Effects of Text Grade Level on Oral Reading with Mildly Handicapped Students

by

Bill Wolking, Carolyn Harris, Jolenea Ferro and Jack Scott[1]

Readability has been a central construct of reading instruction for the past half century or more. It is determined by an analysis of the structure of text. Counts of the number of words per sentence, number of syllables per word, number of times words appear in a passage, and number of subordinate clauses per paragraph are typical of the data used in readability formulas (Giordano, 1987; Yarington, 1979). Grade level classifications are the outcome of applying readability formulas to reading passages. Structurally determined grade levels are the foundation on which basal reading series are built. Many school systems require teachers to use basal reading series and to use the instruction found in their manuals when placing students and teaching reading (Goodman, 1987). It is estimated that between 80 and 95% of elementary school teachers use basal readers (Cassidy, 1987; Yarington, 1979). When most teachers speak of the readability of a passage, they mean the assumed difficulty based on structural characteristics of the text.

In spite of their common use, structural approaches have been criticized. Some educators believe readability formulas sacrifice meaning and comprehensibility to the control of text dictated by their use (Bell & Willems, 1986). More emphasis is placed on writing a 'grade level' passage than on writing an interesting passage. Strict control over student progress by time on grade level without giving enough weight to student mastery is another criticism of the structural approach (Cadenhead, 1987).

Functional approaches to readability are developing. Mounting criticism of structural approaches combined with a growing awareness of direct measurement and curriculum based assessment strategies are two possible reasons for the explorations of functional approaches to readability. Functional approaches are based on an analysis of student reading performances rather than on the structural characteristics of text passages. Mills (1988) and Kaufman (1989) have studied the effects of a range of grade levels of text on the speed and accuracy of student oral reading performances. Both Mills and Kaufman found that as grade level of passages increased there was a moderate decrease in the speed of oral reading, but only a slight increase in error rates. Kaufman also studied the effect of grade level on comprehension, finding a moderate decrease in comprehension with increasing grade level. Scott (1988) & Scott et al. (1990) studied the effect of high-challenge reading passages on rate of learning. High-challenge passages were defined as those that produced frustration level performances on first reading according to traditional placement criteria for percentage of errors made. He found that learning on high-challenge passages was equal or superior to learning on passages at recommended "instructional" placement levels.

The present research extends the exploration of functional approaches to readability. It reports 22 single case experiments (or a single case experiment with 21 systematic replications) of the effects of grade level on the rate of learning and also on the frequency of correct error responses. This work combines the structural and functional approaches. The passages used had previously been assigned grade level classifications based on structural analysis. Student performances were then studied across a range of structurally different passages.

METHOD

Teacher Researchers. Twenty-two preservice trainees of the Department of Special Education at the University of Florida conducted these experiments. They had just completed a six week, ten hour a week, course in precision teaching in which they had conducted a study of the effects of three widely disparate levels of text difficulty on the reading performance of their lab partners. This

1The authors express their appreciation to the teachers and students whose caring and dedicated work produced the studies reported in this paper.

study was a training exercise designed to provide a model for the research reported in this article. Following these experiences, each teacher was assigned to tutor an underachieving or special education student.

Students. The students were referred by their classroom teachers for an after-school program of tutoring by precision teaching trainees. They ranged in age from 6 to 15, two-thirds were male, and most were from lower-middle and lower socioeconomic families. They were referred because they had demonstrated learning, behavior, achievement, or motivation problems. All the students attended a laboratory school that does not classify students by traditional special education labels (learning disabled, behavior disordered, etc.). If these students were attending public school classes in the same county, the majority of them would have been classified as learning disabled or behavior disordered. About 30 percent had attended our after-school program previously, so had some experience with precision teaching.

We are not able to report detailed information on IQs, family socioeconomic status, sensory modality strengths and weaknesses. We intentionally discourage our teachers from using this type of information. We train them to use more immediate and instructionally relevant information obtained from curriculum based assessment timings, and we also encourage them to adopt the philosophy that it is the job of special education teachers to produce substantial rates of learning and retention regardless of the demographic and educational status of a child from the traditional psycho-educational perspective.

