Comprehension of Legal Contracts by Non-Experts: Effectiveness of Plain Language Redrafting

MICHAEL E. J. MASSON and MARY ANNE WALDRON

University of Victoria

SUMMARY

The primary motivation behind the advocated use of plain language in legal documents is to increase comprehension among non-experts. We report empirical evidence regarding the effectiveness of three kinds of simplification of standard legal contracts for increasing comprehension among naïve readers. A set of legal contracts was redrafted in three stages to produce three modified versions. In the first stage we removed or replaced archaic and redundant terms; in the next stage simplified words and sentence structure were introduced; in the final stage legal terms were defined or replaced with simpler terms. Comprehension, as measured by paraphrasing and question-answering tasks, was reliably enhanced by the use of simplified words and sentence structure, but absolute levels of comprehension were still very low. An examination of erroneous responses suggested that, quite apart from the constraints of language, non-experts have difficulty understanding complex legal concepts that sometimes conflict with prior knowledge and beliefs.

Efforts to promote clarity in communication have a long history. As early as the seventeenth century, scientists have expressed concern about aspects of speaking and writing that obfuscate rather than clarify the explanation of ideas:

In 1667, The Royal Society of London was five years old. Laying the foundation of modern science and technology, the fellows rejected ‘amplifications, digressions, and swellings of style’. They urged members to use a ‘close, naked, natural way of speaking; positive expressions; clear senses’ a native easiness; bringing all things as near the mathematical plainness, as they can’ (Jones, 1991, p. 1).

Nowhere today is this issue of greater concern than in the legal field. Although the consumer movement of post-World War II North America is probably the root of the modern plain language movement in law, criticism of statutes and contracts based upon archaic, obscure language and the use of legal gobbledygook is much older. For example, Sir Matthew Baillie Begbie, Chief Justice of British Columbia from 1870 to 1894, is reported to have said ‘The Statute books are exceedingly

This project was supported by a grant from the Ministry of the Attorney-General of British Columbia. We thank Beth Macdonald for her assistance in redrafting the legal documents, Angela Micco for testing subjects and scoring data, and Tanya Payne for scoring data. We are also grateful to Phil Knight and Jacquelyn Nelson for their encouragement and support regarding plain language research.

Correspondence should be addressed to Michael Masson, Department of Psychology, University of Victoria, P.O. Box 3050, Victoria, British Columbia, Canada V8W 3P5.

CCC 0888-4080/94/010067-19
© 1994 by John Wiley & Sons, Ltd.

Received 27 April 1993
Accepted 25 May 1993
muddled. I seldom look into them' (Hamilton and Shields, 1979, p. 515). But it was not until the mid-twentieth century that serious efforts directed to improving the readability of legal documents began.

The question was: How could it be done? Early work emphasized the development of readability tests (Felsenfeld, 1981–2; Flesch, 1948) and general exhortations to lawyers to use plain words and short sentences (Flesch, 1962, 1972, 1973, 1979; Wydick, 1979). In many cases, American states passed plain language acts regulating the wording of certain contracts (e.g. New York General Obligation Law, 1980–1). In Canada, although plain language was not required by legislation, some businesses began to see merit in simpler language and one lender attracted considerable attention by producing a plain language mortgage (Dick, 1980). Much effort also has been invested recently in trying to teach lawyers to write plain English. For example, the Canadian Bar Association and the Canadian Bankers Association (1990) compiled a set of rules for clear writing.

At the same times, plain language is not without its critics. Readability tests, critics argue, provide no guarantee of readability (Black, 1981; Ross, 1981). In fact, incomprehensible sentences could score well. Moreover, using simpler language and abandoning legal terminology could introduce uncertainty into documents that were currently well understood (Prather, 1978). It may be that complex ideas require complex words (Aiken, 1959).

Economic analysis, too, may lead to questions about the usefulness of the plain language movement. Where risk is allocated by the common law to the consumer, Cohen (1981–2) suggested that the determination to simplify consumer contracts will mean that the commercial enterprise presenting the document will simply omit any reference to that risk, leaving the consumer to rely on her or his presumably inadequate conceptions about how a court would interpret a particular phrase.

Other branches of modern legal theory, influenced by theories of language and literature, also attack the assumption of plain language drafting. Does not the plain language movement depend, to some degree at least, upon the notion that language is capable of objective meaning? This notion receives serious challenge from scholars such as Fish (1989, p. 4), who wrote bluntly:

... to put it in the most direct way possible, ... there is no such thing as literal meaning, if by literal meaning one means meaning that is perspicuous no matter what the context and no matter what is in the speaker’s or hearer’s mind, a meaning that because it is prior to interpretation can serve as a constraint on interpretation.

Is the result of all this that the lawyer is required politically to do that which it may not be possible to do?

