Field Evaluation of Simplified English for Aircraft Workcards

S. Chervak,
Masters Student
Department of Industrial Engineering
SUNY-Buffalo

C. G. Drury
Professor
Department of Industrial Engineering
State University of New York at Buffalo

J. P. Ouellette
Program Leader
G.E. Aircraft Engines

ABSTRACT

The restricted technical language Simplified English (AECMA, 1995) was evaluated using aircraft maintenance workcards. One hundred seventy-five practicing Aircraft Maintenance Technicians (AMTs) were given a comprehension test of four different workcards, each produced in Simplified English (SE) and Non-Simplified English (Non-SE) versions, and with two different layouts. Simplified English versions gave improved performance as measured by comprehension error rate which was reduced from 18% to 14% with SE. Most of the improvement was obtained where conditions were most challenging: for more difficult workcards and with non-native English speakers. No effects of workcard layout on performance were observed.

INTRODUCTION

The importance of good document design practices to the writing of aircraft work control cards (workcards) has already been documented (Bohr, 1978; Patel, Drury and Lofgren, 1994). Patel, et al (1994) showed several deficiencies in structuring, wording, layout and typography, and related these to potential errors by Aircraft Maintenance Technicians (AMTs) in performing their tasks. An improved design was developed and evaluated to demonstrate its superiority. This design was based on the application of the principles of document typography and layout from the human factors literature. Documents produced in the Patel, et al (1994) study had a better choice of case and font, a more consistent paragraph structure and better integration of text with graphics. There are, however, issues in document design which go beyond layout and typography.

Most major transport aircraft manufacturers now use Simplified English (SE) in their documentation. However, the impact of this restricted language on AMTs has not been directly measured. The current study provides such an evaluation to determine whether SE enhances (or degrades) comprehension of workcards by AMTs.

Simplified English

Since Latin faded as the common scientific language, there have been various attempts to produce artificial languages to allow people of different countries to intercommunicate. For general use, the early twentieth century saw Esperanto and later Basic English (Ogden, 1934). More recently, restricted technical languages have appeared, such as Caterpillar Fundamental English (CFE) for the documentation of agricultural vehicles, and Simplified English (SE) for the documentation of procedures on commercial aircraft. More information on the development and details of these restricted languages can be found in Shubert, Spyridakis, Holmback and Coney (in press).

Issues in Evaluation

While restricted languages such as SE make logical sense, there is still a need to evaluate their effectiveness. Despite potentially reduced ambiguity, there are still feelings among some technical writers that SE prevents them from expressing instructions in the most obvious manner. *Restricted* languages can appear *restrictive* to some. Since the documentation is designed for the user, the effect of SE on the AMT is the ultimate criterion. Hence direct evaluation of SE using actual workcards and practicing AMTs is an obvious step.

Evaluation by users has a long history in document design. For example, McLaughlin (1966) compared the readability of two versions of a government pamphlet using a comprehension test. He found the version which had been revised for readability gave improved comprehension. The relationship between readability and comprehension has been further documented in a number of studies reviewed by Klare (1978). A more recent compilation of studies of warnings (Edworthy, Hellier and Stantion, 1995), showed speed of use, accuracy and rating scales as frequently-used measures of the performance of different aspects of warning design.

A major evaluation study of SE, by Shubert, et al (in press) followed a methodology similar to that of McLaughlin (1966) using a comprehension test. SE and Non-SE versions of two maintenance manual procedures were tested on 127 engineering students. While having comparable overall lengths, the two procedures differed in a number of measures of writing complexity, one being more complex than the other. A between-subjects experimental design was used, where each subject was tested on only one of the four documents. The comprehension test was timed and performance on the test was measured both by whether each question had the correct answer and whether the information used for the answer could be located correctly within the maintenance manual procedure.

Analyses were performed separately for native English speakers and non-native English speakers and for the two maintenance manual procedures which differed in complexity. Measures of both the comprehension and the content location showed a significant effect of Simplified English and a significant Simplified English × Procedure interaction. The native English speakers scored higher than their non-native English speaking counterparts. Simplified English gave higher comprehension and location scores than non-Simplified English for the more complex procedure only. Performance time was not a significant factor, except that non-native English speakers were slower overall.

