HUMAN FACTORS, 1986, 28(2), 159–163

Readability of Computer-Generated Fill-Justified Text

STANLEY R. TROLLIP,[1] *Department of Curriculum and Instruction, University of Minnesota, and* GREGORY SALES, *Department of Curriculum and Instruction, University of Cincinnati*

This paper describes two studies designed to explore the effects of reading fill-justified text on reading speed and comprehension. A common contemporary practice is to use the power of microcomputer-based processors to produce printed material that is fill justified(i.e., both left and right margins are straight). Fill justification is frequently accomplished by inserting varying numbers of extra spaces between words. Both studies compared the reading speed of two groups, one that read a fill-justified passage and the other that read the same passage printed with a ragged right margin. Comprehension was investigated in the one study by comparing performance on a test made up of recognition questions; in the other, performance was compared on recall questions. Results indicate significant increase in reading time (that is, slower reading speed) for groups reading fill-justified text. No differences in comprehension were detected.

INTRODUCTION

With the proliferation of microcomputers both at home and in the workplace, the use of word-processing programs has become very common. Consequently, an increasing amount of reading is done from material printed by these devices. Most of these word processors offer the user a variety of options with respect to what the printed document will look like. These options include boldface or ordinary printing, underlining, different-sized fonts, different types of fonts, and several types of justification.

This study investigates the use of two different types of justification on the readability of printed text and on how well the material is comprehended. The two types of justification are fill justification and left justification.

Fill justification is the process whereby the printed text is positioned to provide even left and right margins, creating a rectangular text display similar to that found in most books or newspapers. To accomplish this, most microcomputer-based word processors insert varying numbers of spaces between the words of the line so that the last character of the last word on each line appears in the appropriate position. (See the top paragraph in Figure 1.) These spaces are the same width as ordinary letters, which results in very noticeable gaps appearing between some words.

Some sophisticated word processors and typesetting machines avoid the large interword gaps through use of micro-justification. That is, they insert thin spaces between both letters and words. This distributes the neces-

[1] Requests for reprints should be sent to Stanley R. Trollip, College of Education, University of Minnesota, 159 Pillsbury Drive SE, Minneapolis, MN 55455.

The Dutch did their best to keep the location of the
pepper-growing region a secret, until in 1795 the secret was
pierced by a Yankee skipper from Salem, Massachusetts, who
brought a cargo back that made 700 percent profit on the voyage.
He couldn't keep the secret either, and soon whole fleets were
braving the reefs and pirates of Sumatra to bring tons of pepper
back to Massachusetts. It was one of the vital steps in
establishing a preeminence of the merchant marine of the young
Republic.

The Dutch did their best to keep the location of the
pepper-growing region a secret, until in 1795 the secret was
pierced by a Yankee skipper from Salem, Massachusetts, who brought
a cargo back that made 700 percent profit on the voyage. He
couldn't keep the secret either, and soon whole fleets were
braving the reefs and pirates of Sumatra to bring tons of pepper
back to Massachusetts. It was one of the vital steps in
establishing a preeminence of the merchant marine of the young
Republic.

Figure 1. *Examples of fill-justified (top) and left-justified text (bottom).*

sary spacing evenly throughout the line, causing differences in spacing to be barely discernible. Other processors justify the text by the insertion of thin spaces only between words. Both techniques result in the spacing between words being quite consistent within each line. Because the inserted spacing is very thin, the interword spacing from line to line also appears consistent, even though it may be slightly different.

The alternative to fill justification is left justification, which results in a ragged right margin. In this style, the position of the last printed character on a line depends only on the length of the line (bottom paragraph, Figure 1). In extreme cases of raggedness, resulting from the use of long words, hyphenation can be used to minimize the discrepancies between the ends of lines.

As can be seen from the two printed versions illustrated in Figure 1, the spacing between words with fill justification can vary considerably, whereas the position of the last character of each lines varies with left justification. Note that both methods result in the same number of words on a line.

The most common reason given for using fill justification is that it looks better to have straight left and right margins. This opinion is often reinforced by the producers of word-processing packages. The introductory message to Wordstar™ (Micropro International Corp.), for example, appears on the screen in fill-justified form.

Despite the apparent aesthetic quality of fill-justified text, we believe that the variable spacing between words necessary to accomplish this "neatness" disrupts reading flow. We hypothesize that one result of these disruptions will be to reduce the reading speed. We also hypothesize that these disruptions will divert part of the reader's attention from comprehending the text, and consequently will result in poorer comprehension.

