The Effect of Print Size on Achievement in Mathematics Problem Solving

William H. Nibbelink
Division of Curriculum and Instruction
The University of Iowa
Iowa City, Iowa 52242

Jean A. Gerig
Hoover Elementary School
West Branch, Iowa 52358

Hiram D. Hoover
Department of Psychological and Quantitative Foundations
The University of Iowa
Iowa City, Iowa 52242

It is common for an early elementary mathematics textbook to use an entire pupil's page to present fewer than a half dozen short, easy, traditional story problems and nothing else, often allowing a mathematically adept student named Kim Uhl to complete the page before equally able Henrietta Van Veldhuisen can get past the name blank.

Two obvious barriers to including greater numbers of more interesting problems in early grade mathematics textbooks are the following beliefs held by many: (a) the readability level of story problems should be low, preferably at or below grade level and (b) mathematics textbooks for younger children should use very large type.

Paul, Nibbelink, and Hoover (1986) challenged the first of these widely held beliefs and offered strong evidence that readability level may range considerably without having an effect on pupil achievement in traditional mathematics story problem solving. Indeed, as that study emphasized, such readability tests were never intended for use on narrative passages as short as mathematics story problems usually are.

Assuming that that study provides the freedom to write more interesting problems, the second of the two barriers looms larger than before. Using 18-point type, there is simply too little room on a page to become very interesting. Consider the two versions of a problem given by Figures 1 and 2. Both would require about the same amount of a textbook page. Figure 1 uses 18-point type and Figure 2 uses 12-point type. Figure 2 better captures the pigs' zest for the good life. Notice also that Figure 2 allows for the inclusion of extraneous numerical information, making the problem more a problem than Figure 1.

For the reader who prefers the style of Figure 1, there would still be potential advantages if smaller type sizes could be used, including less cluttered page layout, clearer separation of exercises, more problems per page, more space for art, more space for instruction, and so on. (Keep in mind that number of pages is a very significant factor in textbook cost. Given today's type-size convention, the publisher that decided to provide far more problem solving than is present practice would likely find itself priced out of the market.)

Figure 1. *Problem in 18-point type.*

Nine pigs were in a puddle. Four of the pigs left. How many pigs stayed behind?

Figure 2. *Problem in 12-point type.*

There were nine happy, little pigs
and two big frogs playing in a puddle.
Four of the pigs left to take a nap.
How many pigs stayed behind?

A perusal of today's mathematics textbooks suggests that type size should start at about 18-point for early instruction and then drop about 2 points per year over the first few years of schooling but not much below 12-point (see Figure 3).

Figure 3. *Type sizes.*

18-point type.
14-point type.
12-point type.
10-point type.

Publishers may have little choice on type size if they wish to sell their products. California, for example, proposed at least 18-point type for kindergarten and first grade basic textbooks as recently as 1981 (R.L. Howe, personal communication, August, 1981). Publishers can ill-afford to ignore that market's proposals.

Standardized tests, too, are bound by type-size convention which significantly increases the number of pages needed for earlier grade tests and thus increases the cost of the tests.

Two questions were addressed by this study: (a) how was this type-size convention established? and (b) Is pupil achievement as measured by problem solving in standardized test format enhanced because of it?

Origins and Prior Challenges

About the turn of the century, Shaw (1901) stated, "School life impairs more or less the eyesight of a great number of those who pass through the grades. Our progressive civilization seems destined to tax the eyes to a still greater degree in the future than even at present" (p. 170).

Shaw's answer to this problem was to issue a specific minimum print size for each primary grade level. The school principal was charged with the enforcement of these minimums.

About a decade later, the British Association for the Advancement of Science stated: "The size of the type-face is the most important factor in the influence of books upon vision" (*Report of the BAAS* 1914, pp. 280-281). While it was acknowledged that heredity handed to some a stronger tendency toward myopia than to others, it was also forcefully put forth that myopia was most frequent among the studious. Given that observation, the BAAS sprang to conclusion and issued a chart of specific sizes of types for each age of child, accompanied by a strong statement encouraging the cessation of publication of books that did not comply (*Report of the BAAS* 1914, pp. 281, 288). Apparently not considered upon observing a positive correlation between studiousness and myopia was the alternate cause-effect hypothesis, namely, that myopia causes studiousness due to fewer distant distractions.

Challenges to the large-print-for-small-child convention soon appeared. Gilliland (1923) reported that for speed of reading, "children are not so greatly affected as adults by changes in size of type" (p. 146). (Those over forty resisting a move to trifocals find this statement easy to embrace.)

Technological developments of the 1920s and 1930s allowed researchers to study eye movements and fixations, which were considered to be major factors in reading fluency. Buckingham (1931) reported that, "in spite of the doubling of the size of the type, there is only a slight change in the number of words recognized at each fixation" (p. 104). Uhl (1933) insisted that, "the eye reaches maturity earlier than any other organism in the body", (p. 231) thus diminishing the likelihood that small print in the early grades might negatively influence eye development. (A detailed review of such research on eye movements and fixations is in Weintraub (1977).)