From these studies it was concluded that to evaluate a restricted language we must control both the users' native language and the document complexity. In addition, the evaluation should focus on the accuracy of comprehension using a comprehension test based upon the documents themselves. "Accuracy of comprehension" should measure the correctness of both comprehension questions and location questions.

METHODOLOGY

The basis of our methodology was to extend the comprehension test technique to the use of workcards by practicing AMTs. Differences from the Shubert, et al. (in press) study were the choice of subjects (AMTs *versus* students), levels of document complexity (four workcards *versus* two procedures) and the addition of two levels of workcard layout to provide a test performance of the Patel, et al. (1994) results.

Choice of Workcards

Following discussions with computational linguists at Boeing Inc. and with Aerospace Industries Association of America (AIAA) Simplified English Committee members, it was decided to use actual examples of existing workcards in the evaluation. For two aircraft types, Boeing had produced workcards in pre-SE maintenance manual language and had later modified these to Simplified English standards. Thus the workcards were realistic AMTs and represented actual writing practice by those who write maintenance manual procedures. In this way, difficulties of translating Simplified English workcards back into artificial non-Simplified English versions were avoided. The one drawback of this decision was that the SE interpretation was not always "perfect", *i.e.*, AIAA committee members could still identify a few possible changes needed to ensure full compliance with the latest release of SE. The standard of SE in these workcards was high, and represented real-world "good practice." The benefits of using "real" workcards were considered to far outweigh the few possible non-SE interpretations introduced by real technical writers in their normal writing practice.

Seventeen potential workcards were analyzed for possible inclusion in the study. The Boeing computational linguists and University of Washington technical communications researchers analyzed the non-SE versions of each in terms of total words, mean words per sentence, percentage passive voice, and Flesch-Kinkaid reading score. A task difficulty rating of each workcard by an experienced engineer was also used for guidance. Each of these variables was split at the median to be able to match workcards at the high or low level of each variable. From this analysis, four workcards were chosen: two "easy" on all the measures and two "difficult." Within each pair the document lengths were different, which would presumably mainly affect performance times, although document length could also affect comprehension through additional cognitive load. (Table 15-1, appendix) shows the four workcards chosen.

Each of these four workcards were then prepared in four versions:

1. Simplified English, original layout
2. Simplified English, Patel, et al. (1994) layout
3. Non-Simplified English, original layout
4. Non-Simplified English, Patel, et al. (1994) layout

The four versions were critiqued by our Boeing, University of Washington colleagues and the AIAA Simplified English Committee members. Based on their feedback, minor corrections were made to ensure consistency between versions.

Choice of AMTs

Following pre-tests in a Greater Buffalo International Airport facility to determine the adequacy of the methodology, contacts with airline partners allowed testing to take place at eight facilities of major air carriers. These carriers were chosen to represent the USA from east coast to west coast, from northern to southern states, including the midwestern region. All the AMTs who participated in the study were volunteers, and were assured of anonymity.

One hundred seventy-five licensed AMTs, all with Airframe and Power Plant licenses, from eight major air carrier maintenance sites were tested. The age distribution of this sample shown in (Figure 15-1, appendix) and the AMT experience distribution is shown in (Figure 15-2, appendix). Mean age was 37.7 years, and mean experience 13.2 years. The data from our sample can be compared with demographic data on aircraft mechanics (in all branches of aviation) compiled for 1988 by the Bureau of Labor Statistics (BLS) (Wash, 1991). (Table 15-2, appendix) shows this data comparison. (Table 15-3, appendix) summarizes other characteristics of the AMT sample used in our study.

Wilcoxon tests of the median age in our sample shows that it was not significantly different from the BLS data ($t = 7879$, $p > 0.50$). For the experience distribution, the sample median was significantly greater than the BLS data ($t = 10{,}142$, $p < 0.001$), showing that the AMTs in our study were more experienced than the AMTs of the earlier data. In particular, there were far fewer AMTs with three years or less experience, a finding probably representing reduced hiring patterns in major airlines during the 1990s.