Setting. Tutoring took place in three classrooms in the laboratory school mentioned above. Younger students were taught in classrooms in the elementary wing. Between five and eight tutor-tutee pairs worked in each room. Instruction took place about 10 minutes after the end of the regular school day, three times a week for ten weeks. The duration of the experiments described in this study was between two and five weeks--4 to 16 rated days, within a range of 8 to 35 calendar days.

Research Procedure. Each of the 22 studies conducted was a single subject experiment using an alternating treatments design. This design was used to control for sequence and maturation effects, and also to produce daily opportunities to observe experimental effects as the learning outcomes from each of three levels of text were charted.
All students were taught reading in three consecutive mini-lessons, each 5 to 8 minutes long. Each mini-lesson used a different grade level of text. The within-day sequence of lessons across grade levels was randomized so that each grade level was either first, second, or third approximately the same number of days.

Instructional procedures. In the precision teaching course just completed by these students, we taught and recommended the use of a rereading instructional procedure. It involved having the teacher model how to read the passage correctly, fluently and with good inflection. The readings were either live or tape recorded. The student was asked to mimic the teacher as the reading occurred. Most teachers modeled the passage two or three times. These modeling-imitation tactics were followed by opportunities to practice especially troublesome words or phrases and then by at least two timings, with the best one charted.

Most teachers followed this procedure closely, but some added contingent error correction and additional timings. All made a written record of their procedures and used the same procedure across the three grade levels and for the duration of the study. Clinical supervisors monitored the implementation of teaching procedures on a daily basis to ensure that they were faithful to written plans and that they did not 'drift' from the original implementation over time. We did not force a strict adherence to a single instructional procedure across teachers, but almost all stayed close to the modeling-imitation procedure we recommended.

Grade level placements for the instructional level passage were determined by following a method published by Woods & Moe (1985) as part of a system of individualized reading instruction. The instructional level so determined is designated as the Level 1 passage in this report. Levels 2 and 3 passages were selected by using material interesting to the student that also were a minimum of one and two grade levels higher respectively. We encouraged teachers to be daring in their selection of the higher grade level passages. However, a few were timid and a few were tutoring slow learning first grade students whose behavior was essentially shut down by jump-ups or leap-ups of several grade levels.

All plans and outcomes were recorded on a form specifically designed to accommodate precision teaching instructional procedures and learning outcomes. These forms were used by the teachers when entering their procedures and outcomes into a

desk top computer data base.

The Teachers and Learners - Plans and Outcomes Database (TLPO) Our teachers use a computer data base for all of their instructional plans and learning outcomes. This gives our students experience with desk top computer management of educational data, and it gives faculty supervisors an opportunity to have all plans and outcomes in a standard and easily readable and editable form.

A Macintosh SE computer running the FileMaker II data base application comprised the hardware and software used. Data for each instructional plan (phase of instruction) was entered and became a record in the TLPO data base. Each record includes: identifying information and codes (five fields), the pinpoint and instructional plan (eight fields), the primary learning outcomes data (six fields), and a number of learning outcomes computed by the program (fourteen fields). The computed learning outcomes appear instantly as the last item of information is entered into the record, giving the teacher a full and immediate quantitative description of the effectiveness of the plan. Of course the teacher has seen these outcomes previously in graphic form on the standard celeration chart. The computed learning outcomes serve a different purpose. They are not a prime source of information that controls instructional decisions. They do provide an opportunity for teachers to view and evaluate their teaching from a larger perspective, across plans, students, pinpoints, etc. More importantly from the perspective of this study, they give supervisors an opportunity to monitor all instruction and to provide constructive feedback quickly, in a standard format, that is closely tied to facts of instructional plans and outcomes.

Data from the TLPO data base were exported to JMP and Data Desk Professional, Macintosh statistical applications. These programs produced the variable distributions, descriptive statistics, and correlations on which our results are based.

