Aeronautical Mobile Communications Panel

Working Group D/10 Meeting

Honolulu, Hawaii

19 – 28 January 1999

How Complexity and Format of Air Traffic Control Instructions affect Pilot Recall

Presented by
IFALPA

Published by
Judith Bürki-Cohen, Ph.D.
Engineering Psychologist
USDOT-Volpe National Transportation Systems Center, Cambridge, Massachusetts

Summary

Dr. Bürki-Cohen specializes in air traffic control and high-speed rail automation, decision aids, and communication. She holds M.A. and Ph.D. degrees in experimental psychology from Northeastern University and degrees in German, English, French and Italian from the Interpreters' School in Zürich, Switzerland. She also has completed the Propaedeuticum of the Medical School at the University of Bern, Switzerland. Her professional affiliations include the Human Factors and Ergonomics Society and the Association of Aviation Psychologists.

Section 3.5 and Figures 5 & 6 show the percent errors and miscommunications for radio frequency. Although it is not explicitly noted in this paper, these tests were conducted in an 25kHz DSB AM environment and would probably show error rates significantly higher if it were done in an 8.33 kHz DSB AM environment using the R/T guidelines of Annex 10.

Say Again?
How Complexity and Format of Air Traffic Control Instructions affect Pilot Recall*

1.0 Introduction

1.1 Until recently, controllers were required to communicate all numerical air traffic control (ATC) information in sequential format, that is, digit by digit. For example, an altitude of 17,000 ft had to be transmitted as "Climb to one seven thousand." According to the latest versions of the Air Traffic Control handbook (7110.65G), however, controllers may restate altitude clearances in grouped format as "seventeen thousand" after giving them in sequential format. Although controllers appear to generally agree that grouping numbers improves recall, there is no direct scientific evidence to support this belief.

1.2 This study compared the recall of ATC information presented in either grouped or sequential **format** in a part-task simulation. It also tested the effect of **complexity** of ATC clearances on recall, that is, how many pieces of information a single transmission may contain without resulting in a readback error.

2.0 Procedure

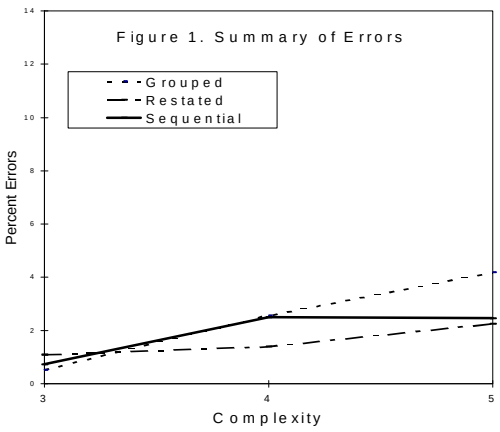
2.1 Twenty-four professional air-line pilots listened to an audiotape containing prerecorded ATC clearances for a low altitude en-route environment. They were tested individually and asked to read the entire transmission back and adjust the settings of a mock-up mode control panel. They could indicate uncertainty regarding a particular setting by pressing the adjacent "say again" button.

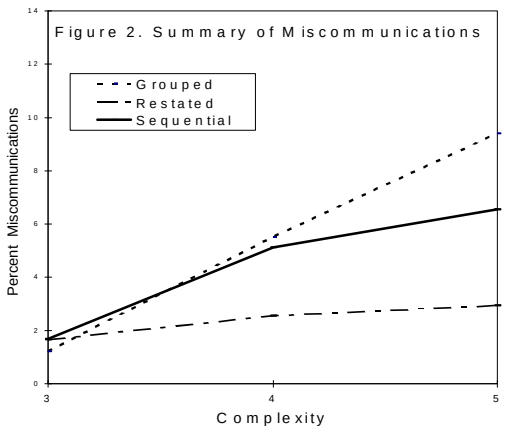
2.2 Clearance complexity ranged from three to five pieces of information. Altitude and frequency information were presented in grouped or sequential format, or both (i.e., restated). Altimeter and air speed were presented either in grouped or in sequential format. Heading was always presented in sequential format, potentially helping to distinguish it from speed, with which it shares part of its numerical range. Altimeter was presented only at the two higher complexity levels. Formats, complexity, order and type of information were carefully counterbalanced to avoid contextual or learning effects. To avoid expectancy effects, the clearances did not follow a coherent flight plan. The tape also contained “filler” clearances intended for other aircraft. All clearances were spoken by the same controller.