(Figure 15-3, appendix) shows the distribution of scores on a reading ability test -- the Accuracy Level Test. Scores are equivalent to reading grade levels. Carver (1987) provides data for this test for two appropriate comparison groups: freshmen undergraduate and beginning graduate students. The mean score of our sample (13.35) was significantly higher than for college freshmen (12.5) with $t(174) = 6.95$, $p < 0.001$. However, it was significantly lower than for graduate students (14.3) with $t(174) = -7.85$, $p < 0.001$. Thus the reading level of our AMT sample was typical of an educated adult group, i.e., above college freshmen but below graduate students. The mean was less than one grade level different from either group, showing that while the differences may be significant, they are not large in absolute terms.

Evaluation Procedure

All the testing took place at airline maintenance facilities, in whatever room was made available. AMTs were tested individually or in groups depending upon their arrival times. Each AMT was given written instructions for completing a demographic questionnaire, a reading comprehension test, the actual workcard comprehension task and a set of workcard rating scales.

Each AMT was given one of the 16 possible workcards, i.e., four complexity levels each in four versions. Workcards were distributed in order, with a different starting point at each carrier. For the comprehension test, each AMT was given the workcard and a set of questions (20 each for three workcards, 19 for the other). Generally, a question concerning specific technical information was followed by a question asking where this information was located in the workcard. The questions demanded either a short answer, a "fill in the blank," or a multiple choice. Because the four workcards represented different procedures, there was no way to match the individual questions across workcards, *i. e.*, ensure that the same questions were asked across all workcards.

Although there were four different versions of each workcard, there was only one version of the comprehension test to eliminate any bias in constructing or wording this test. Also, in some cases, different words were used in Simplified English and Non-Simplified English workcards to refer to the same object. In this case a neutral word with similar meaning was chosen in order to prevent bias. For example, in Simplified English, the term "Do-Not-Operate tag" was used to indicate a card that was placed on an inoperative control lever, whereas in the Non-Simplified English workcard the term "Do-Not-Operate identifier" was used. In this particular case, questions regarding these cards used the term "Do-Not-Operate marker".

The dependent variables measured were defined as follows:

1. Demographic Variables: Age, experience as AMT, experience with different aircraft types, native language.

2. Reading Comprehension: The Accuracy Level Test (Carver, 1987). This was a ten-minute timed vocabulary test which measured the reading level of the AMTs as an equivalent grade level. This test has high reliability (0.91) measured on college students (Carver, 1987) and has a high validity (0.77 to 0.84) when compared to a longer standard reading test (the Nelson-Denny Reading Test).

3. Workcard Comprehension: Accuracy score on comprehension test, called "Test Completion Accuracy" and given as the percentage of correct answers combining accuracy of answers and accuracy of locating the answer in the workcard. Time taken to complete the reading of the workcard and the comprehension test, called "Test Completion Time" and measured with a stopwatch.

4. Rating Scales: Rating scale responses were based upon the evaluation scales used by Patel, et al (1994). They covered ease of use of the workcard and its graphics attachments, the simplicity of the English used and, finally, an overall rating of workcard usability. All were nine-point scales (0 to 8) anchored at each end with an appropriate adjective, and with their midpoints located at a scale value of 4.5.

Experimental Design

This was a three-factor factorial experimental design with AMTs nested under all three factors. The factors were:

- Language at two levels:
  - Simplified English
  - Non-Simplified English
- Layout at two levels:
  - Original layout
  - Patel, et al. (1994) layout
- Workcard complexity at four levels:
  - Easy 1

- Easy
2

- Difficult
1

- Difficult
2

RESULTS

All of the data from each subject were coded using the ACCESS(tm) program, and brought into MINITAB(tm) for statistical analysis. There were three main groups of variables:

