The effects of changes in layout and changes in wording on preferences for instructional text

James Hartley and Mark Trueman

Abstract

Judges rated their preferences for pages of instructional text which varied in terms of their layout, their wording or both of these features. Three different methods of assessing preferences were used. The results suggested that these measures could provide useful rough quantitative data but that they could not be relied upon too greatly when fine judgement was required.

Introduction

A considerable literature now documents how typographical layout can be manipulated to improve the comprehension of instructional text. The evidence shows that readers can search well-designed text more quickly than standard text, that readers prefer well-spaced text to standard text, and that spacing can be used to convey the underlying structure of the text to the reader (e.g. see Hartley, 1978a, b, 1980a, b, Hartley, et al., 1980).

Similarly a considerable literature now documents how text can be re-written in order to make it easier to understand. Experiments have been carried out to show, bearing in mind certain qualifications, that more readable text is read for a greater length of time and with greater understanding than is less readable text (e.g. see Klare, 1976; Hartley, 1980a, 1981b).

When it comes to methods of evaluating changes to text, however, no investigator to our knowledge has sought to partial out the contributions that changes in layout and changes in readability can make to the effectiveness of a particular text. In this paper we have tried to do this by asking readers for their preferences for versions of text which varied in terms of their layout, their readability, or both of these features.
Reader preferences have traditionally received a ‘hostile press’ in typographic research. Hartley (1978a) wrote: “Subjective preferences are interesting but they are not always informative. Not only does one man’s meat appear to be another man’s poison, but also one cannot assume that there will always be a positive relationship between preference, speed of retrieval and ease of use”. And in 1978b he said: “This measure provides additional information but it often seems to be of dubious validity for subjective preferences are usually based on uninformed judgement.”

Today Hartley has somewhat modified these views. In some of his more recent research he has come to rely a good deal on the use of preference measures (e.g. see Hartley, 1980c, 1981a; Hartley et al, 1979; Hartley and Guile, 1981). This research has shown that preference data can be quite sensitive (i) to differences between expert and non-expert judges, and (ii) to the effects of training in using documents with different layouts. Furthermore, one can argue that, although from one particular point of view a judgement might be ‘uninformed’, such a judgement might well affect a person’s initial reaction to a text, and, indeed, determine whether or not that text will be purchased or even read.

This paper explores this last consideration by asking people for their initial reaction to pages of instructional text which vary systematically in their presentation. We planned to use preference measures to enable us to assess the relative contribution of changes in layout and changes in wording to the perceived effectiveness of text. As we shall see, however, these plans were not quite as easy to realise as we expected, and thus this paper reveals some of the problems and pitfalls of using subjective preferences as a dependent measure. Five experiments are described which explore three different ways of obtaining preferences. In each of these experiments an equal number of men and women took part and the results were first analysed for sex differences. Since none were found the overall results are presented in the following account.

**Experiment 1**

Four versions of a page of instructional text were prepared as follows (see Figures 1a, b, c and d):
Version 1 Original Text: Original Layout
Version 2 Original Text: Revised Layout
Version 3 Revised Text: Original Layout
Version 4 Revised Text: Revised Layout
INSULATING RUBBER BLANKETS
DESCRIPTION, MAINTENANCE,
AND INSPECTION

1. GENERAL

1.01 The care, maintenance, and inspection of the insulating rubber blankets are described in this section.

1.02 This section is reissued to delete reference to the KS-16302 cleaner which has been superseded by the B cleaning fluid (AT-8236).

1.03 Insulating blankets are for use as a temporary insulating wrapping on poles which may come in contact with power lines during construction work. The blanket is also for use as an insulating mat on which a workman must stand while operating external derrick controls for a derrick being used in the vicinity of power lines. The use of insulating blankets is described in Section 621-205-010.

1.04 The insulating qualities of blankets are reduced when they become wet. For this reason, insulating blankets shall not be used during periods of rain or to cover pockets of water on the ground.

1.05 The insulating blanket is not a substitute for insulating gloves. Insulating gloves shall always be worn in conjunction with the use of the blanket.

