NASA CR-143836

A PILOT RATING SCALE FOR VORTEX HAZARD EVALUATION

Roger H. Hoh

Systems Technology, Inc.
Hawthorne, California 90250

June 1975

(NASA-CR-143836) A PILOT RATING SCALE FOR VORTEX HAZARD EVALUATION (Systems Technology, Inc.) 33 p HC $3.75 CSCL 01A

N75-24695

Unclas
G3/03 24223

Prepared for

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
Flight Research Center
Edwards, Calif. 93523

1. Report No. NASA CR-143836	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle A PILOT RATING SCALE FOR VORTEX HAZARD EVALUATION		5. Report Date June 1975
		6. Performing Organization Code H-894
7. Author(s) Roger H. Hoh		8. Performing Organization Report No. TR-1025-4/1
		10. Work Unit No.
9. Performing Organization Name and Address Systems Technology, Inc. 13766 So. Hawthorne Blvd. Hawthorne, California 90250		11. Contract or Grant No. NAS 4-1952
		13. Type of Report and Period Covered Contractor Report – Topical
12. Sponsoring Agency Name and Address National Aeronautics and Space Administration Washington, D. C. 20546		14. Sponsoring Agency Code
15. Supplementary Notes NASA Technical Monitor: Herman A. Rediess, NASA Flight Research Center		

16. Abstract

A pilot rating scale is developed for subjective assessment of hazard resulting from wake vortex encounter upsets. The development of the rating scale is based on a survey of 48 pilots regarding the semantic properties of various phrases and a choice of formats for the rating scale. The rating scale can be used to define a hazard/non-hazard boundary as well as to determine a measure of the hazard.

17. Key Words (Suggested by Author(s))	18. Distribution Statement
Pilot rating scale Wake vortex	Unclassified – Unlimited

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price*
Unclassified	Unclassified	33	

*For sale by the National Technical Information Service, Springfield, Virginia 22151

TABLE OF CONTENTS

LIST OF FIGURES

A PILOT RATING SCALE FOR VORTEX HAZARD EVALUATION

Roger H. Hoh

Systems Technology, Inc.

SUMMARY

A pilot rating scale for quantification of the hazard associated with a wake vortex encounter was developed. In addition to the rating scale, this report presents the background and rationale utilized in the development of the scale. This rating scale can be used not only to determine a measure of hazard, but also to provide a means for defining a boundary between hazardous and nonhazardous situations.

The phrases selected for use in the rating scale, as well as the structure of the rating scale, were based on the results of a survey of 48 pilots. The survey included material on the semantic properties of certain word combinations in addition to a choice of format (or structure) for a rating scale.

INTRODUCTION

The need for a new rating scale for vortex upsets stems from two facts. First, existing handling quality rating scales have very poor selectivity at the low or unacceptable end. Second, typical vortex encounters would tend to be rated near the low end of existing handling quality rating scales. Therefore, it seemed unwise to attempt to use these existing scales or even to try to modify them to fit our purpose. Instead, the approach taken was to define an entirely new rating scale. As a result of a survey of 48 pilots, a set of calibrated phrases was obtained in addition to an overwhelming preference for the structure (or format) of the scale. In this survey 59 phrases and two example rating scale structures were presented. The two structures were a decision tree scale (akin to the Cooper Harper scale) and a multiple scale rating system. The survey subjects were asked to rate the various phrases and to state their preference of structure. They were also encouraged to make any comments they considered pertinent.

BASIC CONSIDERATIONS

Qualitative measurements made by human observers are basically estimates which arise from subjective judgment. These estimates are said to lie along a psychological continuum. There are several degrees of sophistication of psychophysical scales (for example, see Ref. 1). A brief description of the four basic levels of psychophysical scales is given below.

- Nominal — This is the lowest level of scale. Its function is usually to identify certain objects such as numbering of football players or assignment of type or model numbers to classes of objects.

- Ordinal — Ordinal scales serve to group objects as greater than or less than in terms of some representative measure.

- Interval — Interval scales are simply ordinal scales which are linear. That is, the distance between points on the scale consist of equal intervals.

- Ratio — A ratio scale allows the determination of the equality of ratios. Absolute temperature is an example of a ratio scale. Such a scale implies that an absolute zero exists.

The primary objective in the present work is to derive a rating scale that will allow certain mathematical operations to be performed on the pilot rating data. In particular, we would like to average the pilot rating data over several pilots as a measure of the overall hazard caused by the wake vortex upset. This implies that the scale must be linear, and from the above definitions it must therefore be interval as a minimum. The scale may or may not be adjectival, that is, we may pick a scale which is simply interval without words or phrases associated with each of the numbers indicated along the scale. Such scales are termed non-adjectival and have been shown to be quite successful especially for non-experienced raters (see Ref. 3). A comparison of adjectival and non-adjectival scales for wake vortex upset hazard is shown in Fig. 1. The main problem with a non-adjectival scale is that physical interpretation of the resulting ratings is almost completely lacking. The adjectival scale in Fig. 1 is one step above a non-adjectival scale in that

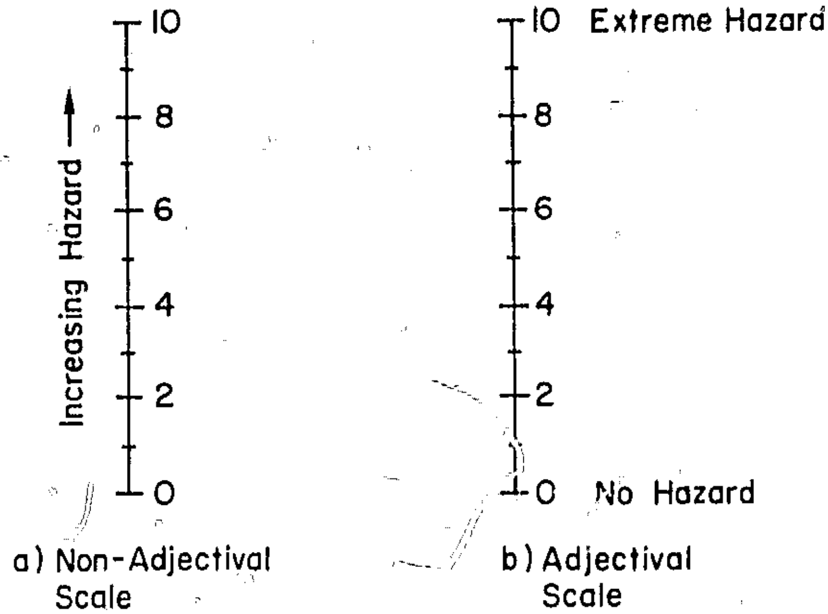


Figure 1. Comparison of adjectival and non-adjectival interval rating scales.

at least the end points are labeled. Physical interpretation of the hazard with this scale while somewhat better than the non-adjectival scale is still less than desirable. For example, it would be difficult to decide at what point on this scale the pilot rating would represent an unacceptable vortex encounter.

To summarize, the rating scale must be interval (or ratio) to allow averaging of the pilot ratings, and must be adjectival to allow physical interpretation of the pilot rating in terms of wake vortex encounter hazard.

Adjectival Scales

The formulation of an adjectival scale involves a basic tradeoff between increasing the number of word descriptors on the scale for physical interpretation, and maintaining linearity. The adjectives or phrases used to define various levels of hazard on the rating scale must have certain basic properties. First, they must relate directly to the area of concern (in our case, the hazard arising from the wake vortex upset). This is a measure of the validity of the adjective or phrase. Second, the adjectives or phrases must be selective, that is they must be distinguishable by the rater in terms of semantic differences in language. The standard deviation of the resulting ability of a group of raters to distinguish semantic differences in the phrases used on a given scale should be low, and should be approximately equal for all phrases. If the adjectives or phrases are to represent even increments on an interval scale, they must be distinguishable in equal amounts from the case where there is no hazard to the extreme hazard case. It is shown in Ref. 2 that this is essentially impossible, with the primary area of difficulty occurring at the low or unacceptable end of the scale (which happens to be the region of primary interest). The basic problem is that there is insufficient language to provide a constant sensitivity scale where the equal intervals are units related to noticeable semantic differences.