1. Independent Variables.

   Language

   Layout

   Workcard complexity

2. Dependent Variables.

   - *Performance Measures on comprehension test:*

     Test completion time

     Test completion accuracy

   - *Ratings of workcard:*

     15 rating scale results

3. Possible performance predictors or co-variates.

   Age

   Experience as mechanics

Experience with different aircraft

Native language

Reading comprehension score

In this report, no distinction was made between scores on correctness of answers and correctness of location. Only a single score was derived, called test completion accuracy (or just "accuracy"). The first analyses presented here assess statistically the effects of the three independent variables on the two performance measures, using selected performance predictors as co-variates. Subsidiary analyses explore the role of some of the co-variates further, e.g., native language and experience with different aircraft. These analyses are followed by those of the effects of the independent variables on rating scale scores. All analyses used analysis of variance or covariance procedures, specifically the General Linear Models technique which allows for unequal sample sizes between conditions. Statistical significance is defined here as odds of greater than 1 in 20 against an effect having arisen by chance ($p < 0.05$).

Performance Effects

For the analysis of the three major factors, a co-variate was used to reduce the expected variability between individual AMTs. The four possible individual variables which may affect performance, and therefore could be useful co-variates, were: AMT experience, inspection experience, age and reading ability score. An intercorrelation matrix of these and the two performance variables (time, accuracy) showed that inspection experience was uncorrelated with other variables and that AMT experience was highly correlated with age. The other two variables, age and reading ability were moderately correlated with time and accuracy. Correlation coefficients were calculated as 0.217 between age and task completion time and -0.158 between age and task completion accuracy. Age and reading ability were tested, singly and together, as co-variates, and gave almost identical results. Only the analyses using age as a cavort are presented here for simplicity.

(Table 15-4, appendix) shows a summary of the significant effects for task completion time and task completion accuracy. The cavort (age) was significant in each case showing that times increased and accuracy decreased with increasing age. Both performance measures (time and accuracy) showed a significant workcard effect and a significant interaction between Simplified English/non-Simplified English and workcard, as shown in (Figure 15-4, appendix) and (Figure 15-5, appendix). For times, Figure 15-4 shows that each workcard had a somewhat different effect of Simplified English. Workcards Easy 1 and Difficult 2 gave slower performance times, and the others faster performance times. However, for accuracy (Figure 15-5, appendix) the effects were much clearer. For the two Easy workcards, there was no significant change in accuracy between Simplified English and non-Simplified English versions, but for the two Difficult workcards, Simplified English gave clearly superior accuracy.

To determine whether the AMTs' experience on Boeing aircraft had an effect on their performance on the comprehension test, a factor of whether or not each AMT had worked on Boeing aircraft in the past two years was added to the ANCOVAs of time and accuracy. No main effect or interaction with Boeing experience was found to be significant.

In the Shubert, et al. (in press) study it had been found that SE was of greatest benefit to non-English speakers, so that a similar test was appropriate in our study. Of the 175 AMTs tested, 157 were native English speakers and only 18 non-native English speakers. Because there was an even distribution of the 16 workcards to AMTs, nine non-native English speaking AMTs were given SE workcards and nine non-SE workcards. The number of non-native English speakers was too small for this characteristic to be used within the main ANCOVA, either as a co-variate (Boeing experience), or as a fourth factor. Hence, a separate ANOVA was performed with only two variables, each at two levels:

Language of workcard: SE or non-SE
Native language: English or non-English

(Table 15-5, appendix) shows the significance of the main effects and their interaction for task completion time and task completion accuracy. Only the AMT's native language affected task completion time significantly. Native English speakers took an average of 20.5 min. while non-native English speakers took longer, an average of 24.7 min. to complete the comprehension test. Accuracy was different between the two types of English, between native and non-native English speakers, and for the interaction of these two factors. (Figure 15-6, appendix) shows all of these effects. There was a clear superiority for Simplified English, with accuracy increasing from 76% to 86% overall. Equally important is the finding that the effect was most marked for non-native English speakers, where the improvement in accuracy was from 69% to 87%. Indeed, Simplified English allowed non-native English speakers to achieve about the same level of performance as native English speakers. Performing multiple comparisons among the four means in (Figure 15-6, appendix) shows that only the differences between the lowest mean (non-SE/non-native English speakers) and the other three were significant at $p = 0.05$. Thus, the scores for both groups of native English speakers and the SE non-native English speakers were essentially the same, *i.e.*, use of SE brought the non-native English speakers up to the some accuracy as native English speakers.