1.06 When using an insulating blanket as a mat, care must be taken not to place it directly on sharp gravel, glass, or other sharp objects which will cause cuts. Either sweep the area to remove such objects or place boards to protect the blanket.
INSULATING RUBBER BLANKETS

Description, Maintenance, and Inspection

1.0 General

1.1 The care maintenance and inspection of the insulating rubber blankets are described in this section.

1.2 This section is reissued to delete reference to the KS-16302 cleaner which has been superseded by the B cleaning fluid (AT-8236).

1.3 Insulating blankets are for use as a temporary insulating wrapping on poles which may come in contact with power lines during construction work. The blanket is also for use as an insulating mat on which a workman must stand while operating external derrick controls for a derrick being used in the vicinity of power lines. The use of insulating blankets is described in Section 621-205-010.

1.4 The insulating qualities of blankets are reduced when they become wet. For this reason insulating blankets shall not be used during periods of rain or to cover pockets of water on the ground.

1.5 The insulating blanket is not a substitute for insulating gloves. Insulating gloves shall always be worn in conjunction with the use of the blanket.

1.6 When using an insulating blanket as a mat, care must be taken not to place it directly on sharp gravel, glass or other sharp objects which will cause cuts. Either sweep the area to remove such objects or place boards to protect the blanket.

Figure 1b. Version 2 Original text : Revised layout
INSULATING RUBBER BLANKETS
DESCRIPTION, MAINTENANCE
AND INSPECTION

1. GENERAL

1.01 This section describes the care, maintenance and
inspection of insulating rubber blankets.

1.02 This section is re-issued to delete reference to the
KS-13602 cleaner; this has been superseded by the
B cleaning fluid (AT-8236).

1.03 Insulating blankets are used to provide temporary
insulation around poles that might come into
contact with power lines during construction work. The
blankets are also used as insulation mats for workmen
to stand on when they are operating the external
controls of a derrick near power lines. The use of
insulating blankets is described in Section 621-205-010.

1.04 The insulating quality of the blanket is reduced
when it gets wet. For this reason do not use
insulating blankets to cover pools of water on the
ground, or when it is raining.

1.05 An insulating blanket is not a substitute for in-
sulating gloves. Always wear insulating gloves
when using an insulating blanket.

1.06 When using the blanket as a mat take care not to
place it directly on sharp gravel, glass or other
sharp objects which might damage it. Either sweep the
area to remove such objects, or put down boards to
protect the blanket.
INSULATING RUBBER BLANKETS

Description, Maintenance and Inspection

1.0 General

1.1 This section describes the care, maintenance and inspection of insulating rubber blankets.

1.2 This section is re-issued to delete reference to the KS-13602 cleaner; this has been superseded by the B cleaning fluid (AT-8236).

1.3 Insulating blankets are used to provide temporary insulation around poles that might come into contact with power lines during construction work. The blankets are also used as insulation mats for workmen to stand on when they are operating the external controls of a derrick near power lines. The use of insulating blankets is described in Section 621-205-010.

1.4 The insulating quality of the blanket is reduced when it gets wet. For this reason do not use insulating blankets to cover pools of water on the ground, or when it is raining.

1.5 An insulating blanket is not a substitute for insulating gloves. Always wear insulating gloves when using an insulating blanket.

1.6 When using the blanket as a mat take care not to place it directly on sharp gravel, glass or other sharp objects which might damage it. Either sweep the area to remove such objects, or put down boards to protect the blanket.
One hundred mature Open University students\(^1\) attending a summer school at the University of Keele were each asked individually to compare two of the above versions (e.g. Version 1 with Version 2) and to state which one they preferred. The actual instructions asked them to consider "Which one you prefer, which one do you think the easiest to use, to understand, which one do you think is the clearest? When you have done this can you then give each one a mark out of ten for clarity." The marks allocated recorded, and sufficient data was collected to allow us to make the following comparisons:

**Comparisons focusing on layout changes**
- Version 1 versus Version 2 (N = 20)
- Version 3 versus Version 4 (N = 20)

**Comparisons focusing on text changes**
- Version 1 versus Version 3 (N = 20)
- Version 2 versus Version 4 (N = 20)

**Comparisons focusing on both**
- Version 1 versus Version 4 (N = 20)

The results obtained are displayed in Table I. The comparisons were made using the Wilcoxon Matched Pairs Signed Ranks test (Seigel, 1956). This test gives a T value, and a z score. The z scores and their significance levels (for one-tailed tests) are reported in Table I. In almost every case changes in layout led to significantly greater preferences and, similarly, changes in wording led to significantly greater preferences. Combining the changes, however, did not lead to a wider disparity in the marks allocated: indeed the results were very similar.