RESULTS

The results are presented in two ways. Table 1 presents a summary of the data for each of the 22 single case studies. These results are presented in a table rather than on the standard celeration chart to save journal pages. Although a preferred format, it would take many chart pages to present the data in Table 1. Table 2 presents aggregated data for the 22 studies. Chart 1 depicts the aggregated data on the standard celeration chart. Table 3 presents correlation coefficients between selected variables.

Data for Individuals. Results of the 22 single case experiments appear in Table 1. Starting with the column on the left and moving a column at a time to the right, the columns contain: the case code and number of rated days, grade levels of the three passages used, first and last frequencies for correct responding, first and last frequencies for error responding, celerations of correct responding, and celerations for error responding in the last column on the right. The 22 experiments are arranged in ascending order by the grade level of the instructional (lowest) level passage for each case.

A criterion of .20 (20% per week) in the difference between celerations for the lowest and highest grade levels was used to evaluate the effect of grade level on rate of learning. This criterion was selected to represent a difference in the rate of learning which would be educationally significant in brief, two to three weeks, periods of instruction. The educational significance of a .20 difference in celeration over two and three week periods of instruction is easier to appreciate with an example. If a student had celerations of x1.25 and x1.45 for words read correctly (a .20 difference in celerations) respectively on their Level 1 and 3 passages for two weeks, there would be a 54.5 wpm difference. At the end of three weeks the difference in rate of reading words correctly would be 110 wpm.

Grade level and correct celerations. In 18 of the 22 studies, the celerations of the lowest and highest grade level passage did not meet this criterion. That is, by our criterion there was not an educationally significant difference in the rate in learning when Level 1 and 3 passages were compared. In all four cases that met the criterion the celeration of correct responding was higher on the highest grade level passage. These celerations are followed by double asterisks in the fifth column from the left of Table 1.

Grade level and error celerations. Using the same criterion, .20 with the celerations for error responding, 11 studies found no educationally significant difference in the rate of error reduction comparing the lowest and highest grade level passages. In 7 of the 11 cases that do meet the criterion, the higher rate of error reduction is associated with the most difficult level of material. These celerations are marked with double asterisks in the sixth column from the left of Table 1.

Table 1: Data for 22 teacher-student cases
Arranged in ascending order by Level 1 text grade level

Case Code Rated Days	Text Grade Levels	Correct Responses First-Last*	Error Responses First-Last	Celerations Correct Responses	Celerations Error Responses
haman-	0 Preprimer	92-150	3-0	x1.15	/ 1.05
forse	1	88-158	2-0	x1.10	/ 1.20
8	2	60-121	6-3	x1.20	/ 1.20
morke-	0 Preprimer	75-224	0-0	x1.85	/ 1.00
kesal	1	52-108	0-0	x1.55	/ 1.00
9	3	28-134	0-0	x1.70	/ 1.00
newmi-	1	97-112	0-0	x1.15	/ 1.00
booja	2	82-132	0-0	x1.17	/ 1.40
6	4	144-173	3-0	x1.20	/ 1.00
brava-	2	57-154	0-0	x1.20	/ 1.00
simka	3	62-150	0-0	x1.20	/ 1.00
9	4	50-135	0-0	x1.15	/ 1.00
tyrke-	2	88-125	0-0	x1.90	/ 1.00
booce	3	58-104	0-0	x1.80	/ 1.00
6	4	31-97	0-1	x3.00**	/ 3.00**
hayja-	3	127-168	2-1	x1.20	/ 1.80***
gusje	4	106-200	2-0	x1.10	/ 1.45
16	5	91-220	2-0	x1.25	/ 1.00
jonmi-	3	118-251	2-2	x1.19	/ 1.45***
brepa	4	118-219	2-1	x1.25	/ 1.80
7	6	87-207	0-0	x1.24	x1.06
folbr-	4	113-212	1-0	x1.15	/ 1.00
thosh	7	112-208	4-0	x1.20	/ 1.10
16	9	94-204	6-0	x1.15	/ 1.15
mccro-	4	196-206	3-1	x1.18	/ 3.40***
gadsa	7	122-218	4-1	x1.40	x1.75
7	13	109-204	3-2	x1.17	/ 1.65
milra-	4	112-232	4-4	x1.38	/ 1.00
doumi	5	106-199	5-6	x1.50	/ 1.10
9	6	150-171	5-3	x1.25	/ 2.15**
pribe-	4	96-229	2-5	x1.35	x1.96
mcdmi	7	100-160	1-0	x1.30	/ 1.15
6	9	127-218	0-1	x1.55**	/ 1.00**
harle-	5	146-245	3-9	x1.35	x1.54***
waljo	8	122-278	10-7	x1.40	/ 1.30
7	12	123-229	4-6	x2.35**	x1.76
milki-	5	150-223	3-0	x1.25	/ 1.80
rojsh	7	147-226	8-1	x1.45	/ 1.85
6	9	112-177	8-0	x1.10	/ 1.90
prica-	5	82-137	3-2	x1.10	/ 1.15
oweca	7	71-122	6-1	x1.15	/ 1.50
7	8	54-130	7-2	x1.35**	/ 1.00