To test these hypotheses, we conducted two experiments.

METHOD

Experiment 1

The purpose of Experiment 1 was to compare the reading speed of two groups, one

(Group F1) reading fill-justified text and the other (Group L1) reading left-justified text, and to assess whether there was any difference between groups in the scores on a series of recall questions.

Subjects volunteering for the study were enrolled in two introductory computer classes in the College of Education that were open to both advanced undergraduates and beginning graduate students. Students were assigned at random to one of the two groups. Group F1 consisted of 21 subjects (14 female and 7 male, with an average age of 28.3). The second group (Group L1) comprised 25 subjects (15 female and 10 male, average age 30.8).

All subjects read an extended passage taken from an article in the *Smithsonian* on the history of pepper (Wernick, 1983). The article was processed on the Applewriter IIe wordprocessor and was printed using a Pica 10 font at 10 characters per 2.54 cm, 66 characters per line, on a Diablo 640 daisy-wheel printer. The passage was printed with 23 double-spaced lines per page and was 14.5 pages long.

At the beginning of each session, the experimenter explained the procedures for the study. Subjects had to record the time at which they started reading and the time that they finished the passage. Timing did not include the answering of questions. Subjects were not permitted to turn back to the body of the text once they started answering questions.

On completing the reading, both groups answered 15 recall (fill-in-the-blank) questions covering factual information contained in the passage.

Experiment 2

Experiment 2 was conducted in an identical manner to Experiment 1, except that the 15 questions asked at the end of the reading tested recognition, not recall. All of the questions were multiple-choice. Subjects were all volunteers from an undergraduate rhetoric class and were assigned at random to the two groups.

The group reading the fill-justified text in this study was labeled F2, and consisted of 46 subjects (35 females, 11 males, average age 23.6). The group reading the left-justified text, L2, consisted of 41 subjects (31 females, 10 males, average age 22.8).

RESULTS

Experiment 1

Group F1 had a mean time to completion of 918.8 s with a mean score of 7.7 out of 15 on the test. Group L1 completed the reading in an average time of 815.2 s with a score of 7.2. Figure 2 illustrates the difference in time. Table 1 contains all the data discussed in this section.

Performance of the two groups with respect to time taken to complete the reading and the score on the posttest was compared using independent t tests.

In Experiment 1, time to completion of the reading for Group F1 was significantly longer, $t(44) = 2.107$; $p < 0.05$, than for Group L1. Difference in posttest scores was not significant.

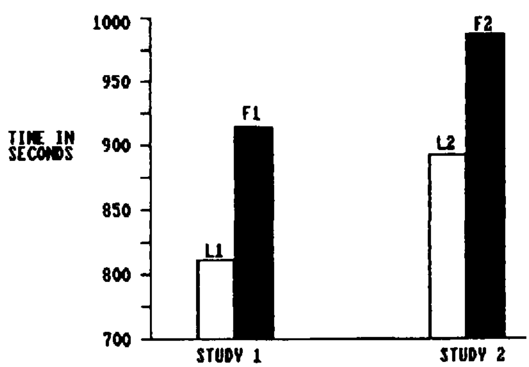

Figure 2. *Times to completion of reading.*

TABLE 1

Reading Times and Posttest Scores as a Function of Justification Condition and Type of Question

		Question Type	
		Recall	*Recognition*
		Group F1	*Group F2*
Fill justification	Time	918.8	990.5
	Score	7.7	9.8
		Group L1	*Group L2*
Left justification	Time	815.2	899.0
	Score	7.2	9.5

Experiment 2

Group F2 had a mean time to completion of 990.5 s with a mean score of 9.8 out of 15 on the test. Group L2 completed the reading in an average time of 899.0 s with a score of 9.5. See Figure 2.

Performance of the two groups with respect to time taken to complete the reading and the score on the posttest was compared using independent t tests.

The fill-justified group (F2) took significantly longer than the group (L2) reading the left-justified text, $t(85) = 2.228$; $p < 0.05$. Once again, there was no significant difference in the posttest scores.

As one means of checking the quality of our questions, we compared performance of the two fill-justified groups with each other, and the two ragged groups with each other. Independent t tests comparing F1 with F2 and L1 with L2 yielded significant differences for performance on the posttest, $t(65) = 3.218$, $p < 0.005$, and $t(64) = 3.307$, $p < 0.005$. This was expected because the first study required subjects to recall answers to questions and the second provided recognition-type questions. The reading times for the groups were not significantly different.