2.3 The results were analyzed in terms of type of information, recall errors, and miscommunications, which include both errors and requests for repeats. Errors include all instances where either readback or setting or both were incorrect, or both were omitted. All variables were analyzed as categorical in a two-way repeated measures analysis of variance (SAS proc glm). Where indicated, Bonferroni post-hoc analyses at a significance criterion of $p<.05$ were conducted.

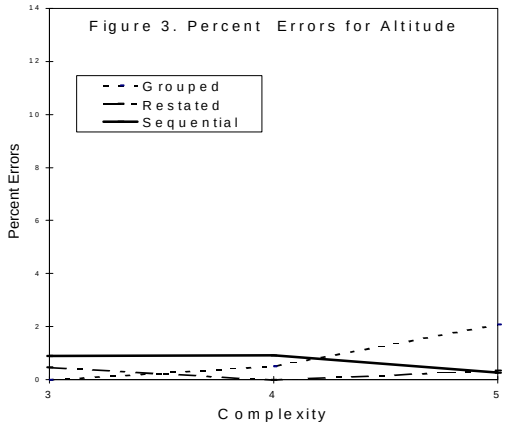
3.0 Results

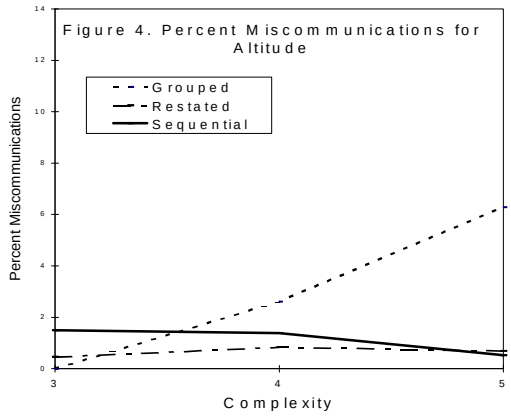
3.1 Figures 1 and 2 show the summary results. For errors, the effect of complexity was highly significant [$F(2,46)=9.94$; $p<.001$], whereas the effect of format just failed to reach significance [$F(2,46)=2.35$; $p>.10$]. The interaction was significant, however [$F(4,92)=2.99$; $p<.05$]. The post-hoc comparison showed that with 4.18 percent errors, grouped information was significantly less well recalled at complexity level five than sequential and restated information. Also, with a significant increase in errors going from complexity level three to four to five, grouped information revealed itself as sensitive to an increase in complexity. Restated information was not affected by complexity, and sequential only when going from complexity level three to higher levels.





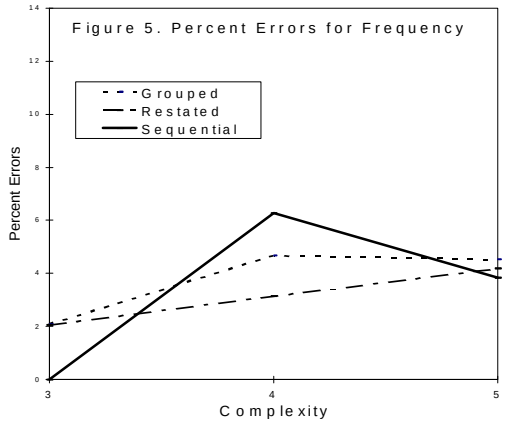
3.2 For miscommunications, both the effects of complexity and format were highly significant [F(2,46)=16.57 and 10.57; p<.001]. Their interaction remained significant [F(4,92)=6.04; p<.001]. The results of the post-hoc comparisons corresponded to the ones for errors, with an emerging advantage of restated over sequential information at complexity level five.

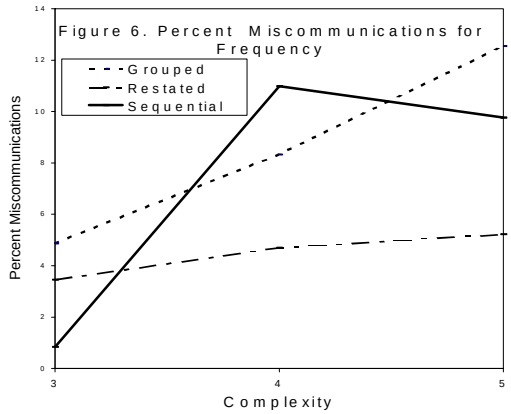




3.3 Figure 3 shows the percent errors for altitude. Error rates never exceeded one percent, except for grouped clearances at complexity level five with more than two percent errors. There were no significant overall effects of either format or complexity [F(2,46)=1.17 and .69]. There was, however, a trend of an interaction between format and complexity [F(4,92)=2.33; p<.10], indicating that with more data, effects might emerge for specific combinations of complexity and format.