An attempt to quantify a large number of adjectives and phrases in terms of their selectivity and sensitivity was made in Ref. 2. A technique called the method of successive intervals was used to modify the raw data so as to achieve constant standard deviations (sensitivity) and constant intervals (selectivity) for each of the adjectives and phrases used in the study. One result of that study is shown in Fig. 2. The adjectives shown on this scale all exhibited a standard deviation of approximately one rating point (14% of full scale). With the exception of "bad," the adjectives fall on the scale in approximately equal intervals. Unlike the widely used Cooper-Harper handling quality rating scale (see Fig. 3), the adjectives in Fig. 2 are not assigned integer numerical ratings and instead are allowed to fall at the appropriate place along the scale. In this way, the rater has several reference points along the interval scale to relate the physical significance of the ratings to the numerical values. Note that there seems to be no word to satisfy a numerical rating in the region of 5. This scale represents a compromise between a purely adjectival and a purely interval scale.

The adjective "uncontrollable" in Fig. 2 does not fall on the interval scale as it represents an ultimate value or endpoint which cannot be averaged as a scale value. That is, "uncontrollable" could just as easily be rated a 15 as an 8. It is therefore treated as a discrete rating (subject checks box) to indicate that the ultimate in poor handling qualities has been achieved.

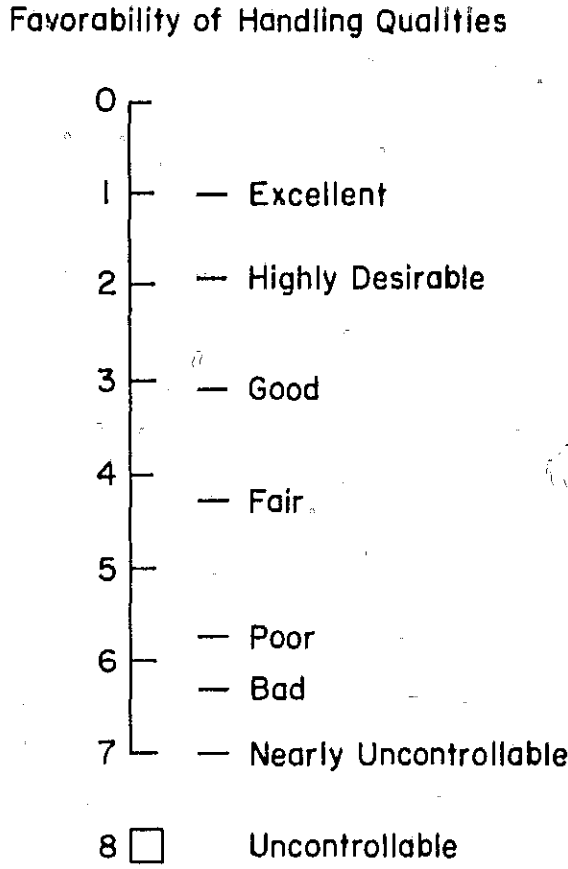


Figure 2. An interval rating scale with adjectival reference points.

Scaling Concepts

The pilot opinion rating for a given task is usually the result of the consideration of several factors. For example, when evaluating handling qualities on the Cooper-Harper scale, the pilot must consider aircraft characteristics, demands on the pilot, and adequacy for selected task or required operation. In deriving a rating scale for evaluating the hazard due to wake vortex upsets we have considered two possible approaches. The first was to use a "multiple scale rating system" where the pilot is asked to separately rate each element comprising the overall hazard. The second approach was to formulate a "decision tree" akin to that used in the Cooper-Harper scale (see Fig. 3).

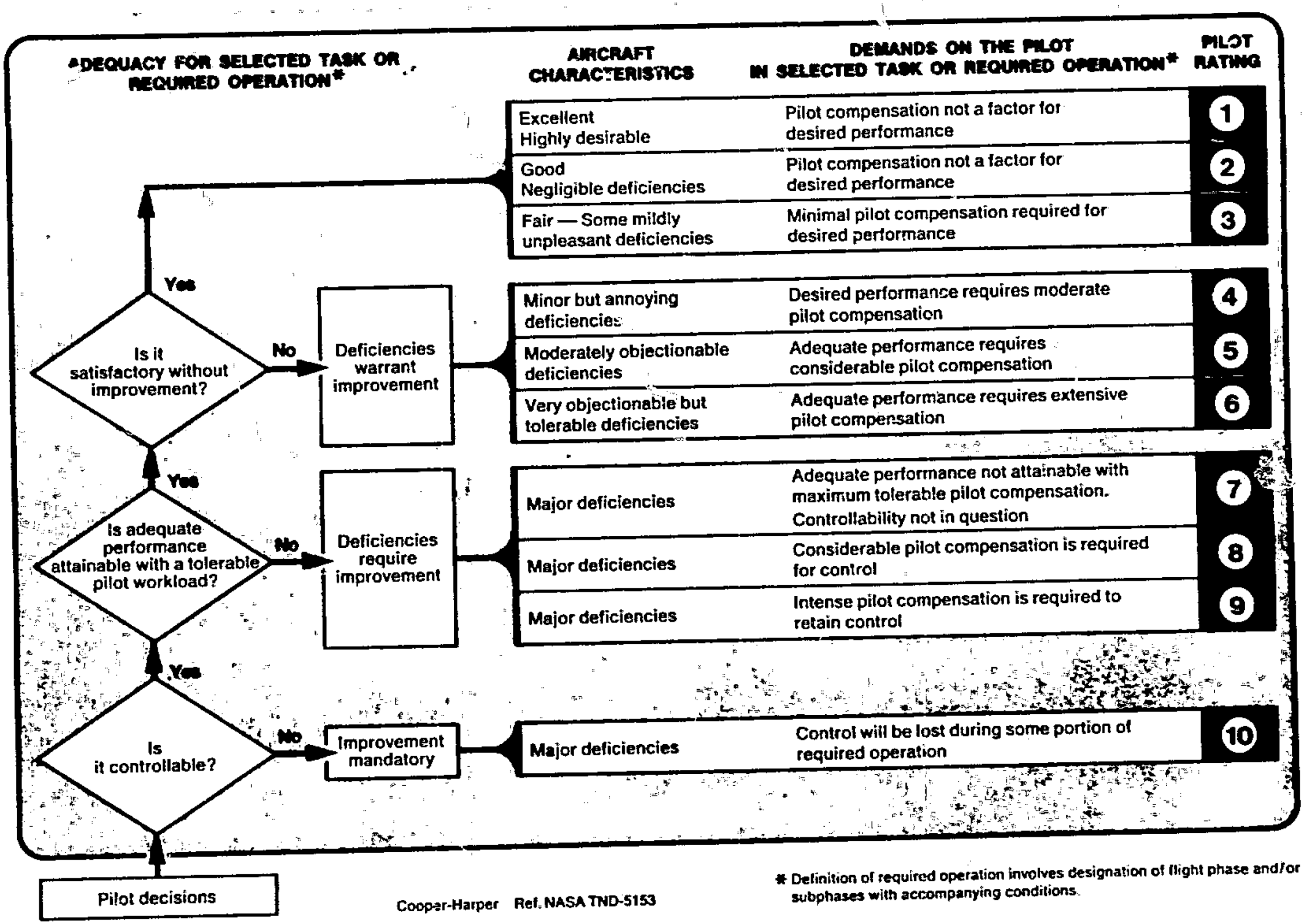


Figure 3. Cooper-Harper handling qualities rating scale (Ref. 5).

Multiple scale rating system. — Each of the components of a multiple scale rating system defines a separate rating boundary. The summation of these boundaries will define a composite boundary which accounts for all the pertinent factors which combine to define a hazard. This technique was used in recently completed automobile handling qualities experiments, Ref. 4. The boundaries had the general shape shown in Fig. 4 where it is seen that different rating factors define the minimum acceptable boundary for various values of the evaluation parameter. With a decision tree approach, the rater is forced to define these boundaries himself by internally evaluating and weighting each of the pertinent factors which make up the overall rating.