Today, the plain language movement remains very much a gospel in need of a theology. Admonitions to write more plainly, and criticism of those who admonish, are based largely upon untested assumptions. What actually works in legal writing? Very little empirical work exists to answer this question. The chief area of investigation, to date, has been the effect of simplification on jury charges (Diamond, 1993; Kramer and Koenig, 1990; Severance, Greene, and Loftus, 1984; Severance and Loftus, 1982; Steele and Thornburg, 1988). Generally, these studies concluded that using psycholinguistic methods to rewrite pattern jury charges resulted in a
significant improvement of juror comprehension. Even with revised instructions, however, jurors continued to make a considerable number of errors.

Similarly, in a report of two studies on the simplification of consumer-credit contracts, Davis (1977) demonstrated that both increased readability and reduction of information load were important factors in improving understanding in both high and low income groups. Davis also recognized the inherent difficulty of the task of changing consumer behaviour. He concluded:

Whether we can accomplish this goal remains to be seen. Perhaps the entire venture will prove to have been a mistake. But one thing is certain, it is too late to turn back. The enormous amount of societal energy expended in the hope of changing consumer behavior requires us to take every reasonable additional step toward accomplishment of that goal ... To be sure, we risk throwing good effort after bad. Yet, if simplifying consumer contracts is feasible, as I think it is, the choice seems clear (Davis, 1977, p. 907).

If the legal profession is to attempt to simplify its language in statutes and contracts, we suggest that further empirical research needs to be done to seek answers to the questions that are still debated. First, can we demonstrate that changes in style of drafting legal documents increase the comprehension of an ordinary reader? One of the objectives of the study reported here was to test the hypothesis that people with no formal legal training will more effectively comprehend the content and implications of a legal document if that document is drafted in simpler language.

Second, if this hypothesis is correct, what changes make the most difference? This issue is perhaps the most significant one for the lawyer faced with the task of simplifying the language in a legal document. As noted above, much of the literature simply assumes that certain words, phrases, and sentence structures are more easily understood than others and that all such changes are equally effective. This study was also designed to measure differences in comprehension resulting from different operations used to simplify a document.

Third, what limits exist to plain language as a means of increasing comprehension? Even if the use of plain language is capable of improving comprehension of legal documents, it is possible, and indeed past research and present theory suggest, that strict limits may exist on how well an untrained reader can understand even the simplest expression of legal obligation. The purpose of the study described in this paper was to begin to address some of these issues.

In this study a sample of subjects read a set of legal documents drafted in four different versions, representing varying degrees of simplification of language. Three steps were involved in redrafting the original documents. In the first step archaic terms (e.g. hereinafter) were replaced or removed, producing the archaic-terms-removed version. This step constituted a simple remedy that does little to alter the substance of a document, but that has been suggested as a possible benefit in improving the comprehension of documents. The second step involved a more radical change in style. The extremely long sentences that characterized the documents were broken into shorter sentences and difficult words were replaced with simpler terms, creating the plain-language version. These changes were intended to make comprehension more successful by reducing demands on lexical knowledge and working memory (Just and Carpenter, 1980, 1992; Kintsch and van Dijk, 1978). In addition, references to contracting parties, such as 'mortgagee' and 'mortgagor', were replaced by the
personal pronouns ‘you’ and ‘I’ to make the reader’s role and obligations under the agreement more clear.

Another potential benefit of the insertion of personal pronouns was to move the flavour of the documents toward oral language. Explorations of the differences between oral and written language suggest that a critical distinction between these two forms is that oral language places emphasis on interpersonal involvement, whereas written language focuses on information transmission (Tannen, 1985). An important implication of this distinction is that a written document is separated from its author and takes on a meaning that is somewhat independent of the author’s intention (Olson, 1980). Consequently, Olson suggests, a reader relies to a greater extent on what was said in a document than on the author’s intention. By moving the wording of a document towards oral language, therefore, we may be increasing readers’ appreciation of the intention underlying the document. Understanding the intention of a contractual agreement can serve as a useful heuristic when interpreting and applying its contents.

In the final step of redrafting the documents, specialized legal terms (e.g. subrogated) were replaced by simpler phrases or were defined in the text, to construct the legal-terms-defined version. Although we did not test specifically for knowledge of these terms, increased understanding produced by this modification may serve to clarify the context in which other elements of the document are interpreted.

We assessed the effectiveness of these modifications by measuring how quickly subjects were able to read the documents and how well they understood their content. Comprehension was measured in two ways. First, subjects were asked to paraphrase a segment of each document using their own words, and the completeness of these paraphrases was assessed. This procedure has been used effectively in earlier research on comprehension of modified jury instructions (Goodman and Greene, 1989). Second, subjects answered a set of questions about each document in which hypothetical scenarios were described and judgements were elicited concerning the rights of the parties involved.

In analysing the data, we emphasized three planned comparisons, each of which tested the effectiveness of one of the redrafting methods. The benefit of removing archaic terms was assessed by comparing performance on the original and archaic-terms-removed versions. The effect of using simpler words and sentence structure was tested by comparing the average of performance on the original and archaic-terms-removed versions with the average of performance on the plain-language and legal-terms-defined versions. Finally, the influence of defining legal terms was examined by contrasting performance on the plain-language and legal-terms-defined versions.