In 1931, the National Society of the Study of Education published a massive study using data gathered from over 2,000 children. The study was financed by three major publishing companies and was undertaken to discover what relationship type size, line length, and spacing had to the readability of text. Considering reading speed alone, 12-point type showed slight advantage over 14-point type, with 18-point type coming in a dismal third (Buckingham, 1931). The study found little in favor of using large type. Buckingham celebrated the case made for smaller type by exclaiming that, "The two great factors of readability and cost are not in opposition but are in reality in alliance" (p. 122). As he put it, "the case seems to be closed." (p. 122) And, closed it was, but not in the way Buckingham had in mind. The large-print-for-small-child convention easily withstood his attack. In the *Thirty-Sixth Yearbook of the National Society for the Study of Education*, Uhl (1937) boldly stated that, "we have evidence to show that earlier beliefs are ill-founded" (p. 233). He claimed that eye stress and eye fatigue among younger children have no direct tie to print size. He, too, made a case for using smaller print but a case having no more impact on practice than had Buckingham's.

No research was found to support the use of large print sizes for smaller children nor does there appear to be much recent interest in the question. A computerized ERIC search on the subject did not reveal any sources. Concern with the topic was found only in books that had not been checked out of the library in over 30 years.

Purpose of The Study

The purpose of this study was 2-fold: (a) to re-visit the question of a type-size effect, this time focusing on traditional problem solving as measured by the format common to nationally standardized achievement tests and (b) to re-visit the question, this time also considering 10-point type, very small type indeed by today's convention.

Research Methodology

Working with the Iowa Testing Programs staff, problem-solving tests were developed and field tested for use in this study. The first-grade test contained 18 items, the second-grade test 25 items, and the third-grade test 25 items. Items were chosen to represent a wide range of mathematical processes and to be grade-level appropriate. All items were presented in multiple-choice format, offering three numerical choices and the answer-not-given choice. Listed below is a sample of items used (Gerig, 1988):

Grade 1, Item 9. Ann caught 9 frogs. She gave 2 of the frogs to her friend and 3 hopped away. How many frogs did she have left? Answer choices: 14, 7, 4, N.

Grade 2, Item 5. Jan walks 8 blocks to school, Mark walks 3 blocks and Anne walks 6 blocks. How many less blocks does Mark walk than Ann? Answer choices: 3, 5, 9, N.

Grade 3, Item 16. Jamie bought 5 apples for 10¢each and a cookie for 30¢. How much money did he have left from $1.00? Answer choices: 20¢, 80¢, $1.00, N.

Three type sizes (18-point, 14-point, and 12-point) were used to derive three forms for the first-grade test. Two type sizes (12-point and 10-point) were used for the second- and third-grade tests. (Using a 10-point version for the first-grade test was considered but then abandoned for fear of negative reactions from teachers or school administrators.)

The tests were randomly assigned to students from Iowa schools who were participating in the Iowa Testing Programs' pilot testing program. (As part of its test development procedure for the *Iowa Test of Basic Skills*, the Iowa Testing Programs routinely collects data on new item sets by sending them to sets of schools representative of the state as a whole. The tests used by this study were included along with other tests in such a pilot run.) The request to the Iowa testing programs was to send out at least 30/of each of the seven forms of the tests.

Beyond the obvious advantage of sampling a rather wide population, administering the tests through the Iowa Testing Programs' pilot testing procedure insured that all tests would be administered in the same way, there being a well-defined protocol for administration.

Results

Table 1 shows the number of students who took each test and the percent correct for each test. Reliability coefficients (KR-20) by grade were 0.78 for first grade, 0.89 for second grade, and 0.83 for third grade.

Table 2 provides evidence that there does not exist a statistically significant influence of type size on problem-solving performance as measured by standardized test format.

Table 1

Grade, Numbers of Students, Type Size, and Percent Correct

Grade	*N*	18 point	*N*	14 point	*N*	12 point	*N*	10 point
1	34	48%	31	58%	31	52%		
2					44	52%	42	55%
3					38	58%	39	58%

Table 2

One-way ANOVA on the Type-size Treatment Effect Using Pupil Mathematics Test Scores by Grade

Grade	Variance	df	SS	*F*	*Pr>F*
1	Type Size	2	58.1314	1.90	0.16
	Error	93	1421.6082		
2	Type Size	1	7.7526	0.20	0.66
	Error	84	3315.0032		
3	Type Size	1	0.9370	0.04	0.84
	Error	75	1794.3097		

Limitations

Possible limitations to wide generalization of the results would include the following:

1. The sample was chosen to be representative of the State of Iowa, a state which consistently scores relatively well on standardized tests and which has a more homogeneous population than the nation as a whole.
2. The results are limited to what can be determined about problem solving from typical standardized test format. (Some educators hold the standardized test format in relatively low regard.)
3. Clearly, the results may not apply to some children who suffer special visual problems.

