

PILOT ESTIMATES OF GLIDE PATH AND AIM POINT
DURING SIMULATED LANDING APPROACHES

by

C.W. ACREE, JR.

S.B., Massachusetts Institute of Technology

(1973)

SUBMITTED IN PARTIAL FULFILLMENT

OF THE REQUIREMENTS FOR THE

DEGREE OF

MASTER OF SCIENCE

AND

ENGINEER IN AERONAUTICS AND ASTRONAUTICS

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September, 1978

© Massachusetts Institute of Technology 1978

Signature of Author _____

Department of ~~Aerodynamics~~ and Astronautics

Certified by _____

Thesis Supervisor

Certified by _____

Thesis Supervisor

Accepted by _____

Chairman, Departmental Graduate Committee

PILOT ESTIMATES OF GLIDE PATH AND AIM POINT
DURING SIMULATED LANDING APPROACHES

by

C.W. Acree, Jr.

Submitted to the Department of Aeronautics and Astronautics on September 1, 1978, in partial fulfillment of the requirements for the degrees of Master of Science and Engineer in Aeronautics and Astronautics.

ABSTRACT

Aircraft landing approaches are the most difficult and dangerous phase of flying. Better knowledge of the perceptual response of pilots to deviations from a nominal landing approach would be of help in the development of better simulation and training techniques. Deviations in the vertical plane present in theory the greatest perceptual difficulties, so experiments were performed to measure pilot perceptions of glide path and aim point during simulated landing approaches.

Safety, cost, and convenience favored the use of a fixed-base cockpit simulator, with landings pre-recorded on video tape from a runway model board and projected with a video projector. Subjective estimates of the magnitudes of the approach deviations were obtained. Analysis of variance techniques were used to construct statistical models of the pilots' responses.

Pilots could estimate glide path errors quite well, but had difficulties estimating aim point errors. While the low accuracy of aim point estimation could have been due in part to problems with the experimental design, two other possibilities seemed more likely: fundamentally poor perception of aim point, and pilot preference for re-aligning with the originally desired glide path to correct for aim point errors. The data make plausible the hypothesis that pilots are little concerned with aim point during most of an approach, concentrating instead on remaining close to the nominal glide path and trusting this technique to guide them to the proper runway touchdown point.

Thesis Supervisor: Walter M. Hollister
Associate Professor of
Aeronautics and Astronautics

Thesis Supervisor: Laurence R. Young
Professor of
Aeronautics and Astronautics

ACKNOWLEDGEMENTS

I wish to express my gratitude to the people who helped and guided me through this thesis. Professor Renwick E. Curry initiated the research, Professor Walter M. Hollister guided me through its completion, and Professor Laurence R. Young oversaw the entire project as Director of the Man Vehicle Laboratory. Manuel J. Queijo of NASA Langley Research Center and David C. Nagel and John D. Stewart of NASA Ames Research Center were grant monitors for the research and assisted me in the use of NASA's simulator facilities. Their support is gratefully acknowledged.

Thanks also go to Mark E. Connelly, who patiently introduced me to the cockpit simulator, and to William A. Morrison, who was always ready with advice for using the Amphicon projector and with help in solving the many technical problems which arose. Many subjects from the MIT community generously donated their time to the experiments. Their contributions are all greatly appreciated.

Sherwood A. Modestino and Kerry M. Campbell deserve credit for creating the manuscript and were tremendous aids in solving a myriad of bureaucratic difficulties, for which they have my sincere thanks.

This research was performed under NASA Grants NGR 22-009-701 and NSG 2230.