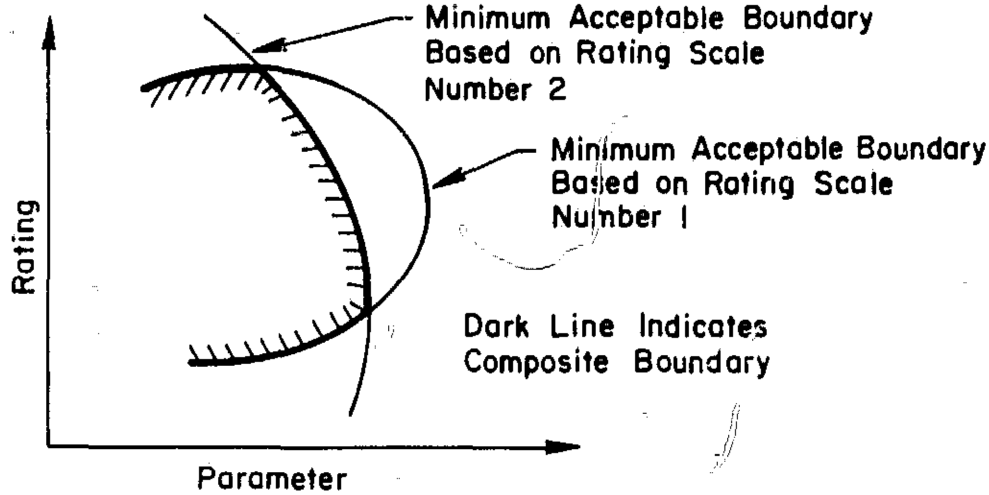


Figure 4. Minimum acceptable boundaries derived from multiple scale rating system.

Five tentative scales were initially proposed based on the assumption that the components which define a hazard for wake vortex encounters are as follows:

1. Chances for successful recovery
2. Aircraft excursions
3. Subjective opinion
4. Demands on pilot
5. Aircraft control

Decision tree rating scale. — The decision tree rating scale is probably best exemplified by the well known Cooper-Harper handling qualities scale shown in Fig. 3. Here the rater is asked to make initial yes/no decisions which will lead him to one of four categories. Three of these categories contain adjectives and phrases which describe aircraft characteristics and demands on the pilot. The rater must then decide which of these series of adjectives and phrases best describes the current situation. Thus, the rater must make the basic tradeoffs between the various factors which make up the overall rating. For example, the pilot may feel that deficiencies for a particular airplane "require improvement," but that the aircraft characteristics may be described as "very objectionable but tolerable." He is then forced to decide whether to give the airplane a 6 based on the latter or a 7 based on the former characterization. Another problem with this type of scale is that the larger number of adjectives and phrases which go into making up the final rating reduce the probability of obtaining a linear scale. However, because of the considerable success enjoyed by the Cooper-Harper handling qualities rating scale, it is felt that we should at least consider the possibility of using a decision tree type scale for the evaluation of wake vortex hazard.

Pilot Survey

Adjectives and phrases were formulated to indicate various levels of hazard in terms of each of the categories discussed in the following section. The selectivity and sensitivity of the adjectives and phrases were evaluated by a survey of 48 pilots who were asked to locate each phrase on a non-adjectival scale of increasing hazard. (See Appendix A.) Since the basic purpose of this evaluation was to determine the semantic properties of each phrase independently, the phrases were placed in random order. A provision was also made to allow the subjects to grade the suitability of each phrase on a scale from A to F. Finally, examples of a decision tree scale and a multiple rating scale system were presented using identical sample phrases and adjectives. The subjects were asked to state their preference.

An example of the survey questionnaire is given in Appendix A. The results are discussed in the following section of this report.

SELECTION AND VALIDATION OF FINAL SCALE

The results of the questionnaire in Appendix A are given in Appendix B where the tested phrases are listed under the following categories.

- Chances for Successful Recovery
- Aircraft Excursions
- Subjective Opinion
- Demands on Pilot
- Aircraft Control

Each of the phrases included in the questionnaire (Appendix A) were evaluated on a 7 point non-adjectival scale of increasing hazard. The means and standard deviations were calculated to determine the selectivity of each phrase. In addition, the average letter grades were computed to obtain an estimate of the acceptability of each phrase to the pilots as a group (see Appendix B). Those phrases which had means falling at approximately equal intervals and which exhibited the lowest standard deviations were selected for use on the rating scales. Where two or more phrases were approximately equal, the best letter grade was the determining factor.

The above "rules" were tempered with a certain amount of judgement in order to provide consistency on a given scale. For example, on the "aircraft excursions" scale, the word excursions appears in the selected phrases for ratings near 4 and 5 (see Fig. 5). Therefore, phrase number 1 in Fig. B-3 (negligible excursion) was picked over phrase number 2 (similar to light turbulence) even though number 2 had a slightly better average letter grade (B- compared to C+). Likewise, two phrases in the "demands on the pilot" category were shortened in order to emphasize the key adjectives (low, moderate, and extreme) describing the pilot effort.

One phrase was given twice in the questionnaire as a brief check on the validity of the results (phrases 34 and 52 in Appendix A). The results are given in Fig. B-2 where it is seen that "aircraft control required moderate pilot effort" received identical average values (3.0) and letter grades (C+) and reasonably close values of standard deviation (0.51 compared to 0.68).

Scale Format

The multiple scale rating system format was selected over the decision tree approach for the following reasons.

- 37 out of 47 pilots (one offered no preference) favored the multiple scale rating system format. (See Appendix A for scales presented for comparison.)
- Written commentary from pilot respondents indicated that most subjects felt that the multiple scale system was less confusing to use and would yield results that could be more easily interpreted.
- Written commentary from several pilot subjects indicated dissatisfaction with the decision tree scale because the single rating obtained with this scale assumes that the pilot agrees with all the items leading to any final branch.

Of the ten pilots who selected the decision tree scale on the questionnaire (Appendix A), at least two felt that their choice was at least partly due to a bias based on extensive use of the Cooper-Harper scale (Fig. 3). It is expected that acceptance of the multiple scale rating system by research pilots will generally be a matter of overcoming the bias.

Selection of Phrases

The category "chances for successful recovery" was found to be unacceptable for use on a vortex hazard rating scale. As shown in Appendix B, none of these phrases yielded an average rating less than 2.75. This would indicate that if recovery is even a question, the hazard level is considered to be quite high. Finally, the variability for these phrases was large compared to phrases in other categories.

The category "Subjective Opinion" was rejected because of very poor pilot acceptance (indicated by an average letter grade of D+). Many of the phrases in this category exhibited rather large variability among pilots.

"Aircraft Control" was a fairly good category yielding several phrases with low variability ($\sigma < 0.6$) and well separated means. The final rating scale is shown in Fig. 5 where the "Aircraft Control" column utilizes phrases 2, 9, 15, and 25 (from Appendix A). These phrases yielded an average standard deviation* of 0.51 and an average letter grade of B.

The second category utilized in the final rating scale was "Demands on the Pilot." The Fig. 5 phrases (numbers 32, 34, and 51) yielded an average standard deviation of 0.47 and an average letter grade of C+. Only three phrases in this category (out of 13) had means and standard deviations suitable for use on the final scale. However, since the purpose of the phrases is simply to provide anchor points on the interval scale, three is felt to be an adequate (minimum) number.

The category "Aircraft Excursions" utilizes phrases 23, 24, 48, and 56 resulting in an average standard deviation of 0.45 and an average letter grade of C+. Unfortunately, the phrases tend to have means that are somewhat unequally spaced. However, it is felt that having phrases with low semantic variability among pilots (low σ) was more important than equal spacing of the phrases on the scale.