Discussion

As Buckingham (1931) said it, "the case seems to be closed", (p. 122) now for problem solving as measured by a standardized test format as well as for the variables examined by those earlier challengers of the large-type-for-small-child convention. Furthermore, the case may be closed for even smaller type than was considered earlier.

Given this study and the earlier studies reported, perhaps the more fascinating question is this--Why do we persist in clinging to a type-size convention that apparently offers only unnecessary restriction and increased cost for classroom materials and tests? The advantages to abandoning this convention are clear. Beyond potential advantages already mentioned above are similar cost textbooks that supply more alternatives for students of varying ability, greater opportunity to do more significant things on a single page, and so on.

A companion study to this one used exactly the same methodology to examine the effect of print size on reading comprehension. That study, too, found no effect whatsoever (Gerig, 1988).

Considering this result, along with that of Paul, Nibbelink, and Hoover (1986) which strongly suggested no effect of readability level on problem-solving performance, it would appear that publishers of tests and textbooks should be set free to write both more problems and more interesting problems without a need to increase production costs.

Should today's type-size convention be abandoned at once? Absolutely. Is it likely to be abandoned? Not really. After all, conventions are conventions, and the convention in question easily survived past attacks. There is little reason to believe that it will not as readily survive this one. Then again, who knows which straw it will be that finally breaks the camel's back.

References

Buckingham, B. R. (1931). New data on the typography of textbooks. In *Thirtieth Yearbook of the National Society for the Study of Education* (pp 93-125). Bloomington, IL: Public School Publishing Company.

Gerig, J. A. (1988). *The effects of adjusting print size on reading comprehension and mathematical problem solving performance.* (Doctoral dissertation, University of Iowa, 1988). *Dissertation Abstracts International, 356*, 50/02-A.

Gilliland, A. R. (1923). The effect on reading of changes in the size of type. *The Elementary School Journal, 24*, 138-146.

Paul, D. J., Nibbelink, W. H., & Hoover, H. D. (1986). The effects of adjusting readability on the difficulty of mathematics story problems. *Journal for Research in Mathematics Education, 17*(3), 163-171.

Report of the Eighty-Third Meeting of the British Association for the Advancement of Science. (1914). London: John Murray.

Shaw, E. R. (1901). *School hygiene.* New York: The Macmillan Company.

Uhl, W. L. (1937). The materials of reading. In *Thirty-sixth Yearbook of the National Society for the Study of Education* (pp. 207-253). Bloomington, IL: Public School Publishing Company.

Weintraub, S. (1977). Two significant trends in reading research. In H.A. Robinson (Ed.), *Reading and writing instruction in the United States* (pp. 207-253). Newark, NJ: International Reading Association.

Summary of the SSMA Board of Directors' Meeting
Marquette, Michigan--October 10-11, 1992

All the officers and directors were present, along with six visitors and former officers.

Art White and Donna Berlin reported on several outcomes of the 1991 SSMA Wingspread Conference and distributed an overview of the products and presentations which have been made or are planned at several national and regional conventions.

The Executive Secretary reported the current financial status is stable. The process of selecting a new Executive Secretary was discussed.

The Executive Committee will meet to consider modifying the By-Laws due to changes in the association over the past several years.

About 1,200 people registered for the convention at Marquette. MCTM and SSMA will each realize a profit of $2,000.00.

Plans for the October 14-16, 1993, convention at Alexandria, Louisiana, were discussed. Plans have not been finalized for 1994. Upon a request from Bob Underhill, the Fall 1995 convention location was approved at Charlottesville with the Virginia teachers' organizations.

At the Past-Presidents' Breakfast on Friday, Joe Kennedy received the George G. Mallinson Distinguished Service Award, and both Phil Wagreich and Howard Goldberg were recognized with the Excellence in Integrated Mathematics and Science Award.

Larry Enochs presented the Editor's report. He described the essential role played by Carol Borchers and the associate editors who worked with him at Kansas State University. Bob Underhill reported that the transition to his editorial operation is going smoothly.

Norma Hernandez reported on the *SSMART* newsletter. She, along with James Milson, assumed the production this summer.

The Publications Committee asked to allocate $1,000.00 in the budget for a survey to determine publication needs.

The Membership Committee suggested that members be reminded to state their SSMA affiliation when making professional presentations.

The Spring Board and Publications Committee meeting was set for near O'Hare Airport in Chicago on April 23-25, 1993.

Candidates on the Spring 1993 ballot for President-Elect will be Lloyd Barrow and Patricia Blosser. Candidates for directors will be Terry Goodman, Norbert Kuenzi, William O'Donnell, and Steve Oliver.

The Executive Committee will consist of Jerry Becker, Bob McGinty, Larry Enochs, Darrel Fyffe, Bonnie Litwiller, and Frances Lawrenz.

Members who wish a full set of meeting minutes are asked to contact the Executive Secretary.

Respectfully submitted,

Darrel W. Fyffe
Executive Secretary