TABLE OF CONTENTS

CHAPTER NUMBER		PAGE
1	Introduction	9
	1.1 Reasons for the Experiment	9
	1.2 Definitions	11
2	Choice of Equipment	17
	2.1 Introduction	17
	2.1.1 Preliminary Experiments	18
	2.1.2 Description of Equipment	19
	2.2 Models versus Computer Image Generation	23
	2.3 Video Taping Considerations	30
	2.4 Projectors versus Collimated Monitors	31
3	Experimental Design	34
	3.1 ANOVA and the Statistical Model	34
	3.1.1 Definition of the Model	34
	3.1.2 ANOVA Techniques and Significance	39
	3.2 Magnitude Estimation	41
	3.3 Practical Limitations of the Experimental Design	47
	3.4 Choice of Stimuli	50
	3.5 Experimental Procedure	53
4	Results	56
	4.1 Results for Grouped Data	63
	4.2 Interpretation	66

CHAPTER NUMBER	PAGE
4.2.1	Glide Path Estimates 66
4.2.2	Aim Point Estimates 67
5	Conclusions and Suggestions for Further Research 77
5.1	Pilot Performance 77
5.2	Experimental Procedure 79
APPENDICES	
A	Subject Data 82
B	Instructions to Subject 83
C	Individual Results 86
C.1	Individual Estimate Models 86
C.2	Raw Data Statistics 114
D	Instructions for the Operation of the Equipment 121
D.1	Operation of the Projector 121
D.2	Use of the Simulator Room and Screen 123
D.3	Suggestions for Video Taping 125
References	130

FIGURES

FIGURE NUMBER		PAGE
1.1	Illustration of terms describing an approach	12
1.2	Illustration of stimulus deviations	15
2.1	Side view of simulator and projector layout	21
2.2	Overhead view of simulator and projector layout	22
2.3	Photograph of the cockpit simulator taken from the mirror	24
2.4	Photograph of the mirror taken from the cockpit simulator	25
2.5	Photograph of the projector head taken from beside the mirror	26
2.6	Photograph of the pilot's seat taken from inside the cockpit simulator	27
4.1	Glide path estimate model for all subjects combined	57
4.2	Glide path estimate model for high-time pilots	58
4.3	Aim point estimate model for all subjects combined	59
4.4	Aim point estimate model for high-time pilots	61
4.5	Estimate of angle and ground distance by an ideal pilot	69
4.6	Illustration of a nulling approach	72
C.1	Glide path estimate model for subject #1	90
C.2	Aim point estimate model for subject #1	91

C.3	Glide path estimate model for subject #2	93
C.4	Aim point estimate model for subject #2	94
C.5	Glide path estimate model for subject #3	96
C.6	Aim point estimate model for subject #3	98
C.7	Glide path estimate model for subject #4	100
C.8	Aim point estimate model for subject #4	101
C.9	Glide path estimate model for subject #5	103
C.10	Aim point estimate model for subject #5	104
C.11	Glide path estimate model for subject #6	105
C.12	Aim point estimate model for subject #6	106
C.13	Glide path estimate model for subject #7	108
C.14	Aim point estimate model for subject #7	109
C.15	Glide path estimate model for subject #8	110
C.16	Aim point estimate model for subject #8	112
C.17	Glide path data statistics for all subjects combined	115
C.18	Glide path data statistics for high-time pilots	116
C.19	Aim point data statistics for all subjects combined	117
C.20	Aim point data statistics for high-time pilots	119

SYMBOLS

α	glide path angle
γ	flight path angle
Δ	unit stimulus deviation
K_1, K_2, K_3	constants in typical statistical model
L	significance level (specified value)
p	significance level (probability)
r_1, r_2, r_3	ranges from the runway

The following symbols and subscripts describe the linear statistical model used for data analysis:

A	main effect of first factor (range)
B	main effect of second factor (glide path)
C	main effect of third factor (aim point)
e	model error
μ	overall mean
()	parentheses indicate interactions between enclosed factors
Y	response
\hat{Y}	predicted response

Subscripts:

i	stimulus level index of first factor (range)
j	stimulus level index of second factor (glide path)
k	stimulus level index of third factor (aim point)

CHAPTER 1

INTRODUCTION

1.1 Reasons for the Experiments

Construction of a model of a pilot's visual information processing during VFR landing approaches would help develop advanced techniques of flight simulation. Such a model would determine specifications for the quantity, fidelity, and exact types of visual information presented to the pilot during simulation of any particular flight condition, thereby allowing more efficient, cost-effective, and subjectively realistic flight simulations for research and training. It would also aid in understanding safety aspects of visual landings: 50% of all aircraft accidents occur during the 2% of flight time spent in approach and landing (Hasbrook, 1975). But before such a model can be constructed, it is necessary to determine exactly which visual cues, out of all visual information, are actually important and how they interact with each other. Therefore, the goal of these experiments, as a first step in this research, was to obtain the perceptual responses of humans to deviations from a nominal landing approach.