---

\(^1\) Open University students are very different from conventional British University students. They have much more varied backgrounds and range in age from their 20s to their 80s (see Mackintosh, 1974). As such they are more typical of the normal population than conventional university students, although clearly they are exceptional in many ways.
Table I
Medians and ranges of the marks given out of ten for each version in each comparison (Experiment 1).

<table>
<thead>
<tr>
<th></th>
<th>Median</th>
<th>Range</th>
<th></th>
<th>Median</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Layout changes</strong></td>
<td></td>
<td></td>
<td><strong>Text changes</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Version 1</td>
<td>5.5</td>
<td>2–10</td>
<td>Version 3</td>
<td>6.0</td>
<td>2–8</td>
</tr>
<tr>
<td>Version 2</td>
<td>8.0</td>
<td>4–10</td>
<td>Version 4</td>
<td>8.0</td>
<td>4–10</td>
</tr>
<tr>
<td>z value</td>
<td>1.34</td>
<td></td>
<td>z value</td>
<td>3.02</td>
<td></td>
</tr>
<tr>
<td>Significance</td>
<td>n.s.d.</td>
<td></td>
<td>Significance</td>
<td></td>
<td></td>
</tr>
<tr>
<td>level</td>
<td></td>
<td></td>
<td>level</td>
<td>p&lt;.005</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Median</th>
<th>Range</th>
<th></th>
<th>Median</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Version 1</td>
<td>5.0</td>
<td>2–8</td>
<td>Version 2</td>
<td>6.0</td>
<td>4–8</td>
</tr>
<tr>
<td>Version 3</td>
<td>7.0</td>
<td>4–9</td>
<td>Version 4</td>
<td>8.0</td>
<td>4–10</td>
</tr>
<tr>
<td>z value</td>
<td>2.86</td>
<td></td>
<td>z value</td>
<td>3.12</td>
<td></td>
</tr>
<tr>
<td>Significance</td>
<td>p&lt;.005</td>
<td></td>
<td>Significance</td>
<td></td>
<td></td>
</tr>
<tr>
<td>level</td>
<td></td>
<td></td>
<td>level</td>
<td>p&lt;.001</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Median</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Both changes</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Version 1</td>
<td>6.0</td>
<td>2–10</td>
</tr>
<tr>
<td>Version 4</td>
<td>8.0</td>
<td>4–10</td>
</tr>
<tr>
<td>z value</td>
<td>2.95</td>
<td></td>
</tr>
<tr>
<td>Significance</td>
<td>p&lt;.005</td>
<td></td>
</tr>
<tr>
<td>level</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Experiment 2**

Experiment 2 was a replication of Experiment 1 (and was carried out before the results from Experiment 1 were analysed) using a different text. Again four versions of this text were prepared in the same way (see Figures 2a, b, c and d):

Version 1 Original Text: Original Layout
Version 2 Original Text: Revised Layout
Version 3 Revised Text: Original Layout
Version 4 Revised Text: Revised Layout

The main difference between the texts used in Experiments 1 and 2 lay in the nature of the changes made. In Experiment 1 the main changes were typographical: in Experiment 2 the main changes were textual. The original version of the text used in Experiment 1 had a Gunning Fog Index of 14.7 and for the revised version it was 13.5. The original version of the text used in Experiment 2 had a Fog Index of 11.0 and for the revised version it was 8.6.
Table II

Medians and ranges of the marks given out of ten for each version in each comparison (Experiment 2).