**METHOD**

**Subjects**

Two samples of subjects were obtained, one consisting of clerical staff at the University of Victoria and the other comprised of city residents who were taking courses through the University’s Division of University Extension. Participants were solicited

---

1 We thank an anonymous reviewer for drawing our attention to this issue.
through a letter that announced a research project on the comprehension of legal documents and offered a payment of $30 for taking part. There were 24 subjects in each of the two samples, with 16 females in the clerical sample and 20 females in the student sample. The mean age and educational level of the subjects in these samples were 36.6 and 14.7 years, respectively, for the clerical sample and 31.3 and 15.8 years, respectively, for the student sample. The ranges of age and educational levels, which were similar for both samples, were 20–50 years and 12–19 years.

Materials

Five documents were selected; four critical documents were to be used in actual testing and a fifth was included to give subjects practice at the procedure. The four critical documents were excerpted from typical legal agreements that might be read by a member of the public: a mortgage, an agreement for sale of property, a bank loan, and a renewal of a lease. The excerpts that we used as documents ranged in length from 391 to 495 words; the average length was 422 words.

In their original form, the documents were in a formal legal style, replete with the faults identified in many of the plain writing texts noted above. Each document was then redrafted in three different styles. The object of the redrafting process was not to produce a perfect draft nor even to improve the document as much as possible. Rather, the object was to make specific, well-defined changes at each stage so that each change could be related to any measurable changes in comprehension. Redrafting was carried out in a sequence of three steps:

1. All archaic terms were replaced or removed. For example, ‘indenture’ was replaced with ‘agreement’ and ‘hereinafter’ was often deleted. Redundant words or phrases were deleted. At this stage, the structure of the sentences was left as close as possible to the original; the language was simply ‘cleaned up’.

2. The segments were subjected to a more dramatic change to a plain English style. Difficult words were replaced with simpler ones, long sentences were divided into shorter sentences, and the passive voice was replaced with the active voice. These changes were made specifically to improve the readability of the documents as measured by traditional readability tests. Evidence of the effectiveness of these modifications is shown in Table 1, which presents a summary of some of the grammatical characteristics of each of the versions of the documents. An additional change involved replacing references to the parties by references to ‘I’ and ‘you’. ‘I’ referred to the reader, who in all cases played the role of a consumer, and ‘you’ referred to the party entering into the agreement with the reader (e.g. bank, landlord). Even at this stage, however, the order of ideas was not rearranged any more than was needed to carry out the changes described.

3. In the final draft, the plain language changes were retained but, in addition, all legal terms were replaced by simpler, more common words or were explained in the text.

These procedures resulted in four versions of each document: the original version and a redrafted version at each stage of redrafting. For convenience, the versions were designated original, archaic-terms-removed, plain-language, and legal-terms-defined, respectively. None of these versions employed any restructuring of ideas,
nor was the design of the document altered. Matters such as print face, paragraphing style, justified or unjustified margins, bold faced type, headings, and so on can make significant differences to readability. In this study, however, we attempted to avoid all such changes, narrowing the field of inquiry to only wording and syntactic change. The average lengths of the archaic-terms-removed, plain-language, and legal-terms-defined versions were 299, 272, and 360 words, respectively. For illustrative purposes, a section taken from each version of one of the documents is shown in Appendix 1.

Two further items were created for each document set (i.e. each document in all four versions). First, a brief introduction was written that described the nature of the transaction to which the document related (e.g. ‘Assume that you are in the process of buying a house’). The introduction also described the nature of the document (e.g. that same introduction continued, ‘you and the seller of the house agree you will pay the price over time’).

The second item prepared for each document set was a set of four questions. These questions were designed to test the subjects’ understanding of the passage. They were prepared according to several guidelines. First, each question could be answered from reading any of the four versions of a document. For example, the questions did not ask about the interpretation of legal terms that would not be known to lay persons and that were defined only in the legal-terms-defined version. Second, each question required more than a simple repetition of the contents of the document. The subject was asked to give an answer based on a hypothetical scenario and on the contents of the document. Third, each question had two parts: a part that could be answered ‘yes’ or ‘no’, and a second part that required a justification for the answer. Finally, each justification contained one to three steps in reasoning to be used in scoring the questions. As an example, one of the questions is provided in Appendix 1.

In addition to asking questions, comprehension was assessed by having subjects provide a paraphrase of a segment of each document. One segment was identified for each document for this purpose. The major criterion in selecting a segment was that its content was comparable across all four versions of a document, although the length of the segment varied with version. The average lengths of these segments were 104, 76, 64, and 79 words for the original, archaic-terms-removed, plain-language, and legal-terms-defined versions, respectively. Part of one of these segments is shown in Appendix 2. The excerpt is shown in the original and in the legal-terms-defined versions. We also include a list of the idea units or propositions (Kintsch, 1974) that represent the content of the legal-terms-defined version of the excerpt. The list of propositions from the legal-terms-defined version was used in scoring
all paraphrase responses provided by the subjects. Our rationale was that the legal-terms-defined version consisted of the major, unembellished concepts common across all versions.