In practice, deviations in the vertical plane are the most important, and in theory the most difficult to determine (Naish, 1971). These experiments tested perception of

altitude displacement deviations from a nominal glideslope, and aim point deviations from a nominal touchdown point. All experiments were run on a simulator, and the method of magnitude estimation, coupled with analysis of variance techniques, allowed maximum efficiency in data collection and analysis.

Running the experiments on a simulator instead of a real aircraft allowed major reductions in time and cost, and important improvements in safety and precision of the visual stimuli. It was desired to eliminate all seat-of-the-pants and time-integrated motion cues to force the subjects to concentrate on their instantaneous visual perceptions, so only brief segments of landing approaches were actually shown to the subjects, and there were no motion cues. This took the subjects "out of the loop", but was completely acceptable for these experiments and had the additional benefit of reducing the time required of the subjects as much as possible.

The subjects sat in a fixed-base cockpit simulator and looked out of the pilot's window at a rear projection screen which covered the entire field of view. A video projection system projected recorded black-and-white television scenes onto the screen, simulating the view out of a real aircraft during a landing approach. The television scenes were recorded on video tape with a moving camera and model terrain board visual scene generating system.

Boeing donated a fixed-based cockpit simulator to MIT, and NASA Langley Research Center loaned the Man Vehicle Laboratory an Amphicon 260 television projector system. With this equipment already in hand, it was affordable to integrate the two systems at MIT. It was also more convenient to run the experiments at MIT than at NASA Ames or NASA Langley Research Centers. The cockpit was modified to accept the Amphicon unit's projected image as an out-the-window display. Video tapes of landing approaches made on one of NASA's Redifon simulators were shown on the combined system. Programmed sets of landing approaches were simulated and the resulting visual images recorded for each experiment, then modified where necessary by video editing. This allowed the experiments to be run at MIT independently of NASA's facilities.

1.2 Definitions

From among the variations in terminology in the literature and among potential subjects, the following set of definitions were chosen for use in instructions to subjects and in this thesis. The "glide path" is the path through space from the aircraft to the nominal runway touchdown point located 1000 ft beyond the runway threshold. The "nominal" (or correct) glide path is the "glideslope", which