### Layout changes

<table>
<thead>
<tr>
<th>Version</th>
<th>Median</th>
<th>Range</th>
<th>Version</th>
<th>Median</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6.0</td>
<td>3–9</td>
<td>3</td>
<td>5.0</td>
<td>3–9</td>
</tr>
<tr>
<td>2</td>
<td>8.0</td>
<td>4–10</td>
<td>4</td>
<td>8.0</td>
<td>2–9</td>
</tr>
<tr>
<td>z value</td>
<td>2.47</td>
<td></td>
<td>z value</td>
<td>1.96</td>
<td></td>
</tr>
<tr>
<td>Significance level</td>
<td>p&lt;.01</td>
<td></td>
<td>Significance level</td>
<td>p&lt;.05</td>
<td></td>
</tr>
</tbody>
</table>

### Text changes

<table>
<thead>
<tr>
<th>Version</th>
<th>Median</th>
<th>Range</th>
<th>Version</th>
<th>Median</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6.5</td>
<td>2–9</td>
<td>2</td>
<td>5.0</td>
<td>2–9</td>
</tr>
<tr>
<td>3</td>
<td>8.0</td>
<td>4–9</td>
<td>4</td>
<td>8.0</td>
<td>5–10</td>
</tr>
<tr>
<td>z value</td>
<td>1.51</td>
<td></td>
<td>z value</td>
<td>2.97</td>
<td></td>
</tr>
<tr>
<td>Significance level</td>
<td>n.s.d.</td>
<td></td>
<td>Significance level</td>
<td>p&lt;.005</td>
<td></td>
</tr>
</tbody>
</table>

### Both changes

<table>
<thead>
<tr>
<th>Version</th>
<th>Median</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6.0</td>
<td>2–8</td>
</tr>
<tr>
<td>4</td>
<td>8.0</td>
<td>6–10</td>
</tr>
<tr>
<td>z value</td>
<td>3.55</td>
<td></td>
</tr>
<tr>
<td>Significance level</td>
<td>p&lt;.0005</td>
<td></td>
</tr>
</tbody>
</table>

A further one hundred Open University students (50 men and 50 women) took part in this Experiment as in Experiment 1. The results obtained are shown in Table II. These results reflect almost exactly those obtained in Experiment 1.

The results from both of these experiments suggest (i) that people prefer revised versions to the original ones, whether the revisions are to the layout, to the text, or to both, but (ii) that the measure made is not sensitive to the number and to the kind of changes made. It seems that most students gave themselves a baseline of about 5 marks for the version they liked least and a ceiling of about 8 marks for the version they liked most, irrespective of the versions being compared. Thus versions 2 and 3 for example scored about 5 when they were presented as original texts and about 8 when they were presented as revisions. In short, (although clearly the revised
IMPORTANT INFORMATION FOR OUR PASSENGERS

Even though you may be an experienced air traveler, there are certain features of this airplane with which you may not be familiar.

AUTOMATIC OXYGEN SYSTEM

The higher altitudes at which this aircraft operates require the prompt use of the automatic oxygen system in case of any sudden change in cabin pressure. Should a decompression occur, oxygen masks will drop down. Take nearest mask and promptly place over nose and mouth. BREATHE NORMALLY (NO SMOKING PLEASE).

SEAT BELTS

Even if the "SEAT BELT" sign is turned off in flight, it is recommended that you keep your seat belt fastened, whenever you are in your seat.

FLOTATION SEAT CUSHIONS

The cushion on which you are sitting is designed to keep you afloat. In the event of a water landing, grasp the cushion at the rear, pull it forward and take it with you.

EMERGENCY EXITS

There are nine exits provided for your use. The chart below will show you the one closest to your seat. The exits over the wings are removable windows. For easy access to the window, push seat back ahead of the window forward. The two exits at each end of the cabin are doors equipped with fast operating evacuation slides. There is also a door in the rear of passenger cabin. REAR CABIN EXIT (STAIR). (If usable, will be opened by a crew member.)
IMPORTANT INFORMATION FOR OUR PASSENGERS

Even though you may be an experienced air traveler, there are certain features of this airplane with which you may not be familiar.

AUTOMATIC OXYGEN SYSTEM

The higher altitudes at which this aircraft operates require the prompt use of the automatic oxygen system in case of any sudden change in cabin pressure. Should a decompression occur, oxygen masks will drop down. Take nearest mask and promptly place over nose and mouth. BREATHE NORMALLY (NO SMOKING PLEASE).

SEAT BELTS

Even if the "SEAT BELT" sign is turned off in flight, it is recommended that you keep your seat belt fastened, whenever you are in your seat.