All materials were typed in double-spaced format for presentation to subjects. Each introduction, document, question, and selected segment appeared on a single, separate page. The document title (Agreement for Sale, Mortgage, Bank Loan, or Lease Renewal) appeared at the top of each of these pages. A similar set of materials was constructed for the fifth document, which was used to familiarize subjects with the tasks. This document concerned a lost cheque disclaimer and was written in a plain language style.

Procedure

Each subject was tested individually in a quiet room in a single session lasting approximately 90 minutes. The subject was informed that the purpose of the study was to assess comprehension of legal documents by having people attempt to paraphrase parts of documents and to answer questions about them. The procedure involved five cycles of document reading and comprehension testing. The first cycle involved the practice document to familiarize the subject with the procedure and each of the remaining four cycles involved one of the critical documents. For each subject, each of the four critical documents appeared in a different version, so that each subject was tested once with each document version.

The reading phase of each read–test cycle began by having the subject read the introduction to the document. Then the subject read the entire document while being timed. The subject was asked to read the document in its entirety one time and to indicate when she or he was finished. The document was then removed and the reading time recorded. The subject was then presented with the selected segment of the document (in the same version as the complete document) and asked to read it through one time. Reading time was recorded and the page was removed when the subject indicated that she or he had finished reading.

In the first part of the testing phase, the subject was asked to provide a written paraphrase of the selected segment of the document. In providing a paraphrase the subject was instructed to write down the ideas from the segment in her or his own words. The subject recorded the paraphrase on a blank page of an answer booklet. Finally, the subject provided written answers to the series of four questions associated with the document. The questions were presented one at a time and the subject wrote the answer to each question on a separate page of the answer booklet.

After completing the read–test cycle for the practice document each subject completed the cycles for the four critical documents which were presented in a random order. Within each of the two samples of 24 subjects, document versions were assigned to subjects so that each version of a document was read by six different subjects.

RESULTS AND DISCUSSION

We report the results of statistical analyses of three aspects of the behavioural data: (1) reading rate; (2) completeness of paraphrased descriptions of document content; and (3) accuracy of question answering and justifications for answers. For each
behavioural measure our interest was in comparing performance on documents written in the four different drafting styles. Therefore, for each measure we present the mean score associated with each of the document versions.

The statistical analyses we conducted involved three comparisons in each case. The first comparison was concerned with the effect of using plain language and consisted of a comparison of average performance on the original and archaic-terms-removed versions against average performance on the plain-language and legal-terms-defined versions. The second comparison was directed at the question of whether simply removing archaic and redundant terms was an effective device, and consisted of comparing performance on the original and archaic-terms-removed versions. The third comparison was between the plain-language and legal-terms-defined versions and addressed the issue of whether defining legal terms produced any benefit for comprehending aspects of the documents that did not involve the definitions themselves. Two analyses are reported for each of the comparisons. One analysis is based on scores computed for each subject and permits some degree of generalization from our sample to other similar members of the population. The other analysis is based on scores computed for each document and permits some generalization to other similar documents (Clark, 1973). A type I error rate of 0.05 was used in all statistical tests. Preliminary analyses indicated that there were no significant differences between the two samples of subjects so all data reported here represent averages computed across the two samples.

Reading rate

The mean reading rate (measured in words per minute) for each of the document versions is shown in Table 2. One set of means is presented for reading rates for entire documents and the other set represents reading rates for the segment that was read in preparation for the paraphrase task.

The use of plain language did not have an effect on reading rate for either the entire document or for the paraphrase segment; the comparison of the average of the original and archaic-terms-removed conditions against the average of the plain-language and legal-terms-defined conditions did not produce a statistically reliable difference. The difference between the original and archaic-terms-removed versions was not significant either for the entire document or for the paraphrase segment. The difference between the plain-language and legal-terms-defined versions was significant when reading rate for entire documents was considered, $F(1, 47) = 8.83$, $MS_e = 1630.5$, for the analysis of data based on individual subject scores, and $F(1, 3) = 42.63$, $MS_e = 57.5$, for the analysis based on individual item (document) scores.

The difference between these two conditions approached significance in the reading
Table 3. Mean percentage and standard deviation (in parentheses) of propositions included in paraphrases of documents

<table>
<thead>
<tr>
<th>Document version</th>
<th>Original</th>
<th>Archaic-terms-removed</th>
<th>Plain-language</th>
<th>Legal-terms-defined</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>15.2</td>
<td>24.0</td>
<td>29.8</td>
<td>36.0</td>
</tr>
<tr>
<td></td>
<td>(15.9)</td>
<td>(23.6)</td>
<td>(27.1)</td>
<td>(24.9)</td>
</tr>
</tbody>
</table>

rates for paraphrase segments, $F(1, 47) = 3.03, MS_e = 5.678.7, p < 0.09, F(1, 3) = 8.33, MS_e = 526.1, p < 0.07$.