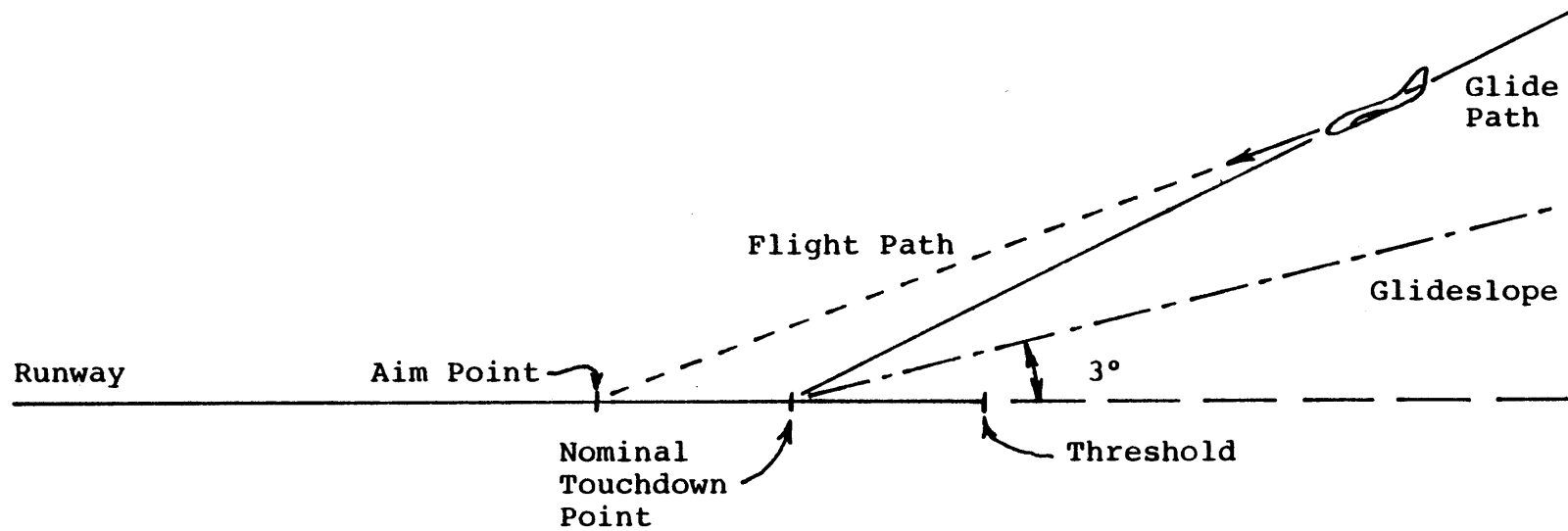


Figure 1.1 Illustration of terms describing a landing approach. (Angles are exaggerated.)

here makes a three degree angle from the horizontal. The "flight path" is the extension through space of the aircraft's instantaneous velocity vector. The "aim point" is the place on the ground intersected by the flight path (see Figure 1.1). In a nominal approach, the glide path, flight path, and glideslope all coincide, and the aim point is the runway touchdown point.

Pilots have direct control over the attitude of an aircraft, but have only indirect control over altitude or aim point. Yet pilots usually think of errors in approach in terms of height above (or below) the nominal glideslope, and in terms of aim point distance from the nominal touchdown point. Height and distance measures of error vary with ground distance (range) from the runway, unlike angular measures which remain constant. Furthermore, aim point errors vary highly nonlinearly with changes in flight path angle. These are fundamental problems in presenting stimuli to the subjects, and in their estimation of that stimuli.

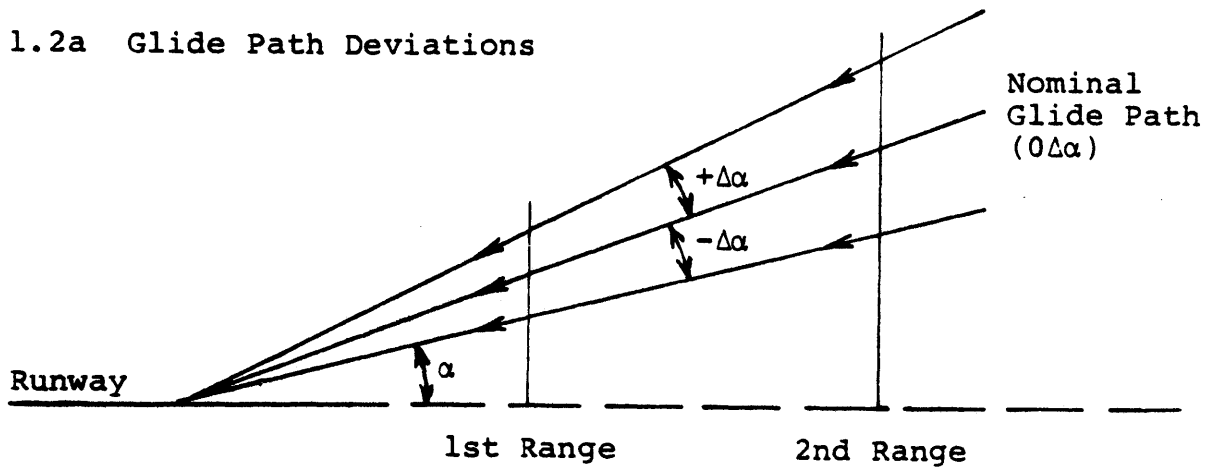
It was decided to base the stimuli on angular deviations, but to allow the subjects to respond in terms of height and distance. Subjects were told of the angular basis of the deviations, but were asked to estimate glide path errors as being "high" or "low" with respect to the glideslope, and aim point errors as "long" or "short" with respect to the