Rating Scale Analyses

In the Patel, et al (1994) study, the rating scales were used to compare new and old workcard layouts. For such a simple comparison, a non-parametric statistical test could be used. However, the current study had a more complex multi-factor experimental design so that analyses of variance or covariance were the only feasible statistical analyses. This meant that the rating scales had to be assumed to produce normally-distributed responses. Histograms of the responses to each scale were plotted and no marked departures from normality noted.

In the rating scale data there were few significant effects noted in the ANCOVAs. (Table 15-6, appendix) shows the significance levels for the main effects; only a single interaction was significant. There were significant layout differences for six of the fifteen rating scales, all in favor of the original rather than the Patel, et al (1994) layout. (Table 15-7, appendix) shows the mean ratings for these significant measures. Of the four ratings which gave significant workcard effects, (Table 15-8, appendix) shows that AMTs gave low ratings to the Difficult 2 workcard on the measures listed in the table. Amount of graphics information and simplicity of English used were both rated close to the center of the scale for all workcards. The single significant interaction was workcard × SE/non-SE for the overall rating ($p = 0.027$). Of the two Easy workcards, one had SE rated better than non-SE while the other was reversed. For the Difficult workcards, both had the SE version rated better overall. No clear pattern emerges from this significant interaction.

DISCUSSION

This large and realistic study measured the effects of SE across a range of AMT backgrounds, types of workcard and workcard layouts. The aim was to determine whether SE helps (or hinders) comprehension of workcard information, and whether it does so uniformly or mainly in particular circumstances. In doing so, it was intended to confirm and extend existing comprehension studies, and to make sound recommendations on the use of SE by the aviation maintenance community.

The major result was that SE was indeed useful, having a positive effect on comprehension accuracy without any consistent negative effect on the speed of performance. On a representative sample of 175 practicing AMTs from sites across the USA, it was accuracy which was impacted by SE, showing that performance changes with SE would be in the direction of error reduction. In this aspect, the current work mirrors that of Shubert, et al (in press), where comprehension, correctness and content location (accuracy measures) were also the affected outcome measures.

In terms of which factors interacted with the SE factor, again previous research was confirmed and extended. Both the native language of the AMT and the complexity of the workcard interacted with the SE/Non-SE factor. As (Figure 15-6, appendix)showed, the effect of SE was to improve the accuracy by about 2% for native English speakers, but by about 18% for non-native English speakers. If we consider error rates, the inverse of accuracy rates, the results look even more dramatic, as shown in (Table 15-9, appendix).

The conclusion is simple and direct: Simplified English workcards allowed non-native English speakers to achieve the same level of performance as native speakers; the non-Simplified English versions of the workcards did not.

An analogous effect was seen for the interaction between workcard complexity and Simplified English (Figure 15-3, appendix). The two Difficult workcards were the only ones where Simplified English made a significant difference. Again, in terms of error rates, we have the data in (Table 15-10, appendix).

Here, for the Easy workcards there was no difference between Simplified English and non-Simplified English, but for the Difficult workcard the errors were reduced by a third, for all users (native and non-native English speakers).

Overall, Simplified English proved to have the most effect where the most effect was needed, i.e. for those whose native language was not English and where the material was more difficult.

Because of our large and varied sample, this result is generalizable across a range of age and experience levels, and appears independent of the particular make of aircraft with which the AMT is familiar.

There were no effects of layout on performance. From the rating scale data, AMTs preferred the original workcard layout to one incorporating the Patel, et al. (1994) guidelines. This is contrary to the previous finding, but in fairness it should be pointed out that the original workcards in this study were much closer to meeting human factors guidelines than the originals in the Patel, et al. (1994) study. Here, all workcards used an easily readable typeface, laser printer output, high contrast and good paper stock. Given these improvements, AMTs may have preferred to see workcards with a more familiar layout. Also, in the Patel, et al. (1994) study, the inspectors used both types of workcards to perform an actual task on the aircraft. Perhaps rating of layout after a short comprehension test cannot be expected to give the same results as rating after use on an aircraft.