FLOTATION SEAT CUSHIONS

The cushion on which you are sitting is designed to keep you afloat. In the event of a water landing, grasp the cushion at the rear, pull it forward and take it with you.

EMERGENCY EXITS

There are nine exits provided for your use.

The chart below will show you the one closest to your seat.

The exits over the wings are removable windows. For easy access to the window, push seat back ahead of the window forward.

The two exits at each end of the cabin are doors equipped with fast operating evacuation slides.

There is also a door in the rear of passenger cabin. REAR CABIN EXIT (STAIR). (If usable, will be opened by a crew member.)
IMPORTANT!

This aircraft has special safety features. Read this card carefully.

AUTOMATIC OXYGEN

If, during the flight, there is a sudden change in cabin pressure, oxygen masks will drop down automatically. If this happens, take the nearest mask, put it quickly over your nose and mouth, breathe normally, put out all cigarettes.

SEAT BELTS

We suggest that you keep your seat belt fastened when you are seated - even when the SEAT BELT sign is turned off.

FLOATING SEAT CUSHIONS

Your seat cushion will keep you afloat if we make an emergency landing in the sea. Get hold of the cushion at the back, pull it forward, and take it with you.

EMERGENCY EXITS

There are nine emergency exits. The chart on the back of this card shows the exit nearest to your seat. The two exit doors at the end of the cabin are fitted with chutes for sliding down. To get out over the wings you have to take out the windows. To make this easier, put the seat-back down when you are trying to get to the window. The door at the back of the cabin is labelled REAR CABIN EXIT (STAIR). This door will be opened by a crew member.

---

Figure 2c. Version 3 Revised text : Original layout
IMPORTANT!

This aircraft has special safety features.
Read this card carefully.

AUTOMATIC OXYGEN

If, during the flight, there is a sudden change in cabin pressure, oxygen masks will drop down automatically.
If this happens
- take the nearest mask
- put it quickly over your nose and mouth
- breathe normally
- put out all cigarettes.

SEAT BELTS

We suggest that you keep your seat belt fastened when you are seated - even when the SEAT BELT sign is turned off.

FLOATING SEAT CUSHIONS

Your seat cushion will keep you afloat if we make an emergency landing in the sea.
Get hold of the cushion at the back, pull it forward, and take it with you.

EMERGENCY EXITS

There are nine emergency exits.
The chart on the back of this card shows the exit nearest to your seat.
The two exit doors at the end of the cabin are fitted with chutes for sliding down.
To get out over the wings you have to take out the windows.
To make this easier, put the seat-back down when you are trying to get to the window.
The door at the back of the cabin is labelled REAR CABIN EXIT (STAIR).
This door will be opened by a crew member.
versions were preferred to the originals) these results reflect more the problems of subjective scaling than specific differences in the texts (Poulton, 1973).

Such findings, unfortunately, prevent us from saying anything about the relative effectiveness of the layout or textual changes. To make our measure more sensitive to these differences we tried a new approach.

Experiment 3

Experiment 3 replicated Experiment 2 except that in this case each student was given three versions of the airline safety document to compare, and asked to put them in order of preference. The versions used were Versions 1, 2, and 4. After this each judge was asked to suppose that Version 1 had already been given a mark of 5 out of 10 for clarity, and thus to indicate to the investigator what mark he or she would allocate to Versions 2 and 4. A further twenty Open University students took part in this enquiry, 10 men and 10 women.

The results obtained are shown in Table III. These results suggest that the layout changes have little effect relative to the changes in wording. However, this could reflect the fact that Version 3 was not included in this experiment. Consequently Experiment 3 was repeated with 20 more participants using Versions 1, 3 and 4. The results obtained are shown in Table IV.

These results replicate almost exactly those shown in Table III. Compared with the results obtained in Experiments 1 and 2 both of these results now suggest that a double change is seen as more preferable to a single one, but that once again, the data from Tables III and IV suggest that they reflect the nature of the task rather than specific differences between the texts. One point of interest, however, is that in both Experiments 2 and 3 the comparison of Version 1 with Version 3 produced a non-significant result.
### Table III
The median rankings given to Versions 2 and 4 (with a standard mark of 5 given to Version 1).