Endpoints on the Scales

None of the phrases had means which fell at the extreme ends of the five point hazard scale in Appendix A. Considering that phrases such as "crash inevitable..." and "uncontrollable" were included, it was surprising that many pilot subjects did not feel that the "ultimate hazard" was defined. Likewise, such terms as "aircraft control was not a factor" and "negligible excursions" did not fall at the extreme low end of the five point hazard scale. These results are consistent with previous surveys (Ref. 2) and verify the subjective, comparative nature of rating scales (they have no associated absolute values). We have allowed for this in the final rating

*Average standard deviation is defined as

$$\sigma_{avg} = \left(\frac{1}{N}\sum \sigma^2\right)^{1/2}$$

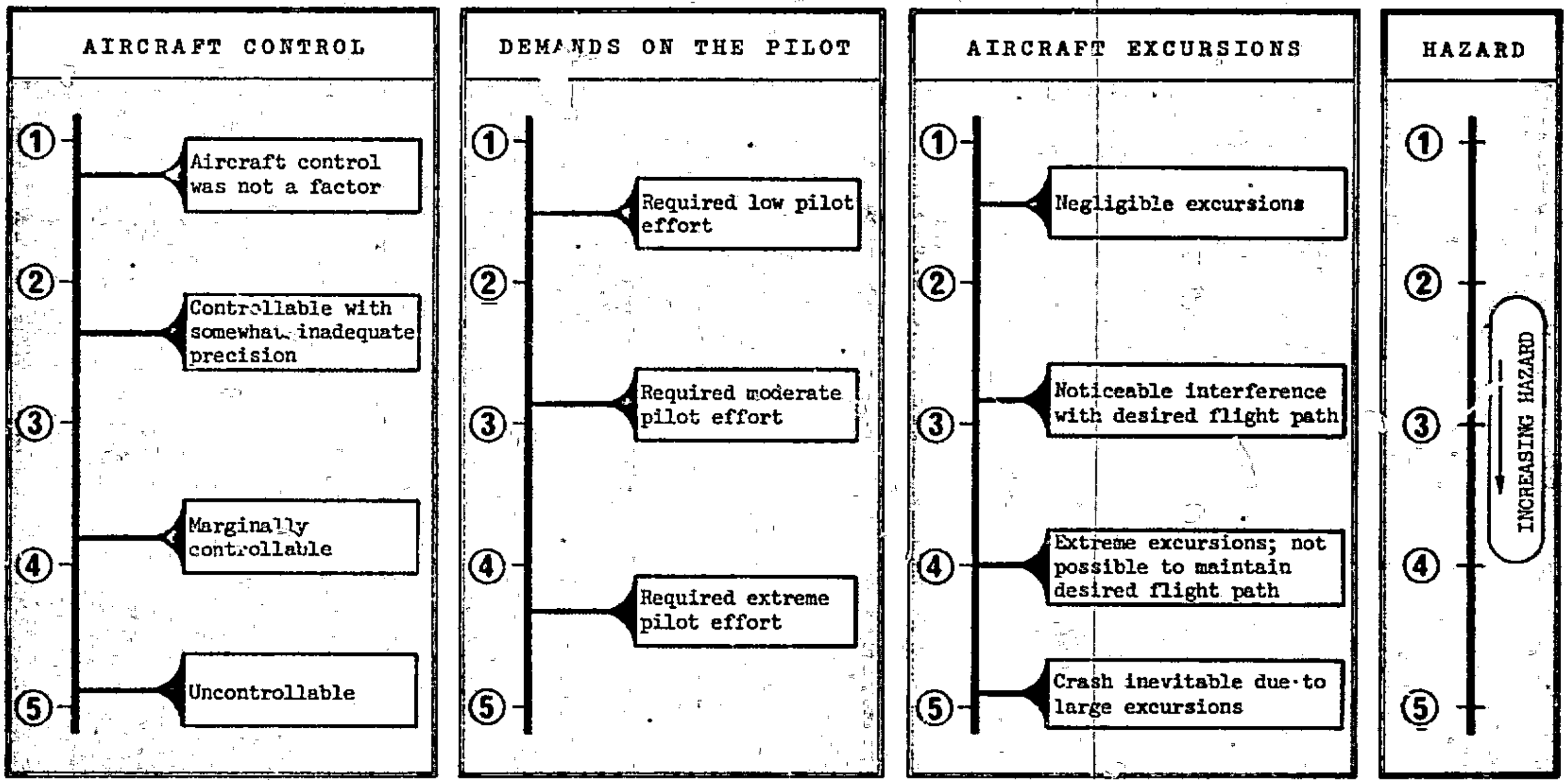


Figure 5. Vortex hazard rating scale.

scale (Fig. 5) by extending the scale below 1 and above 5. Thus, the rater can indicate that the appropriate rating does not fall in the range of 1 to 5 (shown by putting his mark at less than 1 or greater than 5).

Nonadjectival Hazard Scale

The phrases which comprise the adjectival rating scales in Fig. 5 are located along the scales according to their semantic value (mean) in terms of hazard. As with all pilot rating scales, the purpose of these phrases is to attach some physical significance to the scale values. A final scale which is nonadjectival (no phrases) has been added to provide a direct measure of the pilot's opinion of the hazard. This non-adjectival scale is identical to the scale used to evaluate the phrases in the survey. It allows correlation of the first three categories (adjectival scales) with the overall hazard without introducing additional semantic properties.

Experiments with non-adjectival scales (Ref. 3) have shown good sensitivity in discerning changes in the characteristics of systems that are very difficult to control, e.g., an unstable element. It therefore seems very appropriate to apply such a scale to wake vortex encounters where control is frequently marginal.

Validation

Validation of the vortex hazard rating scale can be obtained only through its use in research programs where actual vortex encounters are experienced and rated. The success or failure of any rating scale is its ability to produce consistent ratings within and across pilots. A logical measure of scale validity might therefore be the standard deviation of ratings across pilots for a given vortex encounter situation.

Another measure of validity is a comparison of the semantic meaning of the phrases utilized in the final scale between research pilots (who will be using the scale) and the entire pilot population. It is, of course, desirable that the phrases used have the same mean values (same coordinates on the hazard scale), similar (low) variability, σ, and reasonably good pilot acceptance (letter grades) across both pilot groups. A comparison of the means, standard deviations, and letter grades for the survey results from just research pilots (9) with the results of all the tested pilots (38)* is given in Fig. 6. The mean values are seen to be in good agreement indicating that the phrases used have the same hazard value for research pilots (users of the scale) and the total pilot group. Similarly, the standard deviations and letter grades are in very close agreement for the two pilot groups. A breakdown of the background of the pilots who responded to the questionnaire is given in Appendix B.

QUESTIONNAIRES

As discussed under Basic Considerations, questionnaires will be used to augment the rating data obtained. This will cover such things as pilot background, experience, and the subject's personal interpretation of the severity of the vortex encounter. The proposed questionnaire covering pilot background and experience is shown in Fig. 7. A tentative questionnaire to stimulate pilot commentary required to supplement the rating data is given in Fig. 8. It is intended to aid the experimenter in prompting pilot discussion after each vortex encounter run; and it may be altered as a result of such direct experience.

CONCLUSION

A rating scale has been developed to allow evaluation of the hazard associated with wake vortex encounters. The format used involves four separate ratings, each defining the hazard associated with a specific category. Three of the categories involve interval, adjectival scales, whereas the fourth scale is a non-adjectival measure of hazard. The adjectives and phrases selected for the scale were based on a survey designed to determine the meaning and variability in meaning of descriptive phrases across a group of pilots.

---

*Results from 10 of the respondents were not included in the statistical manipulations because of late arrival or misunderstood instructions.