CONCLUSIONS

1. Aircraft manufacturers and technical operations departments in airlines can use SE for workcards and be confident that it will improve comprehension accuracy.

2. The effectiveness of SE is greatest where it is most needed: for non-native English speakers and for difficult workcards. Under more favorable conditions, i.e. with native English speakers and easier workcards, SE will not adversely affect performance.

3. Workcard layout differences had no effect on comprehension.

ACKNOWLEDGMENTS

The authors would like to acknowledge the active cooperation of many individuals and groups in this project. Special mention is made of ***Boeing*** *(Heather Holmback, Rick Wojcik, Paul Montague)*, to the ***University of Washington*** *(Jan Spyridakis, Serena Shubert)* and the ***AECMA Committee*** *(Kathleen Barthe)*.

REFERENCES

AECMA Simplified English Standard (1995). *A Guide for the Preparation of Aircraft Maintenance Documentation in the International Aerospace Maintenance Language,* AECMA Document PSC-85-16598, Belgium: The European Association of Aerospace Industries.

Bohr, E. (1978). Application of instructional design principles to nuclear power plant operating procedures manuals. In Easterby, R. and Zwaga, H. (Eds), *Information Design*, New York: John Wiley & Sons Ltd, 517-577.

Carver, R. P. (1987). *Technical Manual for the Accuracy Level Test*. REVTAC Publications, Inc.

Edworthy, J., Hellier, E. and Stanton, N. (Eds) (1995). Editorial: Warnings in research and practice. *Ergonomics*, **38(11)**, 2145-2154.

Klare, G. R. (1978). Readability and comprehension. In Easterby, R. and Zwaga, H. (Eds), *Information Design*, New York: John Wiley & Sons, Ltd, 479-495.

McLaughlin, G. H. (1966). Comparing styles of presenting technical information. *Ergonomics*, **9.1**, 257-260.

Ogden, C. K. (1932). *Basic English, A General Introduction with Rules and Grammar*. London: Paul Treber & Co., Ltd.

Patel, S., Drury, C. G. and Lofgren, J. (1994). Design of workcards for aircraft inspection. *Applied Ergonomics* 1994, **25(5)**, 283-293.

Shubert, S. K., Spyridakis, J. H., Holmback, H. K. and Coney, M. B. (in press). The comprehensibility of Simplified English in Procedures. *Journal of Technical Communications*, in press.

Wash, D. P. (1991). The changing workforce to the year 2000. In *Meeting Proceedings of the Fourth FAA Meeting on Human Factors Issues in Aircraft Maintenance and Inspection*, 25-42

APPENDIX

Table 15-1: Characteristics of the four workcards used

Workcard	Word Count	Words per Sentence	Percentage Passive	Flesch-Kinkaid
Easy 1	472	13 (lo)	13 (lo)	9.3 (lo)
Easy 2	254	8 (lo)	3 (lo)	8.6 (lo)
Difficult 1	554	19 (hi)	31 (hi)	10.8 (hi)
Difficult 2	491	17 (hi)	25 (hi)	10.4 (hi)

Table 15-2: Comparison of sample age and experience to 1988 Bureau of Labor Statistics data.

Measure	Range, years	Percent All Av. Mtc	Percent Current AMT Sample
Age	29	31.3	21.8

	30-49	50.7 63.8
	75	18.0 14.3
Experience	<3	22.9 4.6
	4-9	28.5 42.9
	10-19	16.2 32.6
	>20	32.5 20.0

	All Av. Mtc Median	**Current AMT Sample Median**
Age	36.2	35.0
Experience	9.4	10.0

Table 15-3: Characteristics of AMT sample used in our study.