sieed- mcdla 13	5 8 13	137-311 106-233 88-230	3-0 6-2 2-2	x1.15 x1.10 x1.20	/ 1.25 / 1.05 / 1.00
wilcr- stren 8	5 8 9	73-131 69-106 80-108	4-0 17-5 17-2	x1.25 x1.25 x1.20	/ 1.75 / 1.70 / 4.30**
cumka- colma 5	6 7 8	131-157 112-208 94-204	0-0 1-0 4-2	x1.10 x1.80 x1.10	/ 1.00 / 1.00 / 1.40**
ranpa- plato 5	6 7 8	99-181 137-204 148-192	2-1 4-0 2-0	x1.50 x1.60 x1.30	/ 2.00 / 4.30 / 2.50**
tibje- pauev 5	6 8 13	156-238 160-232 162-200	0-0 2-0 2-0	x1.14 x1.20 x1.10	x1.00 / 1.30 / 1.00
ocoal- alfja 10	7 9 13	157-213 100-202 97-187	1-1 4-1 2-0	x1.07 x1.25 x1.14	x1.00 / 1.36 / 1.03
weiga- cresc 8	7 9 10	124-210 142-190 112-189	3-0 3-1 0-2	x1.25 x1.10 x1.20	/ 1.20 / 1.20 / 1.00
viama- wooas 4	9 11 13	211-230 112-132 119-124	1-0 6-3 9-3	x1.10 x1.15 x1.15	/ 1.10 / 1.80 / 3.30**

*First and last frequencies (movements per minute) for each grade level of text.
**Celerations for correct responding exceed the .20 difference criterion and show faster rates of learning with the high grade level of text.
***Celeration for error responding exceed the .20 difference criterion and show faster rates of learning with the low grade level material.

Grade level and first frequencies for correct responding. Grade Level 3 text was associated with a decrease in the first frequency of correct responding in 16 of the 22 cases. These 16 students dropped between 7 and 92 wpm. In four cases students read faster on Level 3, compared to Level 1 passages. Level 3 was read faster by 31 to 49 wpm in these four experiments. In two cases the grade level of the passages had essentially no effect on first frequencies, only 6 and 7 wpm differences between Level 1 and 3 passages.

Grade level and frequencies for error responding. Effects of text grade levels on first frequencies for error responding were less consistent. In 12 studies errors on Level 3 text were higher. Increases ranged from 1 to 13. In six studies there was no difference in errors between Levels 1 and 3 passages. In four studies errors decreased by between 1 and 3 per minute on Level 3 passages.

Grade level and reaching performance standard. While not the only justifiable value, 200 wpm was used as a standard to represent minimally proficient oral reading in this study. The proportion of cases in which students reached or exceeded 200 correct wpm at each level are: Level 1 passages, 13 of 22 (59%); Level 2 passages, 11 of 22 (50%); and Level 3 passages 9 of 22 (41%). These findings must be interpreted conservatively since the number of rated and calendar days was not held constant across the 22 studies. Those teachers choosing to teach longer had a higher probability of reaching performance standard. However, it is interesting to note that within relatively short instructional phases (median 7 rated days within 17 calendar days) more than one-third of the students reached or exceeded the performance standard on text several grade levels above traditional placement levels.

