Improving the Legibility of Textbooks:
Effects of Wording and Typographic Design

Dirk Wendt

It has been demonstrated that the success achieved in a previous study by means of a "visualized" version of a physics textbook lesson was not so much due to typographic changes as to changes in wording that were necessary to fit the text into the new format. These changes in wording are further analyzed.

In a previous study (Wendt, 1979) we found an improvement in the legibility of a physics textbook lesson through changing the typographic design from a one-column standard design to a "visualized" version, both in speed of reading and in comprehension as measured by means of an achievement test.

This successful (visualized) version, designed by Becker/Heinrich (1969), organized a double page into four columns (i.e., two on each of two opposing pages) containing: (1) basic information on general principles, (2) additional explanation and experimental demonstration, (3) figures, and (4) fringe information — e.g., historical remarks, glossary, practical applications, problems for the student to solve. With a text printed in this format, the student can work vertically or horizontally through the lesson as he pleases, going left to right, from general principle through experimental demonstration, and from picture to practical application, or vice versa, from right to left. To design the visualized version according to these principles, it was necessary to rearrange and change the wording of the original text. Thus, in the experimental design to the previous study, two experimental variables were confounded: typographic design ("visualization" vs. standard format) and wording of the text, such that it was finally impossible to say which variable was responsible for the improvement achieved.

In the study reported here we tried to separate these two variables. We had printed an additional version ("E") which contained the same text as the visualized Version ("B" in the previous study) but in standard layout with two columns on each page. We compared this to the visualized Version ("D") and to the original two-column Version ("B").

Thus we have, in total (including the versions tested only in the 1979 study):

Version A: standard text in standard layout with one column on each page, figures lined up at the top.

Version B: standard text in standard layout with two columns on each page, otherwise like Version A

Version C: standard text in one column on each page but with additional spaces between phrases according to the syntactic and semantic structure of the text

Version D: new text in visualized layout by Becker/Heinrich, as described above

Version E: new text (as in Version D) but in standard layout (as in Version B)

Experimental Procedure

Everything was done in the same manner as in the previous study; we even approached the same secondary schools (but, of course, used different students). We used a total of 346 students in ten classes from grades 5 to 7 (age about 12–14 years) in three secondary schools (Realschulen) in Niedersachsen, F.R. Germany.

The experimental procedure took about 90–100 minutes per session (two regular school lessons plus the recess time between them). The experimenter explained the scope and purpose of the experiment to the students and gave them, then, an achievement test on the contents of the lesson to be read (on magnets) to measure the knowledge they had already before reading the lesson, and some silent demonstrations of experiments described and explained in the text to be read.

Each class was then divided at random into three groups of almost equal size, and all members of each group were given the same version of the experimental lesson (Version B, D, or E), face down on the desk in front of them. At a sign given by the experimenter, they all started reading their text, and raised their hands when they had worked through the text and thought they had learned everything that was in it. They were told that there would be a second achievement test on the text, and that *accuracy* of knowledge was more important than speed of reading. When a student had finished reading and raised his hand, the experimenter or a helper recorded the student's reading time, took away the printed lesson, and handed him the second achievement test, which was collected upon completion. Finally, each student was shown three different versions of the lesson ("A" one-column standard layout, "B" two-column standard layout, "D" visualized layout), and asked to rank-order them according to his preference.

Results

The dependent variables in this study were reading time and the difference in achievement test scores before and after reading the experimental lesson. These were analysed by means of analyses of variance. Mean reading times (in minutes) and mean gain in achievement test scores for the three different versions are shown in Table I, together with the respective standard deviations. Results of the respective analyses of variance are shown in Tables II and III.

They indicate that both experimental Versions (D and E) were read significantly faster, and resulted in significantly larger achievement gains than the original Version (B), but the visualized Version (D) was only slightly faster than the Version (E) with the same text but in standard layout. The difference between the experimental Versions (D and E) both in reading time and in achievement gain is statistically not significant, i.e., it may as well be random. In the new study, average reading times are shorter, but average achievement gains are smaller than in the previous one, although still within the range of variation between classes. As in the previous study, this was the main source of variation. Also, in spite of the general trend indicated by the overall means for the various versions, some classes did better with different versions than others, resulting in interactions between versions and classes which cannot be interpreted systematically since the classes were selected at random. Table IV shows the results of the preference rank-ordering: as in the previous study, an overwhelming majority favored the visualized Version D.

Discussion

Apparently, both the shorter reading time and the better achievement reached with the new Version D in the previous study were due less to typographic improvement and more to changes which had been made to the text to fit it into the new format — a kind of by-product in the previous study. Thus, it seems to make sense to take a closer look at what was actually changed from the original Version (A or B) to the improved Version (D or E). First of all, it had become shorter: 1300 words instead of 1710 in the original text. This accounts for the improved reading time; actually it is almost the same if measured in words per minute rather than by overall reading time: $1710/14.55 = 117.53$ wpm for the original, $1300/12.21 = 106.47$ wpm for the revised Version E, and $1300/11.09 = 117.22$ wpm for the visualized Version D.