Question	**Answer**	**Percentage**
Type of Maintenance:	Line	12%
	Hangar	73%
	Shop	14%
	Other	2%
Native Language	English	90%
	Other	10%
Worked in past two years on:	Boeing	87%
	McDonald-Douglas	76%
	Airbus	22%
	Other	14%

Table 15-4: Significance levels for all factors and interactions in GLM ANOVA with age as a covariate.

	Performance Measure	
Factor	**Task Completion Time**	**Task Completion Accuracy**
Age (covariate)	p = .001	p = 0.006
Workcard (W)	p = .001	p = 0.004
SE/Non-SE (S)	not significant	not significant (p = 0.073)
Layout (L)	not significant	not significant
W x S	p = .001	p = 0.024
W x L	not significant	not significant
S x L	not significant	not significant
W x S x L	not significant	not significant

Table 15-5: Significance levels for SE/Non-SE and native/non-native language effects.

Performance Measure

Factor	Task Completion Time	Task Completion Accuracy
Language of workcard	not significant	$p < 0.001$
Native language	$p = .010$	$p = 0.043$
Interaction	not significant	$p = 0.011$

Table 15-6: Significance levels of main factors for rating scale data.

Measure	Workcard	SE/Non-SE	Layout
1. Readability of text	not significant	not significant	not significant
2. Continuity of information flow	not significant	not significant	not significant
3. Ease of information location	not significant	not significant	$p = 0.046$
4. Chance of missing information	not significant	not significant	not significant
5. Ease of understanding	$p = 0.002$	not significant	not significant
6. Ease of location on aircraft	not significant	not significant	$p = 0.024$
7. Ease of relating figure numbers	not significant	not significant	$p = 0.01$
8. Amount of information provided	not significant	not significant	not significant
9. Ease of readability of attachments	$p = 0.038$	not significant	not significant
10. Relating graphics to aircraft	not significant	not significant	$p = 0.016$
11. Consistency of presentation	not significant	not significant	$p = 0.012$
12. Compatibility with attachments	not significant	not significant	$p = 0.001$
13. Amount of graphics provided	$p = 0.001$	not significant	not significant
14. Simplicity of English used	not significant	not significant	not significant
15. Overall ease of usability of w/c	not significant	not significant	not significant

Table 15-7: Mean ratings of both layouts for significant measures.

Measure	Original Layout	Patel, et al (1994) Layout
3. Ease of information location	5.5	5.0
6. Ease of location on aircraft	6.1	5.6
7. Ease of relating figure numbers	5.9	5.3
8. Amount of information provided	5.9	4.7
10. Relating graphics to aircraft	5.7	5.2
11. Consistency of presentation	5.8	5.3
12. Compatibility with attachments	5.9	5.1

Table 15-8: Mean ratings of workcards for significant measures.

Measure	Easy 1	Easy 2	Difficult 1	Difficult 2
5. Ease of understanding	5.8	5.6	5.8	4.7
9. Ease of readability of attachments	5.7	5.5	5.7	4.9
12. Compatibility with attachments	5.5	5.7	5.7	5.0
13. Amount of graphics provided	4.5	5.4	4.2	4.8

14. Simplicity of English used	5.1	5.3	4.8	4.8

Table 15-9: Error rates across native language for Simplified English and Non-Simplified English.

	Error Rates	
Speaker Type	**Non-Simplified English**	**Simplified English**
Native English Speakers (157)	17%	15%
Non-Native English Speakers (18)	31%	13%
Whole Sample (175)	18%	14%

Table 15-10: Error rates for easy and difficult workcards for Simplified English and Non-Simplified English conditions.

	Error Rates	
Workcard Complexity	**Non-Simplified English**	**Simplified English**
Easy Workcards	17%	19%
Difficult Workcards	18%	11%

Figure 15-1: Age Distribution of AMTs in Sample

Figure 15-2: AMT Experience Distribution

Figure 15-3: Distribution of Reading Ability Test Scores

Figure 15-4: TIME - Simplified English vs. non-Simplified English

Figure 15-5: ACCURACY - Simplified English vs. non-Simplified English

Figure 15-6: Simplified vs. non-Simplified English for native English speakers and non-Native English speakers