touchdown point. This ran the risk of complicating the experimental results in exchange for reducing confusion of the subjects.

The experiments were constructed as a full factorial design with three main effects: glide path, flight path and range. Although range was not estimated by the subjects, it had to be treated as an independent variable for proper analysis. The stimuli were defined as deviations from the nominal approach (see Figure 1.2). All stimulus levels were integral multiples of some unit deviation Δ . For glide path (α), the unit angular deviation $\Delta\alpha$ was 0.5° ; for flight path (γ), the unit deviation $\Delta\gamma$ was also 0.5° . Flight path deviations were defined relative to the total glide path, so that the nominal aim point (corresponding to $0 \Delta\gamma$) was always the touchdown point for any $\Delta\alpha$ (it was independent of glide path). The deviations were combined; for example, there could be a positive $\Delta\alpha$ and a negative $\Delta\gamma$, each of a different magnitude, and the combined stimuli could appear at more than one distance. All possible combinations were presented in random order.

The subjects made verbal estimates of the directions and magnitudes of the deviation stimuli in terms of a subjective numerical scale of -10 to +10. The full range of the estimation scale was established by showing the subjects a set of stimuli with deviations more extreme than in the main experiment. Different subjects took different amounts

1.2a Glide Path Deviations



1.2b Flight Path Deviations

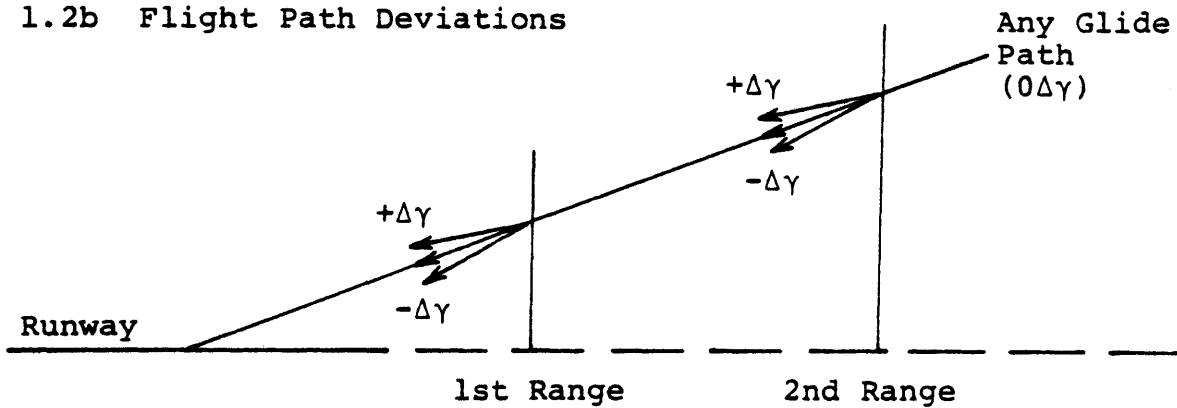


Figure 1.2 Illustration of stimulus deviations.

For clarity only three levels of stimuli are shown at two ranges. (Angles are exaggerated.)

of time to make estimates; the problem was aggravated if subjects made multiple estimates. So, to keep the experiments consistent, each subject saw the full set of stimuli twice - once for glide path and once for flight path - and made only one estimate at a time.