However, this does not explain the better achievement gained with the revised version (unless you assume that students learn more with

Table I: Means (M) and standard deviations (SD) of reading time and achievement gained with three different versions

Version	Reading time		Achievement gained	
	M	SD	M	SD
original two columns ("B")	14.55	30.52	1.77	4.73
visualized ("D")	11.09	15.37	2.45	4.01
Text of "D" in layout of "B"	12.21	14.24	2.58	4.65

Table II: ANOVA of reading times

Source of variation	SSQ	df	MS	F	P
between Versions	591.15	2	295.58	21.69	<0.01
between classes	2117.23	9	235.25	17.26	<0.01
interaction	625.01	18	34.72	2.55	<0.01
error	4306.57	316	13.63		

Table III: ANOVA of achievement gain

Source of variation	SSQ	df	MS	F	P
between Versions	26.91	2	13.45	3.65	<0.03
between classes	242.68	9	26.96	7.32	<0.01
interaction	91.41	18	5.08	1.38	~0.14
error	1164.71	316	3.69		

Table IV: Results of preference ordering

Version	Frequency of rank 1	2	3
A, one column per page	33	106	139
B, two columns per page	37	139	100
D, "visualized"	207	30	38

Table V: Evaluation of our texts by Schulz von Thun

	Simplicity	order/ structure	brevity/ conciseness	stimulating addition
original text	+1	+1	0	0
revised text	+1	+2	+1	−1

shorter texts). In rewording the text for the visualized experimental Version (D) in the previous study, we did not try to produce a text which was less understandable than in the original one: we tried our best, and apparently succeeded more than we wanted to, i.e., produced a better one.

The German literature on educational psychology contains an evaluation system for instructional texts, by Langer (1971) and Schulz von Thun (1973). Because this system has proved to be useful in series of experimental studies (Schulz von Thun/Berghes/Langer/Tausch, 1974; Schulz von Thun/Langer/Tausch, 1972), it seemed appropriate for an evaluation of our experimental texts. The system consists, in principle, of ratings on four scales (or "dimensions"): (1) Simplicity *(Einfachheit)*, (2) order/structure *(Gliederung/Ordnung)*, (3) brevity/conciseness *(Kürze/Prägnanz)*, and (4) Stimulating additions *(zusätzliche Stimulanz)*. Higher ratings on the first three of these four scales are assumed to indicate better understandability of the text rated, whereas "stimulating additions" are supposed to have a more motivational effect (too many may distract from the main objective and thus decrease understanding).

Since the training of raters to evaluate texts on these four scales takes some time and effort, and to be certain of securing the ratings intend-

ed by the authors of this system, we asked one of them, Friedemann Schulz von Thun, to evaluate our texts on his scales. The results are shown in Table V. These ratings indicate that our revised text contained fewer "stimulating additions" (it had to because it is shorter!), but had gained not only on the "brevity" dimension but also in "order/structure."

Thus, after all, the improvement found with the new typographic layout (Version D) in our previous study may be due to the rewording of the text which was necessary to fit it into the new format. This, however, does not account for the success of the visualized Version (D) on the students' preference scale, as found in the previous study and reproduced here in Table IV.

REFERENCES

Becker, D./Heinrich, J. 1969. Untersuchung über die Wirksamkeit typographischer Mittel. *Sem. Arb.* HfbK Hamburg.
Langer, I. 1971. Verständnisfördernde Merkmale der sprachlichen Gestaltung von Unterrichtstexten. Phil. Diss. Univ. Hamburg.
Schulz von Thun, F. 1973. Effektivität und Trainierbarkeit von Verständlichkeit bei der schriftlichen Informationsvermittlung. Phil. Tausch, R. 1974 Diss. Univ. Hamburg.
Schulz von Thun, F./v. Berghes/Langer, I./Tausch, R. 1974. Überprüfung einer Theorie der Textverständlichkeit: Verbesserung der Verständlichkeit von Kurzzusammenfassungen wissenschaftlicher Veröffentlichungen. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie (ZEPP)*, 6, No. 3, 197-206.
Schulz v. Thun, F./Langer, I./Tausch, R. 1972. Trainingsprogramm für Pädagogen zur Förderung der Verständlichkeit bei der Wissensvermittlung. Kiel, Landesverband der Volkshochschulen Schleswig-Holstein e.V.
Wendt, D. 1979. An experimental approach to the improvement of the typographic design of textbooks. *Visible Language, 13,* No. 2, 108-33.