The reading time results indicate that readers were reliably affected only by a change from the plain-language to the legal-terms-defined version of the document. One possible reason for increased reading rate is the absence of undefined, specialized, legal terms in the legal-terms-defined version. On the other hand, it is also possible that subjects tended to read that version faster because it was about one-third longer than the plain-language version. Sensitivity to document length is also hinted at by the non-significant trend toward faster reading rates for the original version, which was the longest of all. A number of factors probably contribute to variations in reading rate, such as self-assessed success at comprehension or frustration at one’s inability to understand complex wording or sentence structure. Our measure does not allow us to draw firm conclusions regarding the factor(s) that determined the observed fluctuation in reading rates.

One important conclusion that we can draw from the reading rate data is that subjects appeared to make genuine attempts at understanding the documents. Their reading rates were well below what would be expected of college students reading easier material such as newspapers or narrative stories in anticipation of comprehension tests (about 250 words per minute; Masson, 1982).

**Paraphrase task**

The written paraphrase protocols were typed and given to two raters for scoring. Each protocol was scored independently by the two raters, without any indication as to the version of the document on which the protocol was based. The scoring procedure involved comparing a protocol to the list of propositions associated with the segment. Credit was given for each proposition that was contained in a protocol, with transformations of wording or syntactic structure permitted. After all the protocols were scored, the evaluations of the two raters were compared. It was found that the raters agreed in their scoring of 89.0% of the propositions. Instances of disagreement were handled by having the raters reach consensus on whether or not to assign credit.

Scores on the paraphrase task were obtained by computing the percentage of target propositions that were included in each protocol. There were 13 instances in which subjects stated that they were not able to provide a paraphrase of the segment. Of these, 11 occurred with original or archaic-terms-removed versions and two involved plain-language or legal-terms-defined versions. In these cases a score of zero was assigned. The mean percentage of propositions included in the protocols from each document version is shown in Table 3.

Statistical analyses, treating each proposition as an item, indicated that using plain
language (plain-language and legal-terms-defined versions) resulted in a significant increase in the percentage of propositions that subjects were able to paraphrase correctly in comparison to versions that did not use plain language (original and archaic-terms-removed). $F(1, 47) = 15.88$, $MS_e = 267.8$, $F(1, 50) = 38.00$, $MS_e = 148.9$. In addition, paraphrases obtained with the archaic-terms-removed version were significantly more complete than paraphrases obtained with the original version, $F(1, 47) = 4.55$, $MS_e = 403.9$, $F(1, 50) = 5.58$, $MS_e = 141.3$. The difference between the legal-terms-defined and plain-language versions was not significant.

In addition to evaluating the percentage of propositions that were paraphrased correctly, we noted that there were instances in which subjects misinterpreted the documents. Two examples of this are shown in Appendix 3. The first example contains the archaic-terms-removed version of a segment that was to be paraphrased and the subject’s paraphrase of that segment. In the second example we provide part of a segment from the legal-terms-defined version of a document and the remainder of the document that followed the segment. We also provide an excerpt from a subject’s paraphrase response that pertained to the material at the end of the document. These examples illustrate some ways in which subjects transformed the intent of the documents in their efforts to comprehend.

Question answering task

Answers to questions were evaluated by comparing them to the ideal answers. Each question consisted of a binary decision (e.g. ‘Can the bank seize your Thunderbird?’) and a justification (e.g. ‘Why?’ or ‘Why not?’); performance on these two parts was scored separately. Answers to the decision part of the questions were scored only as correct or incorrect. In the case of justifications, partial marks were awarded for justifications that were not completely correct, and errors of reasoning were also noted. An error could involve either the use of incorrect information or the inappropriate use of accurate information to justify a decision. Two raters scored the answers independently, their scores were then compared. The raters’ scores were in agreement in 88.9% of the cases and instances of disagreement were resolved by consensus.

For the decision component of the questions, we computed the percentage of cases in which the subjects responded correctly, incorrectly, and failed to give a response. The means for each of these response types are shown in Table 4. We applied statistical analyses to determine whether document versions significantly influenced the percentage of correct decisions. For the analysis by items, each question was treated as a separate item. These analyses indicated that use of plain language improved the accuracy of decisions, $F(1, 47) = 21.75$, $MS_e = 278.4$, $F(1, 15) = 7.72$, $MS_e = 230.1$, but that removing archaic terms or defining legal terms made no significant difference. A second set of analyses applied to incorrect responses indicated that there was a complementary effect of the use of plain language on these scores, with fewer errors committed when plain language versions were used, $F(1, 47) = 10.21$, $MS_e = 195.3$, $F(1, 15) = 14.35$, $MS_e = 51.8$.