Group Data. Highest, median, and lowest frequencies and celerations for correct and error responding are presented in Table 2 and depicted in Chart 1. The median celerations for correct responding on Level 1, 2 and 3 passages are X 1.20, X 1.25, and X 1.23 respectively. Median celerations for error responding on Level 1, 2, and 3 passages are / 1.06/ 1.19, and / 1.11 respectively. Median celerations for both correct and error responding vary by no more than 50% per week.

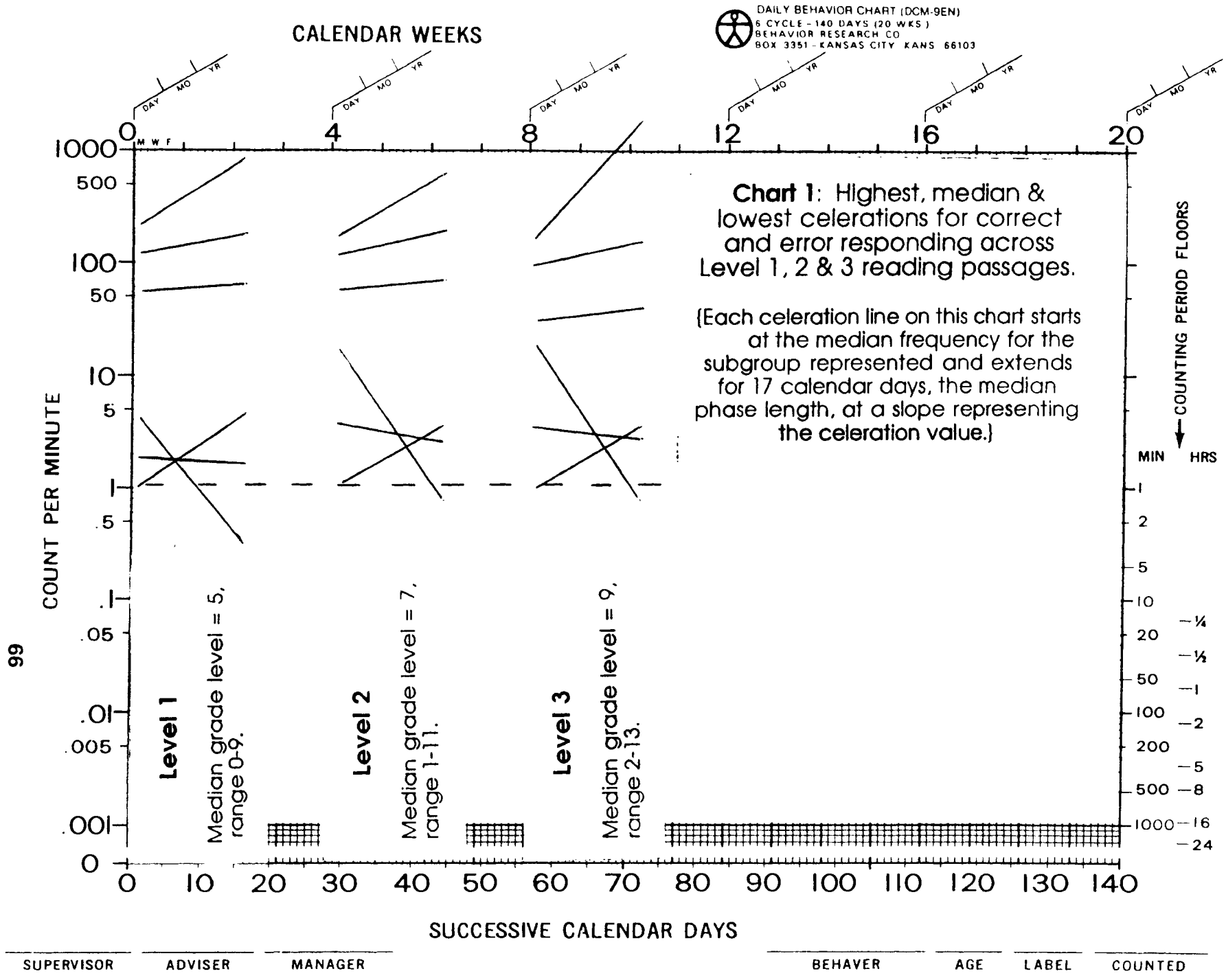
Chart 1 depicts the highest, median, and lowest celerations. Each median celeration line is drawn from the median first frequency for that cell in Table 2, e.g., highest celeration for corrects, median celeration for errors, etc. The celeration lines extend for 17 calendar days to depict the median length of each experiment. The median experiment was seven rated days within a period of 17 calendar days. The overall impression conveyed by Chart 1 is that grade level of material typically has only a small effect of the rate of learning for both correct and error responding. In other words, most students similar to the ones taught in these experiments learn about as well using any grade level of text within a range of four to eight grade levels above instructional placement level. A study of Chart 1 also leads to the observation that the grade level of text used to teach oral reading does have an effect of the speed of correct responding. Reading a Level 3 passage (median 9th grade level) was 23 wpm slower than reading a Level 1 passage (median 4th grade level). This is approximately a four to five word per minute decrease in speed for each increase in grade level. Median error responses for Level 1 and 3 text increased from 1.93 to 3.10 errors per minute. This is an increase of .23 errors per minute for each increase in grade level of the text.