<table>
<thead>
<tr>
<th>Version</th>
<th>Median</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Version 1</td>
<td>5.0</td>
<td>0</td>
</tr>
<tr>
<td>Version 2</td>
<td>6.0</td>
<td>2-8</td>
</tr>
<tr>
<td>Version 4</td>
<td>8.0</td>
<td>1-10</td>
</tr>
</tbody>
</table>

Significance levels:
- Version 1 versus Version 2: $z=2.64$, $p<.005$
- Version 1 versus Version 4: $z=3.16$, $p<.001$
- Version 2 versus Version 4: $z=2.88$, $p<.005$

### Table IV
The median rankings given to Versions 3 and 4 (with a standard mark of 5 given to Version 1).

<table>
<thead>
<tr>
<th>Version</th>
<th>Median</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Version 1</td>
<td>5.0</td>
<td>0</td>
</tr>
<tr>
<td>Version 3</td>
<td>6.0</td>
<td>3-7</td>
</tr>
<tr>
<td>Version 4</td>
<td>8.0</td>
<td>6-10</td>
</tr>
</tbody>
</table>

Significance levels:
- Version 1 versus Version 3: $z=1.25$, not significant
- Version 1 versus Version 4: $z=3.92$, $p<.0001$
- Version 3 versus Version 4: $z=3.92$, $p<.0001$

### Experiment 4
So far, using mature Open University students as participants, we have seen that the method of comparing two versions discriminates between the pairs, but does not discriminate between the number and kinds of changes made. The method of comparing three versions, however, does discriminate between the number of changes, but it does not appear to discriminate between the kinds of changes (with these participants). In order to see if textual changes were preferred to layout ones (or vice versa) we next decided to ask students to compare all four versions of a particular passage and to place them in rank order. In this experiment eighty students from Bristol Polytechnic (aged between 18 and 21) acted as participants. In the first part of the experiment forty students (20 men and 20 women) used the passage on insulating blankets, and in the second part forty students (20 men and 20 women) used the passage on airline safety.
Each student was asked individually to place the four versions of one of the passages in rank order of preference. The versions were labelled p, t, d and h to remove any order effect that might be implied from labelling them a, b, c, d, or 1, 2, 3, 4. After the various versions had been placed in order, the students were asked to suppose that Version p (i.e. 1) had already been given a mark of 5 out of 10 for clarity, and to indicate to the experimenter what mark he or she would allocate to versions t, d and h (i.e. 2, 3 and 4).

The results obtained for the insulating blankets passage are shown in Table V, and for the airline safety passage in Table VI. It can be seen with the insulating blankets passage that Version 1 was significantly less-preferred to Versions 2, 3, and 4 but that there were no significant differences between the preference rankings for the latter versions. The results from the airline safety passages, however, were more striking. Table VI shows that Version 1 was the least preferred and Version 4 the most preferred. There were no significant preferences between Versions 2 and 3, although both were significantly preferred to Version 1 and significantly less preferred than Version 4.

### Table V
The median rankings given to Versions 1, 2, 3 and 4 of the passage on insulating blankets by Bristol Polytechnic students (N=40) with a standard mark of 5 given to Version 1.

<table>
<thead>
<tr>
<th>Version</th>
<th>Median</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5.0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>7.0</td>
<td>2–9</td>
</tr>
<tr>
<td>3</td>
<td>6.0</td>
<td>1–10</td>
</tr>
<tr>
<td>4</td>
<td>7.0</td>
<td>3–9</td>
</tr>
</tbody>
</table>

Significance levels:
- Version 1 versus Version 3 \( z=3.19 \) \( p<.001 \)
- Version 2 versus Version 3 \( z=0.03 \) not significant
- Version 2 versus Version 4 \( z=1.01 \) not significant
- Version 3 versus Version 4 \( z=1.07 \) not significant
Table VI
The median rankings given to Versions 1, 2, 3 and 4 of the passage on airline safety by Bristol Polytechnic students (N=40) with a standard mark of 5 given to Version 1.