The mean percentages correct for the decision component shown in Table 4 seem rather low, with performance on the original and archaic-terms-removed documents near 50%. It is important to keep in mind that although the decision component consisted essentially of a binary choice, 50% correct does not mean that subjects
Table 4. Mean percentage and standard deviation (in parentheses) of responses in the question answering task

<table>
<thead>
<tr>
<th>Question component and response type</th>
<th>Document version</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Original</td>
<td>Archaic-terms-removed</td>
</tr>
<tr>
<td>Decision</td>
<td></td>
<td></td>
</tr>
<tr>
<td>% correct</td>
<td>50.5 (26.5)</td>
<td>49.0 (27.8)</td>
</tr>
<tr>
<td>% incorrect</td>
<td>20.3 (21.7)</td>
<td>23.4 (22.1)</td>
</tr>
<tr>
<td>% missing</td>
<td>29.2 (25.4)</td>
<td>27.6 (21.4)</td>
</tr>
<tr>
<td>Justification*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>% correct</td>
<td>25.6 (17.3)</td>
<td>25.3 (19.3)</td>
</tr>
<tr>
<td>% incorrect</td>
<td>24.0 (19.3)</td>
<td>22.9 (21.2)</td>
</tr>
<tr>
<td>% missing</td>
<td>59.6 (16.2)</td>
<td>59.4 (18.1)</td>
</tr>
</tbody>
</table>

* The percentages do not add to 100 for the Justification component because partial marks could be awarded, and a justification could contain a mixture of correct and incorrect statements. % correct refers to the percentage of justification elements that were correctly stated; % incorrect refers to the percentage of justifications that included at least one erroneous statement (see Appendix 4 for examples); % missing refers to the percentage of questions for which no justification at all was provided.

were merely guessing on these items because there were instances in which no response was given. Considering only those cases in which a response was made, the percentage of correct responses averaged across the original and archaic-terms-removed versions was 69%; averaged across the plain language and legal-terms-defined versions it was 84%.

The justification component of question answering was analysed in the same way as the decision component. The percentage of correct justifications (including partially correct answers), percentage of justifications containing an error in reasoning, and percentage of questions for which no justification was given are shown in Table 4. Statistical analyses were applied to the percentage correct data. It was found that justifications were more accurate when subjects read versions that used plain language rather than versions that did not. \( F(1, 47) =14.68, MS_e = 126.7, F(1, 15) =4.54, MS_e = 101.9, p =0.05 \). The removal of archaic terms or the inclusion of definitions of legal terms did not have a significant effect on performance.

It is interesting to note that, unlike the decision component of the questions, the percentage of justifications that included errors was not influenced by document version. Instead, the error rate was similar across all four document versions, indicating that subjects continued to be susceptible to reasoning errors regardless of the document version. Examples of these errors are presented in Appendix 4.

**CONCLUSION**

The results from the comprehension tasks used in this study demonstrate that simplifying drafting style increases comprehension. Accuracy between the original and the final versions of the documents improved significantly both in answering questions and in producing a paraphrase of a portion of text. The method of scoring the paraphrase task may have to some degree favoured the final version. That final version (the legal-terms-defined draft) was the initial source of the list of propositions
used to determine the accuracy of subjects' paraphrases. Arguably, this could result in raters scoring as more frequently correct those propositions produced by subjects who read the legal-terms-defined draft. Each version was reviewed, however, to determine that the propositions drawn from the legal-terms-defined draft were reflected accurately in the segment used for the paraphrase task. Also, the similarity between the results obtained from the paraphrase task and those obtained from the responses to the questions suggests that this factor was not particularly important.

One might question whether the testing situation in which our subjects were placed is a sufficiently accurate simulation of real life encounters with legal documents. For instance, our subjects may not have shared the same degree of motivation as people who enter into actual, binding agreements. But there is ample evidence to suggest that our subjects were adequately motivated to treat the tasks seriously and to engage significant effort. First, their reading rates were clearly reduced relative to rates that are typical of similarly educated people reading much less demanding material. It was clear during the testing session that subjects were making genuine efforts to comprehend the documents and to express the knowledge they were able to obtain. Second, a number of the subjects remarked that a primary reason for participating was their interest in their own mortgage or bank loan. Our conclusion is that the results of this study provide a reasonably accurate picture of document comprehension, and the changes in comprehension that can be expected by the use of the redrafting procedures we tested.

For the legal profession, perhaps the most useful of our results is the demonstration of which drafting changes produce the greatest effect. For example, undertaking to modernize a law firm's precedents is a major commitment of time and resources. The results of this study suggest that very little is achieved by simply removing archaic terms and legalese. Although these changes may make the document aesthetically more pleasing, to make real gains in reader comprehension, revisions must attack the complexity of language and syntax in a more radical way. The changes we introduced in moving from the archaic-terms-removed to the plain-language version primarily reduced the difficulty of vocabulary and shortened sentences. By using more familiar words we made more concepts accessible to readers, and by using shorter sentences fewer demands were placed on working memory capacity. Theories of reading emphasize the role of working memory both in understanding words in context (Daneman and Green, 1986) and in forming a coherent, integrated representation of text (Fletcher, 1986; Just and Carpenter, 1992; Kintsch and van Dijk, 1978). Breaking the text into sentences that could be maintained in working memory should have enabled readers to more successfully integrate propositions. It is also possible that, by introducing personal pronouns, we moved toward an oral form of language, and thereby highlighted the intention underlying the documents (Olson, 1980; Tannen, 1985).