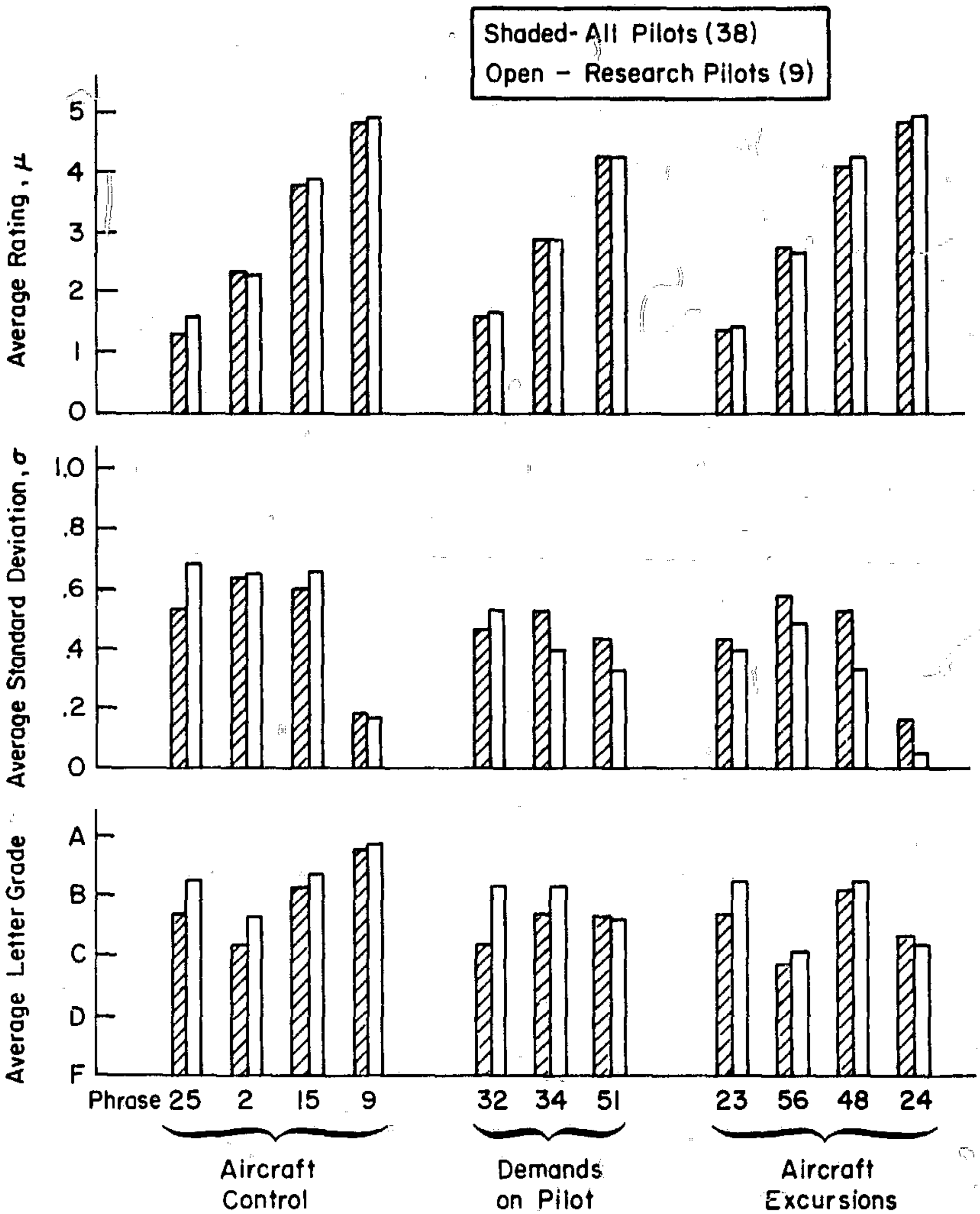


Figure 6. Comparison of research pilots with total population.

Name ______________________________ Date ______________

Occupation ______________________________ Age ________

Total Hours (Approximately) ______________________________

License and Ratings ______________________________

Type Equipment Flown:

- ☐ Military fighter
- ☐ Heavy aircraft
- ☐ Light twins
- ☐ Light singles
- ☐ Research simulator

Have you ever had a wake vortex encounter? _____ If so, briefly describe.

Figure 7. Pilot background.

1. Briefly describe the vortex encounter.

2. Did you consider the run hazardous?

3. Would you have continued on normal flight path or aborted (e.g., discontinue to climb or institute go-around) if this encounter occurred in flight?

4. If the upset was deemed as hazardous, was the primary hazard:
   a) Ground impact?
   b) Structural failure due to vortex?
   c) Structural failure due to recovery attempt?

Figure 8. Questionnaire after each run.
(Verbal response from subject.)

APPENDIX A

VORTEX HAZARD RATING SCALE SURVEY QUESTIONNAIRE

INSTRUCTIONS

The purpose of this questionnaire is to evaluate the meanings of various words and phrases which could be used to describe the hazard associated with a wake vortex encounter. The words and phrases are presented in random order on the next few pages. Alongside each word or phrase is a horizontal scale. We want you to place an X on the scale in a position that represents your impression of the hazard as defined by that particular phrase.

We are also interested in your overall impression of how useful each phrase might be if included in a rating scale. In other words, we want to know which phrases you think ought to be included in a rating scale. Therefore, please give each phrase a grade (in the space provided beside the horizontal scale) using the usual A through F system. That is, an A would imply that the phrase would be an excellent one to include in a rating scale, and an F would indicate that you feel the phrase is totally inappropriate for use on a vortex hazard rating scale. If you have any comments you would like to make, please feel free to mark up the page with whatever comes to mind.

		Increasing Hazard →	Grade
1.	Was disoriented — high probability of incorrect control application.	1 2 3 4 5	
2.	Controllable with somewhat inadequate precision.	1 2 3 4 5	
3.	Similar to moderate turbulence.	1 2 3 4 5	
4.	Barely controllable.	1 2 3 4 5	
5.	Considerable cause for alarm.	1 2 3 4 5	
6.	Easy to apply correct controls.	1 2 3 4 5	
7.	Required considerable pilot effort; improper use of flight controls possible.	1 2 3 4 5	
8.	Significant loss of control resulting in high probability of aborting task.	1 2 3 4 5	
9.	Uncontrollable.	1 2 3 4 5	
10.	Required moderate control to regulate against disturbance.	1 2 3 4 5	
11.	Disturbance resulted in momentary disorientation.	1 2 3 4 5	
12.	Was concerned about safety of flight during some portion of the encounter.	1 2 3 4 5	

	Increasing Hazard →	Grade
13. Aircraft excursions were large enough to cause concern for safety.	1 2 3 4 5	______
14. Serious loss of control.	1 2 3 4 5	______
15. Marginally controllable.	1 2 3 4 5	______
16. Required considerable pilot effort. Flight control usage for recovery straightforward.	1 2 3 4 5	______
17. Control was completely lost — feared for my life.	1 2 3 4 5	______
18. On the verge of losing control.	1 2 3 4 5	______
19. Considered it a "close call."	1 2 3 4 5	______
20. Brief period of loss of control. Little effect on flight progress.	1 2 3 4 5	______
21. High probability of a crash due to large excursions.	1 2 3 4 5	______
22. Extensive loss of control.	1 2 3 4 5	______
23. Negligible excursions.	1 2 3 4 5	______
24. Crash inevitable due to large excursions.	1 2 3 4 5	______

		Increasing Hazard →	Grade
25.	Aircraft control was not a factor.	1 2 3 4 5	____
26.	Disturbance easily corrected with minimum control.	1 2 3 4 5	____
27.	No effect on safety	1 2 3 4 5	____
28.	Very difficult to control.	1 2 3 4 5	____
29.	Completely demanding of pilot attention, skill, or effort.	1 2 3 4 5	____
30.	Similar to light turbulence.	1 2 3 4 5	____
31.	Required nearly full control to regulate against disturbance.	1 2 3 4 5	____
32.	Required low pilot effort.	1 2 3 4 5	____
33.	Feared for my life.	1 2 3 4 5	____
34.	Aircraft control required moderate pilot effort.	1 2 3 4 5	____
35.	Momentarily lost control — some concern for safety.	1 2 3 4 5	____
36.	Aircraft lost.	1 2 3 4 5	____