3.4 A look at the miscommunications confirms this (Figure 4). Format did affect the miscommunication rate [F(2,46)=5.79; p<.01], with significantly more miscommunications for grouped altitudes at complexity level five than for restated and sequential altitudes. Although the overall effect of the complexity of the clearances was still not significant, the interaction between complexity and format was highly significant [F(4,92)=6.02; p<.001]. Miscommunications for grouped altitudes were higher at complexity level five then at the two lower levels, whereas miscommunications for sequential and restated altitudes did not increase with complexity.

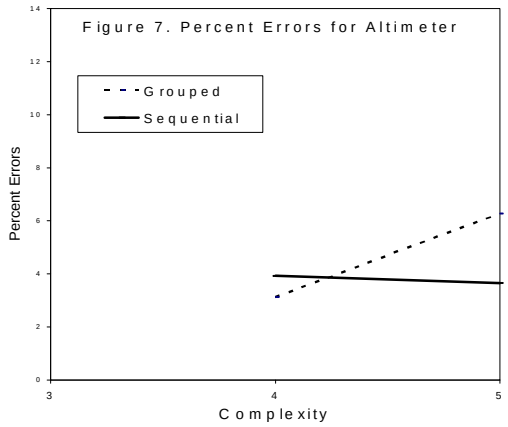


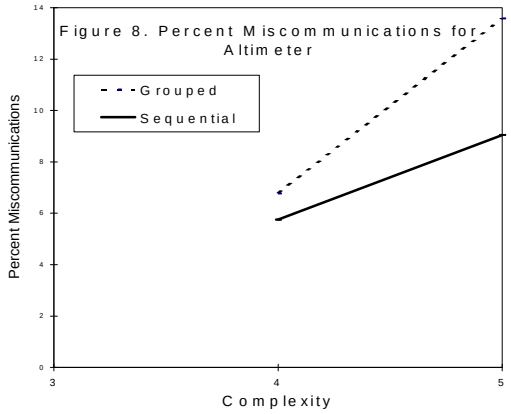


3.5 Figures 5 and 6 show the percent errors and miscommunications for radio frequency. Error rates significantly increased with complexity [F(2,46)=3.9; p<.05], but the format did not matter [F(2,46)<1], nor did format and complexity interact [F(4,92)=1.47]. Further analysis showed a significant increase in errors between complexity levels three and four, but not between four and five.

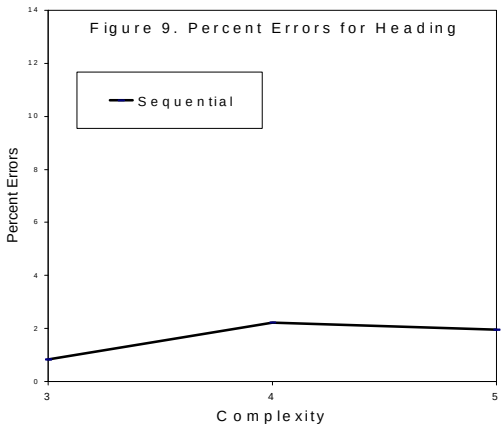
3.6 For miscommunications, both the effects of format and complexity were significant [F(2,46)=5.16 and 7.89; p<.01], as well as their interaction [F(4,92)=3.24; p<.05]. The post-hoc analysis showed an advantage of the restated format over the grouped format at complexity level five and over the sequential format at complexity level four. Also, miscommunications

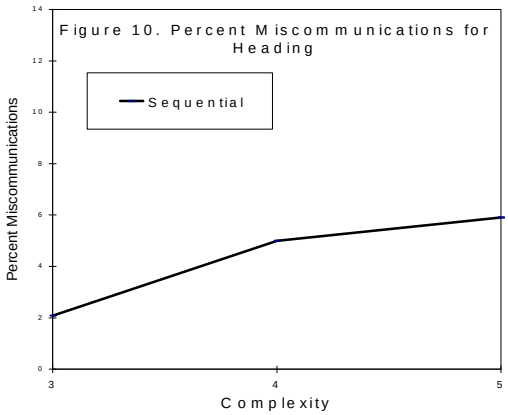
for restated frequencies did not increase with complexity. For grouped frequencies, miscommunications at complexity level five significantly exceeded those at complexity level three. For sequential frequencies, there were fewer miscommunications at complexity three than at both higher levels.