At the other end of the scale, little more improvement seems to come from replacing legal terms with alternative common terms, or providing explanations for them. This observation should be tempered by emphasizing that the questions avoided any issues that depended directly upon an understanding of legal terminology. That structure was adopted purposely to avoid unfairly skewing the results toward the legal-terms-defined version, or unduly favouring those who might have had prior exposure to the terms. But the result is that we can conclude only that comprehension of parts of the document not containing legal terms is not greatly improved by explain-
ing legal terms. We cannot conclude that these explanations fail to improve comprehension of the legal principles involved. In fact, it is self-evident that they would.

The concern with explanation of legal terms raises the difficult issue of what is needed to understand a legal agreement and how much drafters can do about it. Although most lawyers would feel comfortable with eliminating legalese and archaic words, and many might believe that the more radical changes undertaken in our second drafting change were acceptable, few would be prepared to explain in a document the legal concept in which it is based. This attitude reflects a belief that legal principles are too complex, too open to interpretation, and are part of an interrelated system of law and fact that cannot easily be cut into isolated definitions. However much of law's inaccessible nature may be explained by obscurantism, not all of it melts away in the face of plain language.

These views are supported by our findings that despite improved comprehension, performance of subjects on both the question answering and the paraphrase tasks remained relatively poor (in the best cases average performance ranged from about one-third to two-thirds correct, depending on which aspect of comprehension was measured) and misconceptions were apparent across all versions of the documents. Why subjects continued to encounter difficulty even with the most radically modified versions is a matter of speculation. A hint can be taken, however, from a linguistic analysis of interactions between a paediatrician and a child's mother conducted by Tannen and Wallat (1986, 1987). Their analysis indicated that mismatches between the respective knowledge schemas of the paediatrician and the mother, regarding health and the disease in question, were responsible for the mother's concerns, which were highly resistant to attempts at dissuasion by the paediatrician. Moreover, Tannen and Wallat noted that the prior schemas held by the participants were maintained even after the exchange of information.

As in the case of consultations between health-care professionals and consumers, it is possible that legal concepts are difficult to understand because, even when explained in plain language, they are complex or because they are in conflict with folk theories of the law. The subjects in this study (and lay people in general), may have been relying on inaccurate prior knowledge of the law or on their own intuition about justice, which frequently does not reflect the legal reality (see Appendix 4). These results suggest that plain language drafting alone will take us only part way to the goal of making the law more broadly understood. It must be supported by other measures such as public legal education and individual counseling of persons faced with legal obligations.

REFERENCES


APPENDIX 1

The following is a portion of one of the test documents. This example was taken from the Bank loan document. It is included here in all four versions and followed by one of the questions the subjects were asked.

**Original version**

Acceptance by the bank of payments in arrears shall not constitute a waiver of or otherwise affect any acceleration of payment hereunder or other right or remedy exercisable hereunder. No failure or delay on the part of the bank in exercising, and no failure to file or otherwise perfect or enforce the bank’s security interest in or with respect to any collateral, shall operate as a waiver of any right or remedy hereunder or release any of the undersigned . . .

**Archaic-terms-removed version**

Acceptance by the bank of payments in arrears shall not constitute a waiver of any acceleration of payment of this loan or affect any other right exercisable under this agreement. No failure or delay on the part of the bank to file, perfect or enforce the bank’s security interest with respect to any collateral, shall operate as a waiver of any right under this agreement or release any of the undersigned . . .

**Plain-language version**

If you fail to take steps to protect or enforce your security, you do not waive your rights. I am not released by this. You do not lose your right to accelerated payment or any other of your rights under this agreement by accepting overdue payments.

**Legal-terms-defined version**

If you fail to take steps to protect or enforce these interests I have given you, you do not waive your rights. I am still liable to you. If I do not make my payments, you will have the right to insist on being paid the whole loan at once. You do not lose this right or any other of your rights under this agreement by accepting overdue payments.

**Question**

You missed two loan payments. When the next payment was due, you made up one and the bank accepted it. You told them at that time that you were pretty sure you could bring the loan up to date next month. Now the bank has given you notice they intend to seize your Thunderbird. Can they? Why or why not?

---

1 This is a part only of the passage read by the subject.
2 The Thunderbird car was part of the collateral named for the loan.
3 To answer 'correctly', subjects were not expected to know any legal limitations on the rights of lenders. They were asked to answer the questions based only on the information contained in the documents and the given facts.
APPENDIX 2

The following is a portion of one of the document segments used in the paraphrase task. The segment was taken from the Mortgage document. It is included here in two of the versions and followed by a set of propositions that represent the meaning of the legal-terms-defined version.

Original version

... and the decision of the Mortgagee as to the validity or amount of any advance or disbursement made under this mortgage or of any claim so paid off shall be final and binding on the Mortgagor.