		Increasing Hazard →	Grade
37.	Complete loss of control — high probability of a crash.	1 2 3 4 5	______
38.	Nearly uncontrollable.	1 2 3 4 5	______
39.	Poor chance for successful recovery from upset.	1 2 3 4 5	______
40.	Very demanding of pilot attention, skill, or effort.	1 2 3 4 5	______
41.	Lost control for a significant period of time — cause for alarm.	1 2 3 4 5	______
42.	Extreme excursions; definitely an unsafe condition.	1 2 3 4 5	______
43.	Good chance for successful recovery from upset.	1 2 3 4 5	______
44.	Felt I was in trouble.	1 2 3 4 5	______
45.	Extreme concern for safety.	1 2 3 4 5	______
46.	Recovery impossible.	1 2 3 4 5	______
47.	Excellent chance for successful recovery from upset.	1 2 3 4 5	______
48.	Extreme excursions; not possible to maintain desired flight path.	1 2 3 4 5	______

	Increasing Hazard →	Grade
49. Demanding of pilot attention, skill, or effort.	1 2 3 4 5	______
50. Complete loss of control.	1 2 3 4 5	______
51. Required extreme pilot effort; proper control usage difficult and confusing.	1 2 3 4 5	______
52. Aircraft control required moderate pilot effort.	1 2 3 4 5	______
53. Slightly more severe disturbance would cause concern for safety.	1 2 3 4 5	______
54. A precarious situation.	1 2 3 4 5	______
55. Largely undemanding of pilot, relaxed.	1 2 3 4 5	______
56. Noticeable interference with desired flight path.	1 2 3 4 5	______
57. Incident not worthy of mention.	1 2 3 4 5	______
58. Felt I was in serious trouble.	1 2 3 4 5	______
59. Fair chance for successful recovery from upset.	1 2 3 4 5	______

One final item. The following two pages of this questionnaire show examples of two possible types of rating scales. The first type is called a "multiple scale rating system" because it allows the pilot to rate several aspects of the situation separately; and the second type is called a "decision tree scale" (for obvious reasons). We would like you to compare these two types of scales, and indicate which you prefer (and give reasons if you desire). Bear in mind that the particular phrases shown in these example scales are not important, as they will be revised based on this questionnaire. What we really want is to get your opinion regarding the basic structure ---- decision tree vs. multiple scale rating systems.

RATING SCALE PREFERENCE

My preference for a rating scale structure is:

Multiple Scale Rating System (Figure 1) . . . . ☐

Decision Tree Scale (Figure 2). . . . . ☐

Comments:

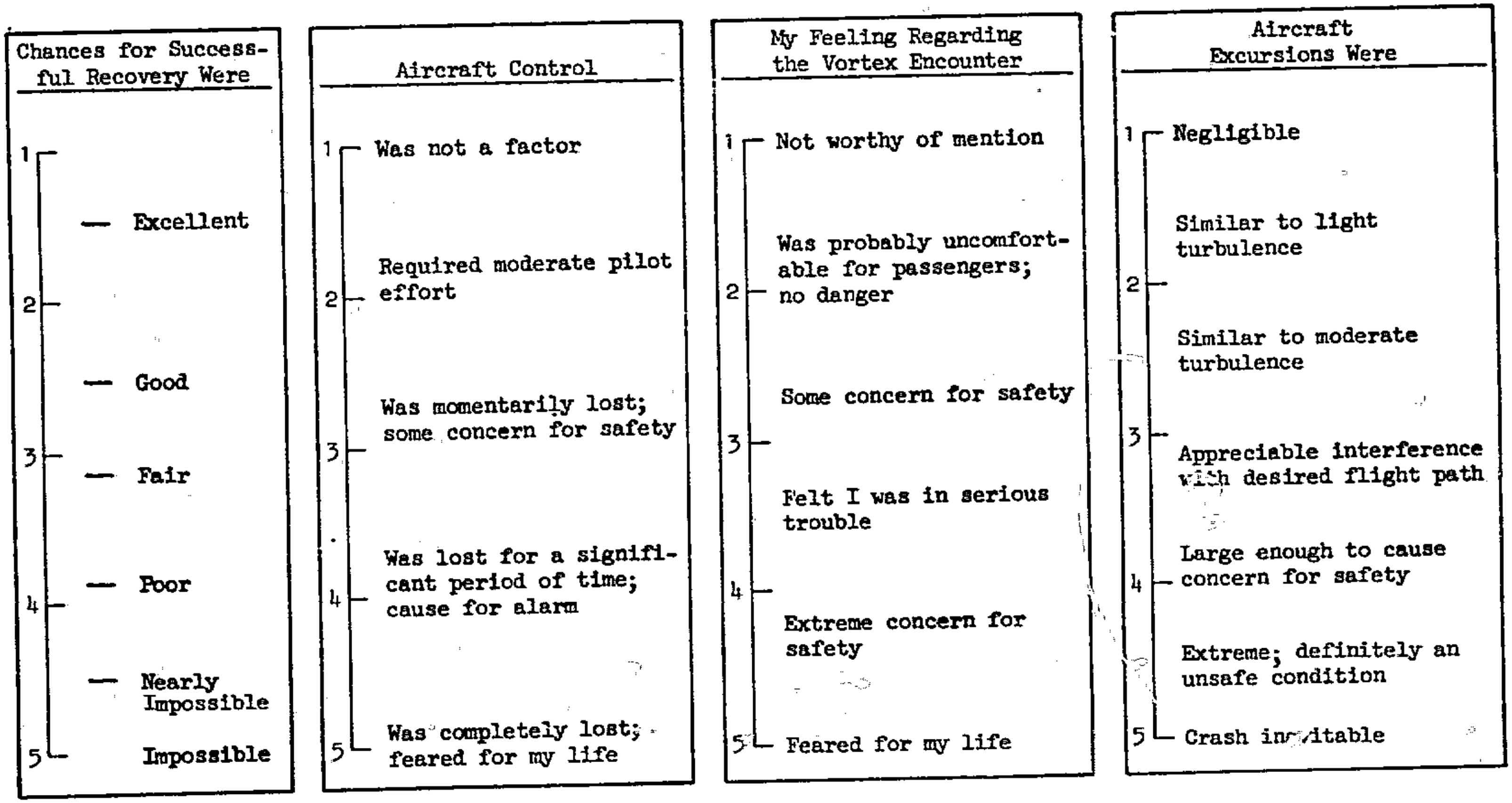


Figure A-1. Example of a multiple scale rating system for evaluation of vortex encounter hazard.