Other Findings. Several questions might be asked in an attempt to understand more completely the effects of grade level of student learning and performance. *What are the relationships between absolute grade level and celerations for correct and error responding?* The Pearson correlation coefficients for correct and error responding respectively were -0.130 and +0.155 (N=66). *What is the relationship between jump or leap-up size and celerations for correct and error responding?* The correlations between jump-up size and celerations for correct and error responding respectively were -0.102 and -.247 (N=22). *What is the relationship between initial frequency of correct and error responding and the celerations for correct and error responding?* The correlations between initial correct frequencies and the celerations for correct and error responding were -.270 and +0.020 (N+66). The correlations between initial error frequencies and the celerations for correct and error responding were -0.119 and +0.437 (N=66).

Initial error frequencies are associated with about 19% of the variance in celerations for error responding. None of the other correlations support a conclusion that would detract from the interpretation that grade level of text is the variable responsible for our results.

Table 2: Highest, median and lowest beginning frequencies, celerations, and ending frequencies for correct and error responding at each of three text grade levels.			
Variable Name	**First Frequencies***	**Celeration# Value**	**Last Frequencies**
Level 1: Grade level text. ^ Median 5th grade, Range from 0 (preprimer) through 9.			
Highest correct	211	x 1.90	311
Median Correct	115	x 1.20	206
Lowest Correct	57	x 1.07	112
Highest Error	4.00	/ 3.40	9.00
Median Error	1.93	/ 1.06	1.23
Lowest Error	0	x1.96	0
Level 2: A step- or jump-up from grade level text. Median 7th grade level. Range from 1 through 11.			
Highest Correct	160	x 1.80	278
Median Correct	101	x 1.25	196
Lowest Correct	52	x 1.10	104
Highest Error	17	/ 4.30	7
Median Error	3.37	/ 1.19	0.50
Lowest Error	0	x1.75	0
Level 3: A jump- or leap-up from grade level text. Median 9th grade level. Range from 2 through 13 (college level).			
Highest Correct	162	x 3.00	230
Median Correct	92	x 1.23	182
Lowest Correct	28	x 1.10	97
Highest Error	17	/ 4.30	6
Median Error	3.10	/ 1.11	1.07
Lowest Error	0	x1.76	0

* Frequency = movements per minute
\# Median number of rated days per experimental phase = 11. Range from 4 to 16.
^ Grade level determined by traditional IRI placement procedures



99

SUPERVISOR ADVISER MANAGER BEHAVER AGE LABEL COUNTED

Lindsley (1990) has held that correct and error frequencies and correct and error celerations are each functionally independent. Based on our data, the correlations between correct and error initial frequencies and correct and error celerations respectively were -0.024 and +.206 (n=66). These data support an interpretation of independence between these two pairs of variables.