Legal-terms-defined version

... You can decide if this money has to be paid or not. Your decision is final and I will abide by it.

List of propositions

... 
8 (bank may decide 9) 
9 (whether money must be paid) 
10 (bank's decision is final) 
11 (borrower will abide by bank's decision)

APPENDIX 3

The following are examples of paraphrase protocols in which subjects showed evidence of misunderstanding the contents of the documents.

Archaic-terms-removed segment (Bank loan)

Actual document

Acceptance by the bank of payment in arrears shall not constitute a waiver of any acceleration of payment of this loan or affect any other right exercisable under this agreement. No failure or delay on the part of the bank to file, perfect or enforce the bank's security interest with respect to any collateral, shall operate as a waiver of any right under this agreement or release any of the undersigned.

Paraphrase

Any acceptance of payment on the side of the bank can count as waiver of any acceleration of payment. Any failure of the bank to accept payment or any new agreements on obligation or collateral at the appropriate time can be taken as a waiver for payment, collateral, or further obligation.

Each proposition is assigned a number for the convenience of expressing embedded propositions, as in proposition 8.
Legal-terms-defined segment (Agreement for sale)

*Actual document (material in italics followed the segment designated for paraphrasing)*

You promise that when the land is transferred, no one will have any rights against the land registered in the Land Title Office. This does not include any costs imposed by the government or utility companies on the day of this agreement or later. *When the government first transferred this property to an owner, it may have kept back some rights over the land or excluded some rights from that transfer. You are not responsible to transfer these rights to me.*

*Paraphrase (this is only a part of the paraphrase response)*

... But actually, it is really not my land even after I pay the original owner, it is owned by the government.

**APPENDIX 4**

The following are examples of justifications for answers to questions in which errors in reasoning were made. The erroneous part of each response is in italics. Although some of these statements were factually correct, they were used incorrectly in reasoning.

**Original version (Bank loan)**

*Question*

You are not making the payments on your 1990 loan. You owe ABC bank $5,000, now due. Your aunt sends you a cheque for your birthday for $1,000. You deposit it in your account at ABC bank. The next week you want to withdraw $500 to buy a portable compact disc player. The bank refuses to pay you the money. Do they have to? Why or why not?

*Correct answer*

No, the bank does not have to give you the money. You gave them the right to set off (right to deduct the amount of your debt directly) against any property of yours they hold for any purpose.

*Subject's response*

If the loan was not secured then an attachment on my wages at a rate of 10% would have been part of my agreement. The deposit of $1,000, however, was placed in an account (private account) which is under a separate agreement and cannot be garnished by the bank.

**Original version (Agreement for sale)**

*Question*

Congratulations! You have just paid off everything you owe under this agreement. However, you find that although your agreement was with Mary Smith, the actual registered owner of the property is Creams and Dreams Limited, a catering company that Mary Smith owns. Does Mary have a problem? Why or why not?
Correct answer
You agreed to accept title signed by the registered owner, even if not Mary. We have to know if she can get Creams and Dreams to sign. If she can (can't) there isn't (is) a problem.

Subject's response
Yes, I think Mary does have a problem because she knowingly committed fraud. Although she is indirectly the owner the real owner is Creams and Dreams and this was not the name on the top of the purchase agreement.

Archaic-terms-removed version (Mortgage)

Question
You now owe $100,000 on the loan. The value of the property, however, has taken a big plunge. Your building and land are worth only $75,000. However, you have some very expensive air conditioning equipment. You decide to abandon the property, but before you do, you rent a truck and load up the air conditioning equipment. Your lender insists you return it to sell with the property. Do you have to give it back? Why or why not?

Correct answer
You have to return the air conditioning equipment. The agreement specifically lists that as deemed to be a fixture.

Subject's response
No, because that was extra money that you had put out to improve your living conditions.

Archaic-terms-removed version (Mortgage)

Question
You made an agreement with a building contractor that she could take your Porsche 911 if you did not pay her for renovations done on your house. You ran into financial trouble and could not pay her. You also take the position that the work she did was substandard and not worth what she claimed. She files a claim against your house at the Land Title Office. The bank that holds the mortgage on your house paid her off. Can the bank take your car? Why or why not?

Correct answer
The bank can take the car. You agreed the bank could decide who to pay and how much. You agreed the bank could have the rights of anyone they paid.

Subject's response
No the bank can't take your car because you offered the car to the contractor and not the bank.

Legal-terms-defined version (Agreement for sale)

Question
A builder you hired to renovate your house filed a lien in the Land Title Office
when you failed to pay him. Your lender paid his claim. In recovering from you the money it spent to do this, it hired a collection agency for $300. Additionally, your lender’s lawyer wrote you a threatening letter for which she charged $200. Which of these charges do you have to pay and why?

Correct answer
You must pay the $300 because you have agreed to pay all costs of collection. You must pay the lawyer’s fee because it is also a part of the cost of collection you have agreed to pay.

Subject’s response
The $300 expense that the collection agency charged, but not the lawyer’s fee as they did not have to go through a lawyer to incur added charges.