Numerical estimates allowed a linear statistical model of the subject's estimates to be constructed along the following lines:

$$\begin{aligned} \text{Estimate} = & \text{Mean} + K_1(\text{Glide Path}) + K_2(\text{Flight Path}) \\ & + K_3(\text{Range}) + \text{Interactions} + \text{Error} \end{aligned}$$

The model was "linear" in the sense that each independent variable (and all interactions) was multiplied by an appropriate constant, then added to the mean and error to get the total response. Nothing is implied about the subject's perception or decision-making processes. Analysis of variance techniques allowed determination of the constants and statistical confidence levels for each main effect and the interactions. The modelled estimates were plotted to allow easier interpretation of the results. The model was not intended to establish any cause and effect relationships, but to establish instead the relative statistical importance of the different visual stimuli.

CHAPTER 2

CHOICE OF EQUIPMENT

2.1 Introduction

Using an aircraft cockpit simulator was obviously the favored method for these experiments. Compared to actual flight tests, it was safe and inexpensive; and equally important for the collection of good data, it allowed the experiments to be exactly repeatable and relatively easy to modify. These same advantages held for using video tapes instead of real-time visual scene generation. The choice of the actual equipment was determined largely by ready availability. Preliminary experiments established the feasibility of the method.

MIT already had a fixed-base cockpit simulator (donated by Boeing) and an Amphicon 260 television projector system (on loan from NASA Langley Research Center). Redifon terrain board simulators were available at both NASA Langley and Ames Research Centers. Video equipment could be initially obtained from the MIT Video Services (administered by the Center for Advanced Engineering Study); the final experiments were run using a video cassette player and a monitor rented from commercial sources.

All of the equipment was expected to be immediately compatible. The Redifon systems used the same video format as the Amphicon projector and standard video tapes. The

cockpit and projector were geometrically compatible (the projected image was the correct size at the correct viewing distance) and could be integrated with a minimum of modification (mostly careful relocation of components).

Details of the use of the equipment are given in Appendix D.

2.1.1 Preliminary Experiments

Two sets of preliminary experiments were run. The first set was used to work out the experimental design and protocol, and the second was used primarily to test the equipment. Both sets used video tapes made at NASA Langley Research Center.

Changes indicated by the results of the preliminary experiments were incorporated into the final experiments, which were recorded at NASA Ames Research Center. Refinements to the experiment are described in Section 3.3, Practical Limitations of the Experimental Design.

For the first set of preliminary experiments, the subjects sat in an open room while estimating both glide path and aim point on each approach. Important results were:

- (1) The subjects were able to make the estimates with the experimental equipment and procedures.
- (2) The illusion of sitting in an aircraft and looking out at a runway could not be sustained without a physical enclosure blocking visual stimuli coming

from the rest of the room beyond the screen - i.e., a cockpit simulator was needed.

- (3) Two estimates per approach were too many, but repeating the full set of experiments twice in one session took too long (nearly two hours).
- (4) The taping method then used (1/2 inch open reel) was awkward and unreliable.

In spite of the problems, it appeared that usable data could be obtained with a suitably refined experiment. Thus, it was decided to install the projection system in the cockpit simulator, simplify the experimental design to allow time for two separate experiments per session, and modify the taping procedure. The basic concept of viewing video tapes of simulated landing approaches appeared to be sound.

2.1.2 Description of Equipment

The Amphicon 260 television projection system accepts standard 525 line black and white video signals from a cable line input (Amphicon Systems, 1967). It uses separate projector unit and electronic control cabinets for flexibility of installation. The projector unit uses Schmidt optics: a small (6 inch), very high intensity cathode ray tube projects onto a spherical mirror, which reflects the light back through a corrector lens, then to the screen. This puts a 7.5 ft by 10 ft image at a focal length of 19 ft, the optimum distance

for best resolution with this system, yielding a standard 3:4 aspect ratio picture. With the screen 12 ft from the viewer's eye, the image covers a $35^\circ \times 45^\circ$ field of view. See Figures 2.1 and 2.2 for diagrams of the layout.