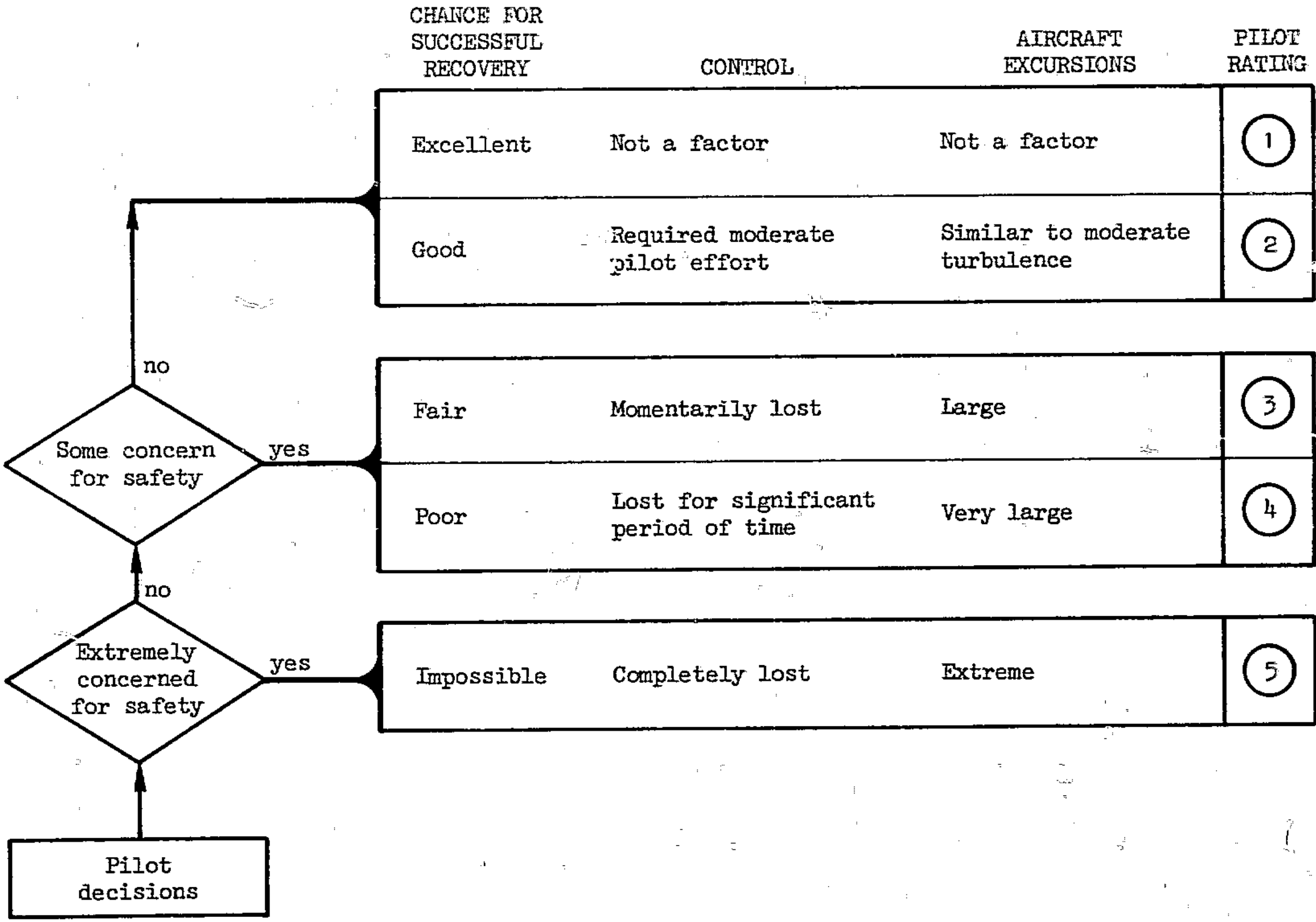


Figure A-2. Example of a decision tree scale.

APPENDIX B

RESULTS OF THE SURVEY

The results of the Appendix A survey are presented in Figs. B-1 through B-4. A total of 48 pilots filled out the questionnaires of which 38 were used to compute means, standard deviations, and an average letter grade for each phrase. Some questionnaires could not be used for data analysis either because they arrived too late to be included in the calculations or because the subject misunderstood the instructions. All of the questionnaires were utilized to decide on the scale format (decision tree vs. multiple scale rating system). A summary of the background of the pilot population responding to the questionnaire was

- 9 research pilots
- 5 airline pilots
- 17 FAA pilots
- 2 military pilots
- 12 nonprofessional pilots (all with engineering degrees)
- 3 flight instructors

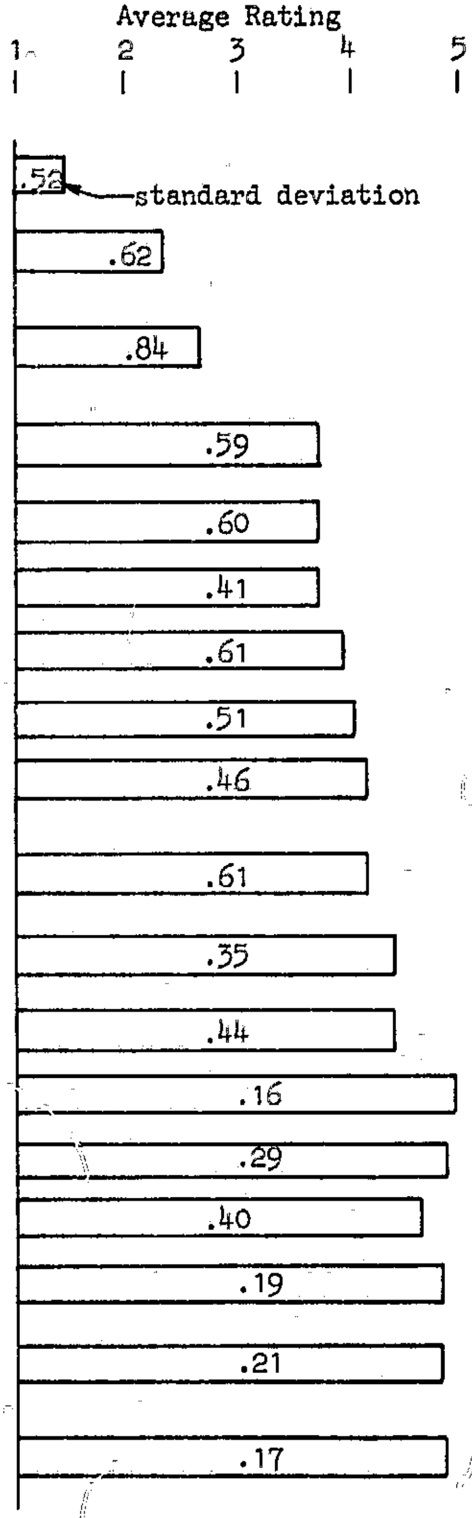
The rating scale format question received the following replies

- 37 favored the multiple scale rating system
- 10 favored the decision tree
- 1 had no preference

Five of the nine research pilots favored the decision tree which probably reflects their extensive use of the Cooper-Harper rating scale.

The general experience level of the respondents was very high with 24 ATP's (Airline Transport Pilot rating). Of the remaining pilots, 17 had a commercial license with an instrument rating. The average number of hours of flying time of all the respondents was 5922.

41 of the 48 respondents indicated that they had experienced one or more vortex encounters during their career. One indicated that the encounter resulted in a crash.



		Letter Grade	Standard deviation
1.	Aircraft control was not a factor.	C+	.52
2.	Controllable with somewhat inadequate precision.	C–	.62
3.	Brief period of loss of control. Little effect on flight progress.	C–	.84
4.	Momentarily lost control — some concern for safety.	C+	.59
5.	Marginally controllable.	B	.60
6.	Very difficult to control.	B–	.41
7.	On the verge of losing control.	C–	.61
8.	Barely controllable.	B–	.51
9.	Nearly uncontrollable.	C+	.46
10.	Significant loss of control resulting in high probability of aborting task.	D+	.61
11.	Lost control for a significant period of time — cause for alarm.	C	.35
12.	Extensive loss of control.	C–	.44
13.	Aircraft lost.	C	.16
14.	Complete loss of control.	B+	.29
15.	Serious loss of control.	B–	.40
16.	Uncontrollable.	A–	.19
17.	Complete loss of control — high probability of a crash.	B–	.21
18.	Control was completely lost — feared for my life.	C	.17

Figure B-1. Aircraft control.

