SD
Retell 1 Score	527	30%	20%
Retell 2 Score	527	29%	22%
Retell 3 Score	527	35%	26%
Ave Retell	526	31%	16%
Retell 1 Word Count	527	38.17	17.70
Retell 2 Word Count	527	32.63	16.23
Retell 3 Word Count	527	37.33	18.44
PRF	527	160.66	34.12
WRF	527	76.82	20.91
TAKS	509	802.02	97.55
Benchmark	497	87.49	11.50
Valid N (listwise)	488		

Note. PRF = Passage Reading Fluency subtest of the TMSFA; WRF = Word Reading Fluency subtest of the TMSFA

Table 2
Correlations Among Measures

	1	2	3	4	5	6	7	8
1. Retell 1 Score								
2. Retell 2 Score	.219**							
3. Retell 3 Score	.321**	.213**						
4. Ave Retell	.694**	.659**	.764**					
5. Retell 1 Word Count	.653**	.271**	.289**	.554**				
6. Retell 2 Word Count	.407**	.518**	.255**	.542**	.660**			
7. Retell 3 Word Count	.376**	.296**	.543**	.582**	.548**	.554**		
8. PRF	-.066	-.095*	.008	-.068	.300**	.250**	.213**	
9. WRF	-.074	.002	.000	-.032	.297**	.276**	.185**	.861**

**p< .01
*p< .05

Table 3

Descriptive Statistics by Condition

	Condition 1			Condition 2			Condition 3			Condition
asure	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean
ell 1	129	20%	18%	151	29%	18%	134	30%	20%	113	37%
ell 2	129	19%	19%	151	28%	20%	134	30%	20%	113	35%
ell 3	129	26%	25%	151	35%	24%	134	40%	30%	113	39%
g Retell	128	22%	14%	151	30%	14%	134	40%	20%	113	37%
ell 1 WC	129	27.86	13.35	151	36.11	15.18	134	42.4	15.8	113	47.71
ell 2 WC	129	24.86	14.21	151	30.39	12.71	134	35.8	16.1	113	40.78
ell 3 WC	129	29.06	15.5	151	36.47	15.46	134	41.8	20.7	113	42.62
F	129	159.82	39.58	151	162.11	34.34	134	158.6	29.3	113	162.13
.F	129	75.46	22.59	151	76.56	20.93	134	76	19.6	113	79.7

Note. PRF = Passage Reading Fluency subtest of the TMSFA; WRF = Word Reading Fluency subtest of the TMSFA

Table 4

One-way Analysis of Variance: Average Retell

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	1.882	3	.627	28.156	.000
Within Groups	11.633	522	.022		
Total	13.515	525			

Table 5

One-way Analysis of Variance: Length of Utterance

		Sum of Squares	df	Mean Square	F	Sig.
Retell 1 Word Count	Between Groups	26992.037	3	8997.346	34.141	.000
	Within Groups	137828.247	523	263.534		
	Total	164820.284	526			
Retell 2 Word Count	Between Groups	17349.081	3	5783.027	24.939	.000
	Within Groups	121278.278	523	231.890		
	Total	138627.359	526			
Retell 3 Word Count	Between Groups	14758.303	3	4919.434	15.678	.000
	Within Groups	164103.906	523	313.774		
	Total	178862.209	526			