DISCUSSION & CONCLUSIONS

We have been careful not to use words that equate grade level of text with the difficulty of teaching or learning. Many teachers use 'grade level' and 'difficulty level' as synonyms, apparently on the assumption that there is a high correlation between the two. The results of the experiments reported in this paper support the conclusion that for most of our students there was only a loose association between the grade level of the text used in teaching oral reading and difficulty of *teaching and learning*. However, our findings indicate that *speed of correct responding* is affected by grade level of text used in teaching. There was a 4.6 wpm (4%) decrease in speed for each increase in grade level.

We draw two general conclusions from these results. First, *rate of learning* is not usually adversely affected by teaching reading with text that jumps- or leaps-up above instructional placement levels by several grades. Second, *absolute performance rates* for both correct and error responding are adversely affected. As far as our findings go we have a trade-off. Increasing grade level above instructional level generally has a positive effect on rate of learning and a negative effect on absolute rate of correct and error responding. To us, the increase in error frequencies seems surprisingly small in almost all cases.

Our findings do not provide relevant data on the issue of whether the higher rates of learning and lower absolute rates of ending correct responding necessarily go together. Our teachers were instructed to end their teaching on all three grade levels when they believed they had enough data to draw a conclusion about the effects of grade level on rates of learning. Although we measured the final performance in relation to a performance standard, we made no attempt to have teachers continue instruction until performance on each level flattened or reached the performance standard. In retrospect this was an error and should be a tactic to consider in future research. Continuing instruction until there was no more improvement or a performance standard was reached could provide information about whether ending performance frequencies on high grade-level passages could equal or exceed those on instructional level passages.

Direct, but unmeasured and unrecorded, observations of our teachers and students lead us to suspect that there are valuable side-effects of learning to read on text above instructional placement levels for both students and teachers. Although initially resistant to reading the high grade level passages, most students were expressing surprise at, and pride in their achievement when the experiments ended. Some said that the stories were more interesting and easier to read, and some were observed to spend more time reading and to request additional reading materials to take home. Many of the teachers were also skeptical about using the above-instructional-level passages, but almost all became enthusiastic about using the higher grade level reading material. We believe it will prove worth the investment to investigate these and related side-effects on teachers and students.

REFERENCES

Bell, M. & Willems, A. (1986). A comparison of the readability level of selected reading materials and their comprehension by eight grade students. *Reading Improvement, 23*, 269-276.

Cadenhead, K. (1987). Reading level: A metaphor that shapes practice. *Phi Delta Kappan, 68*, 436-441.

Cassidy, J. (1987). Basals are better. *Learning '87, 16*, 65-66.

Giordano, G. (1987). Determining readability. *Clearing House, 61*, 32-33.

Goodman, K.S. (1987). Taking charge of your own teaching. *Learning '87, 16*, 63-64.

Kaufman, S. (1989). *Accuracy, speed, and comprehension of third grade students' reading performance across four levels of text.* Unpublished master's thesis. University of Florida, Gainesville.

Lindsley, O.R. (1990). Precision teaching: By teachers for children. *Teaching Exceptional Children, 22*, 3, 10-15.

Mills, S. (1988). *Accuracy and speed for fourth grade students' reading performance across six levels of text.* Unpublished master's thesis. University of Florida, Gainesville.

Scott, J. (1988). *Learning rates for oral reading at instructional and frustration levels by elementary pupils with mild learning problems.* Unpublished doctoral dissertation, University of Florida, Gainesville.

Scott, J., Wolking, B., Stoutimore, J. & Harris, C.(1990). Challenging reading. *Teaching Exceptional Children,* 22, 3, 32-35.

Yarington, D. (1979). *The great American reading machine.* Rochelle Park, NJ: Hayden.

William Wolking

Drs. William Wolking is Professor of Special Education at the University of Florida, Gainesville. Carolyn Harris is a doctoral student in the Department of Special Education at the University of Florida, Gainesville. Jack Scott and Jolena Ferro are members of the faculty of the University of South Florida at Tampa.