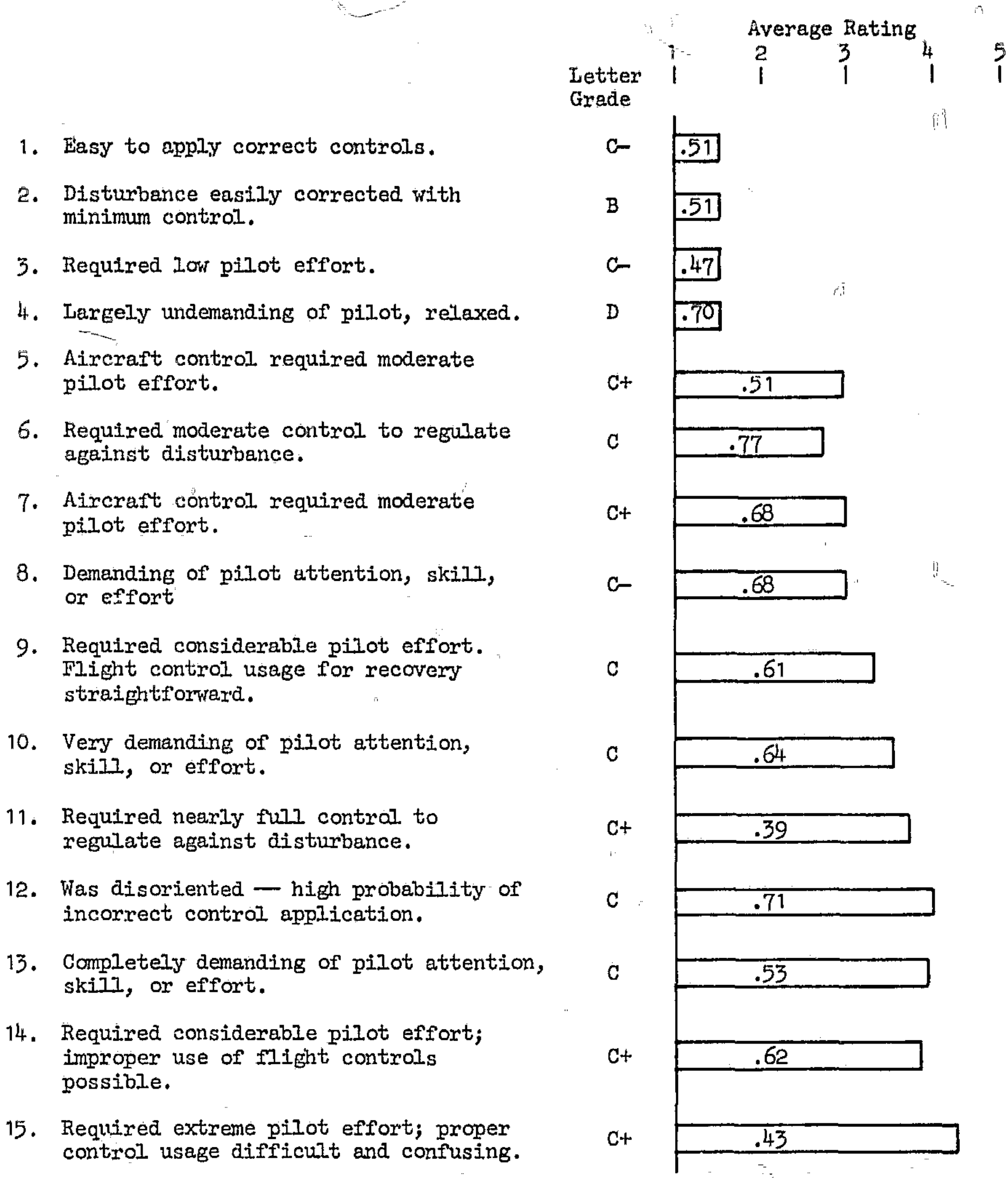


Figure B-2. Demands on pilot.

	Letter Grade	Average Rating (1–5)
1. Negligible excursions.	C+	.43
2. Similar to light turbulence.	B–	.43
3. Noticeable interference with desired flight path.	D+	.58
4. Similar to moderate turbulence.	C+	.70
5. Aircraft excursions were large enough to cause concern for safety.	C	.56
6. Extreme excursions; not possible to maintain desired flight path.	B	.52
7. Extreme excursions; definitely an unsafe condition.	B–	.47
8. High probability of a crash due to large excursions.	C+	.36
9. Crash inevitable due to large excursions.	C	.17

Figure B-3. Aircraft excursions.

	Letter Grade	Average Rating (1–5)
1. Excellent chance for successful recovery from upset.	C–	1.0
2. Good chance for successful recovery from upset.	D+	.85
3. Fair chance for successful recovery from upset.	D	.76
4. Poor chance for successful recovery from upset.	C–	.39
5. Recovery impossible.	B–	.76

Figure B-4. Chances for successful recovery.

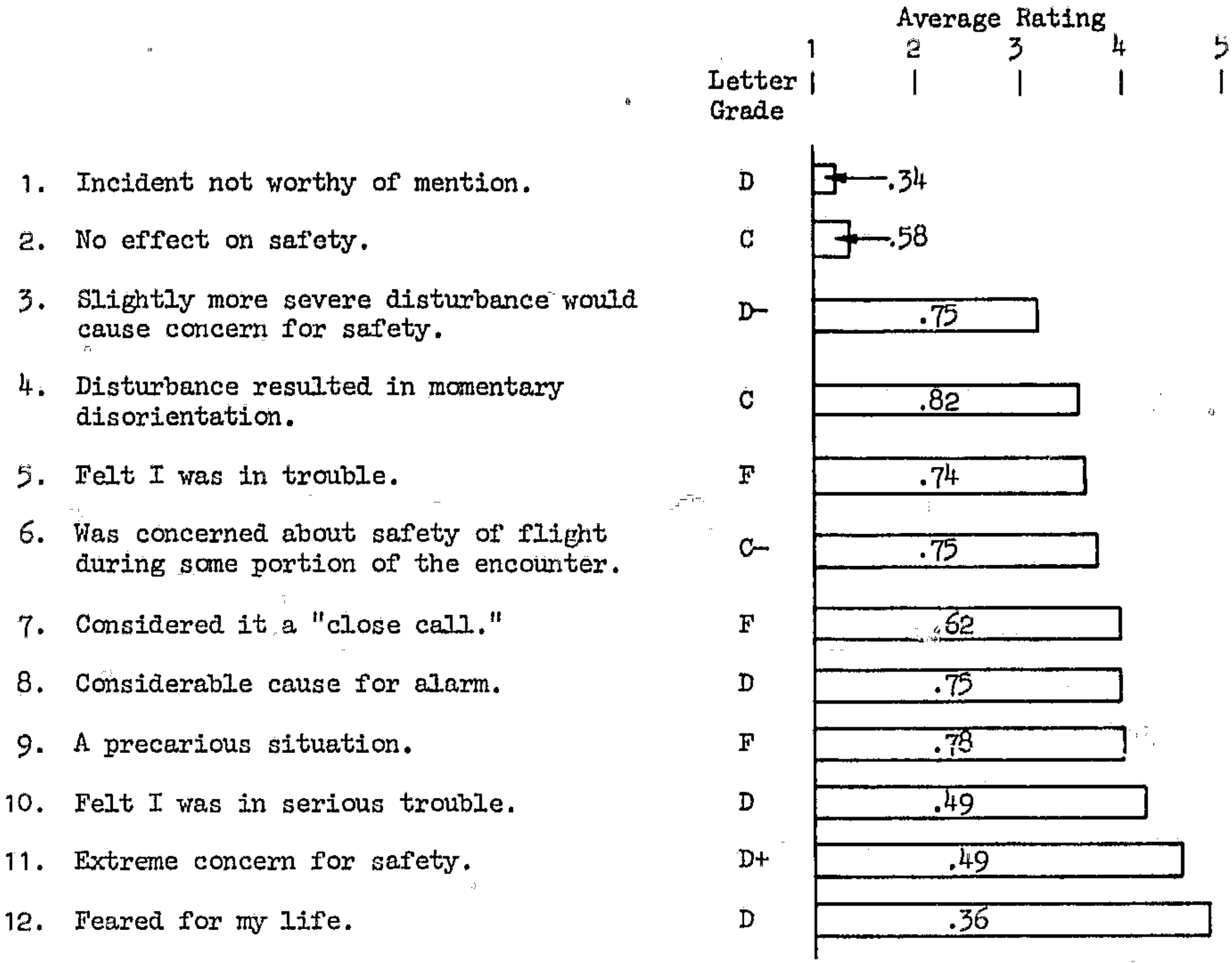


Figure B-5. Subjective opinion.

REFERENCES

1. Steven, S. S., Handbook of Experimental Psychology, John Wiley and Sons, Inc., New York, 1951.

2. McDonnell, John D., Pilot Rating Techniques for the Estimation and Evaluation of Handling Qualities, AFFDL-TR-68-76, Dec. 1968.

3. Hess, Ronald A., "Nonadjectival Rating Scales in Human Response Experiments," Human Factors, Vol. 15, No. 3, June 1973, pp. 275-280.

4. Contract DOT-HS-359-3-762, "Driver Vehicle Response Research," 30 June 1974.

5. Cooper, George E., and Robert P. Harper, Jr., The Use of Pilot Rating in the Evaluation of Aircraft Handling Qualities, NASA TN D-5153, Apr. 1969.