<table>
<thead>
<tr>
<th>Version</th>
<th>Median</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5.0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>6.0</td>
<td>3– 9</td>
</tr>
<tr>
<td>3</td>
<td>6.0</td>
<td>1– 9</td>
</tr>
<tr>
<td>4</td>
<td>8.0</td>
<td>1–10</td>
</tr>
</tbody>
</table>

Significance levels:
- Version 1 versus Version 2 $z=4.11$ p<.001
- Version 1 versus Version 3 $z=2.46$ p<.01
- Version 2 versus Version 3 $z=1.38$ not significant
- Version 2 versus Version 4 $z=3.08$ p< .001
- Version 3 versus Version 4 $z=3.19$ p<.001

Experiment 5
We repeated the first part of Experiment 4 (this time using as participants fifteen undergraduates, aged between 19 and 22, from the University of Keele). In this final experiment, with the passage on insulating blankets, the students were asked to suppose that a mark of 4 out of 10 had been given to version 1. These results obtained are shown in Table VII. These results clearly reflect those shown in Table VI, and suggest that those in Table V are anomalous.

Table VII
The median rankings given to Versions 1, 2, 3 and 4 of the passages on insulating blankets by undergraduates at the University of Keele (N=15) with a standard mark of 4 given to Version 1.

<table>
<thead>
<tr>
<th>Version</th>
<th>Median</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4.0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>6.0</td>
<td>5– 9</td>
</tr>
<tr>
<td>3</td>
<td>6.0</td>
<td>2– 8</td>
</tr>
<tr>
<td>4</td>
<td>9.0</td>
<td>3–10</td>
</tr>
</tbody>
</table>

Significance levels:
- Version 1 versus Version 2 $z=3.41$ p<.0005
- Version 1 versus Version 3 $z=2.78$ p<.005
- Version 2 versus Version 3 $z=0.94$ not significant
- Version 2 versus Version 4 $z=2.47$ p<.01
- Version 3 versus Version 4 $z=3.41$ p<.0005
Concluding Remarks

These five experiments seem to point to four things:
(i) Comparisons can be made to see if any one version of a text seems preferable to another, but
(ii) the judgements will be crude and global (as shown in Experiments 1 and 2).
(iii) One can obtain more refined judgements from individuals by making the task more difficult (Experiments 3, 4 and 5), but
(iv) even here the results appear to reflect the difficulties of subjective scaling, differences between the texts, and differences between the kinds of judges employed.

It would seem that subjective preferences can provide additional information to experimenters about the effectiveness of changes to text, but that this information cannot be relied upon too greatly if fine judgement is required. It may be that a more refined technique, such as the method of paired comparisons, would prevent such difficulties, but it is not usual to use such a technique when only a small number of comparisons have to be made. The cruder methods we have used have not enabled us to separate out unequivocally the relative effects of changes in layout and changes in wording to the perceived effectiveness of instructional text. Our cautious conclusion, however, from the experiments reported here, is that with these texts neither change is seen as more effective than the other, but that both changes in combination are more effective than one alone.

Acknowledgements

We are indebted to all of our participants and in particular to Iona Hill who collected the data reported in Experiment 4.

References

Hartley, J. (1978a)
Space and structure in instructional text.
In J. Hartley (Ed) The Psychology of Written Communication: Selected Readings.

Hartley, J. (1978b)
Designing Instructional Text

Hartley, J. (Ed) (1980a)
The Psychology of Written Communication: Selected Readings
Hartley, J. (1980b)  
Spatial cues in text: a reply to Frase and Schwartz.  

Hartley, J. (1980c)  
Designing journal contents pages: the role of spatial and typographic cues.  

Hartley, J. (1981a)  
Sequencing the elements in references: a reply to Brian Shackel.  
Applied Ergonomics, 1981, 12, 1, 7-12.

Hartley, J. (1981b)  
Eighty ways of improving instructional text.  

Designing journal contents pages: preferences for horizontal and vertical layouts.  

The role of spatial and typographic cues in the layout of journal references.  

Some observations on producing and measuring readable writing.  

Klare, G. R. (1976)  
A second look at the validity of readability formulas.  
Journal of Reading Behavior, 1976, 8, 129-152.

The O U Student.  
In Tunstall, J. (Ed) The Open University Opens.  

Siegel, S. (1956)  
Non Parametric Statistics for the Behavioral Sciences  

Poulton, E. C. (1973)  
Unwanted range effects from using within-subjects experimental designs.  
Psychological Bulletin, 1973, 80, 113-121.