DISCUSSION

The results of these two studies indicate that fill-justified output slows reading speed (increases reading time) significantly. Although no direct reference to this issue could be found (in reviews such as Rayner, 1978), we believe that there are several plausible explanations for this reduced reading rate.

The first concerns the amount of information in the reader's field of vision. Morrison and Rayner (1981) suggest that the number of characters in a fixation remains constant. That is, if the same text is moved closer to or further from the reader, the number of characters in a fixation remains unchanged. The same situation occurs if letters of different sizes are used.

If one makes the assumption that spaces constitute characters, then in order to read the whole text, the reader will have to make more fixations with fill-justified text than with left-justified. This explanation can be easily tested by reproducing the text with more than one space consistently separating words.

A second explanation concerns the variability of spacing between words. The reduced reading rate may occur because the eye has to adjust continually to where the next word starts. That is, a decision has to be made whenever a space is encountered as to where the following word starts. In a study using variable spacing between words, Abrams and Zuber (1972) showed that fixation time immediately prior to a positioning decision (interword space, end of line, or correction at end-of-line return sweep) was substantially shorter (about 180 ms) than the average fixation time during normal reading (about 255 ms). Unfortunately it is not clear from the study whether the interspersed spaces cause additional fixations as well as changing their duration.

The third explanation is based on a disruption theory. Just and Carpenter (1980) have proposed a reading comprehension model that includes a "sentence wrap-up" phase during which the reader attempts to resolve any inconsistencies within the sentence, such as hanging or unassigned referents. If the reader's normal fluency were disrupted by interpreting longer spaces as sentence endings, then this processing would occur, finding many inconsistencies because the sentences or ideas would not be complete most of the time. The resolution of these inconsistencies would take additional time, largely due to an increase in the number of regressions (right-left movements of the eye) required to reassess the content of the text (Carpenter and Just, 1978).

In a sense, the inconsistency of spacing may operate in the same way as spelling errors, which also cause a disruption in fluency. Zola (1984) found that introducing spelling errors into text resulted in an increase in the fixation time at the error, and an increase in the number of regressions (presumably to resolve the error). Both of these factors would increase the reading time.

Of the explanations offered, we tend to support the latter based on our own introspection and on informal data gathered from conversations with subjects and others who had just completed reading fill-justified text.

We had originally hypothesized that comprehension would be lower when reading fill-justified text, due to what we believed to be its disruptive nature—that attention would be diverted from the content to the process of removing ambiguity caused by incorrectly perceiving ends of sentences. We were unable to support this hypothesis. Indeed, our results show minor increases in posttest scores for subjects in the fill-justified treatments. Several of our colleagues have suggested that the hypothesized rereading or reprocessing of the material may actually lead to better comprehension. Further work needs to be done to explore this issue more thoroughly.

It is interesting to note that the groups with an older average age took substantially shorter time to complete the reading, although the differences were not statistically significant. Group F1 averaged 918.8 s and Group F2 averaged 990.5 s. The average ages for F1 and F2 were 28.3 and 23.6 respectively. The mean time to completion of L1 was 815.2 s and that of L2 899.0 s. The mean ages for the two groups were 30.8 and 22.8 respectively. We have no explanation for this difference.

CONCLUSIONS

The obvious recommendation arising from the results of this study is that fill justification should not be used if the justification is accomplished by the insertion of whole spaces. Although it appears that comprehension as measured by recall and recognition questions is not affected, reading time will be significantly increased if text is printed in fill-justified form.

REFERENCES

Abrams, S. G., and Zuber, B. L. (1972). Some temporal characteristics of information processing during reading. *Reading Research Quarterly, 8*(1), 41-51.

Carpenter, P. A., and Just, M. A. (1978). Cognitive processes in reading: Models based on readers' eye fixations. In A. M. Lesgold and C. A. Perfetti (Eds.), *Interactive processes in reading.* Hillsdale, NJ: Erlbaum.

Just, M. A., and Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87,* 329-354.

Morrison, R. E., and Rayner, K. (1981). Saccade size in reading depends on character spaces and not visual angle. *Perception and Psychophysics, 30,* 395-396.

Rayner, K. (1978). Eye movements in reading and information processing. *Psychological Bulletin, 85,* 618-660.

Wernick, R. (1984, February). Men launched 1000 ships in search of the dark condiment. *Smithsonian,* pp. 128-148.

Zola, D. (1984). Redundancy and word perception during reading. *Perception and Psychophysics, 36,* 277-284.