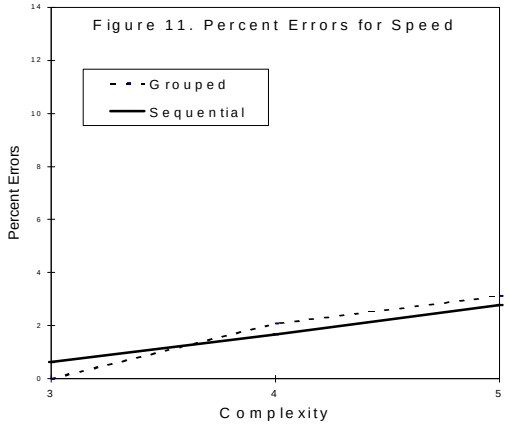


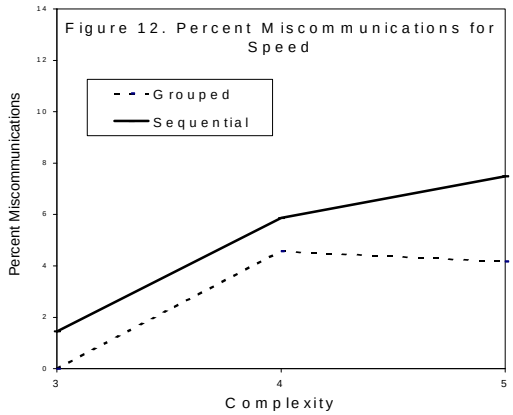
3.7 Figures 7 and 8 show percent errors and miscommunications for altimeter clearances. None of the effects for errors were significant, although there were trends for complexity effects and an interaction between complexity and format effects [F(1,23)=3.01 and 2.93; p=.10]. For miscommunications, however, complexity was highly significant [F(1,23)=17.03; p<.001], and both the effects of format and the interaction between format and complexity emerged as a trend [F(1,23)=3.43 and 3.77; p<.10]. Further analysis showed significantly more miscommunications at complexity level five than for any of the other conditions.





Figures 9 and 10 show the percent errors and miscommunications for heading clearances. Only miscommunications significantly increased with increasing complexity [F(2,46)=5.67; p<.01], and only when going from complexity level three to five.





3.8 Figures 11 and 12 show the percent errors and miscommunications for reading back speed information. Both errors and miscommunications increased significantly with increasing complexity [$F(2,46)=6.08$ and 12.45; $p<.005$], for errors between complexity levels three and five and for miscommunications for levels five and four compared to level three. Miscommunications were also significantly higher for speeds in sequential than for speeds in grouped format [$F(1,23)=8.26$; $p<.01$]. Neither of the interactions were significant [$F(2,46)<1$].

4.0 Conclusions

Five main conclusions can be drawn from this experiment.

1. Usually, an increase in complexity of the clearance did reduce pilot recall, especially in the grouped format.
2. With one exception, grouping numbers resulted in reduced recall compared to restating them or saying them digit by digit, especially at complexity level five.
3. For speed information, grouping helped, reducing requests for repeats. Approximately 10 percent of all speed errors were confusions with heading information. Grouping speed may help reduce confusions with heading, which is always presented sequentially. Before recommending that speed information should be grouped, however, the effects of cockpit noise, transmission quality, and speech rate of the controller have to be investigated (the latter is in progress). Also, numbers are grouped differently in other languages, which might represent a problem for foreign pilots.
4. Restated information resulted in fewer miscommunications than sequential information, at least at complexity level five. This is presumably independent of the format used for restating the information.
5. The error and miscommunication rates across types of information varied considerably. This may be a function of the importance as well as the number of digits and the numerical range of the information.

Acknowledgments

Dr. Kim Cardosi of the Volpe Center originated this work. Of the MIT Flight Transportation Laboratory, Prof. Robert Simpson recruited the airline pilots and Heather Howard assisted throughout the project. John Giurleo of Bedford Systems built the part-task simulator. Dr. Peter Mengert and Dr. Robert DiSario of the Volpe Center helped with the statistical analyses. The work was funded by the Federal Aviation Administration's Office of the Chief Scientific and Technical Advisor for Human Factors. Many thanks to all of them, and to the patient airline pilots who participated in this study!

**Preliminary data have been presented at the FAA/NASA/DOD Workshop on Methods and Metrics of Voice Communications in San Antonio, Texas, May 1994.*