Redifon visual simulators at NASA Ames and Langley Research Centers generated the television pictures for the video tapes. Both simulators have large scale models (terrain boards) of airports and their surrounding terrain, and television cameras which translate in three dimensions to follow scale motions of aircraft. The television pictures showed a view of the modelled areas resembling that seen from a real aircraft cockpit. For these experiments, the cameras followed a preprogrammed set of motions under computer control. The Ames terrain board simulator used for the final set of tapes has a model scale of 600:1. Approaches were made to a 200 ft x 8000 ft runway, simulating a light aircraft flying at 80 mph.

Trial video tapes used for developing the experimental procedure were made at Langley using a 1/2-inch open reel manually controlled tape recorder. Problems with the recording method and desired refinements to the experiment necessitated recording a new set of tapes at Ames, which has more sophisticated video facilities. These tapes were recorded on a 3/4 inch cassette recorder, then edited and copied at the

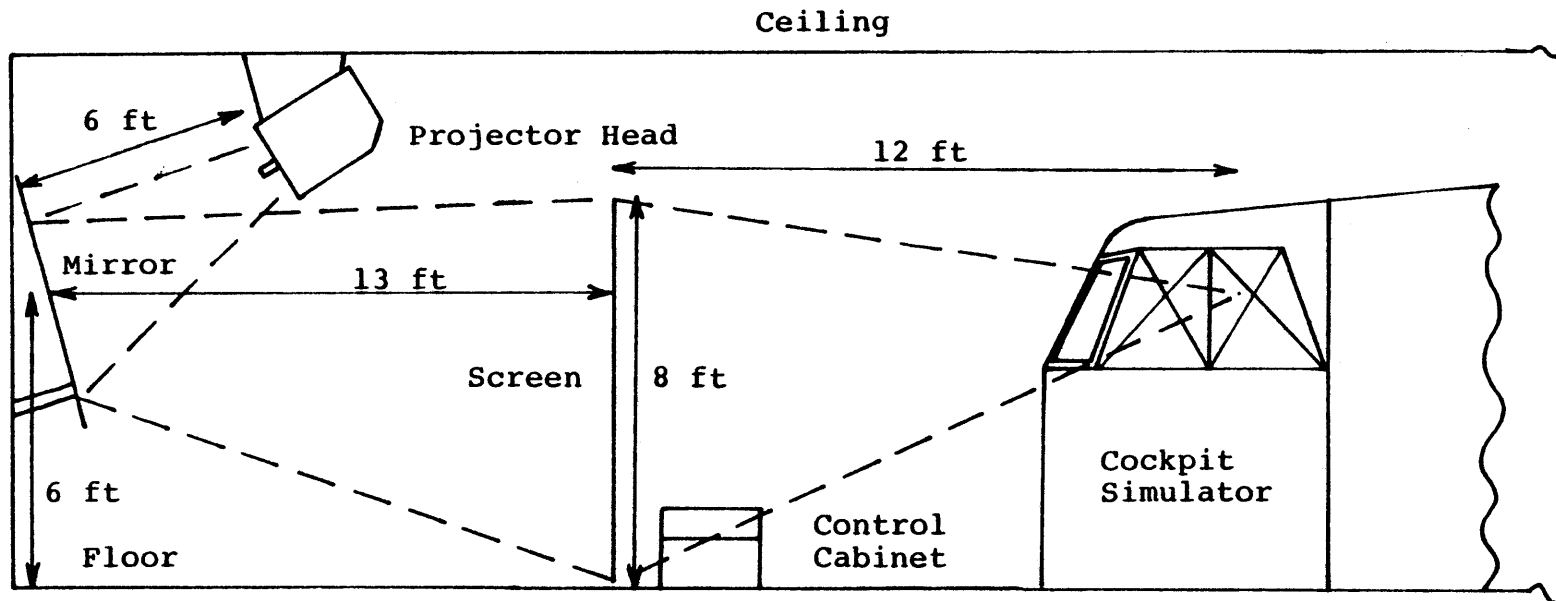


Figure 2.1 Side view of simulator and projector layout.

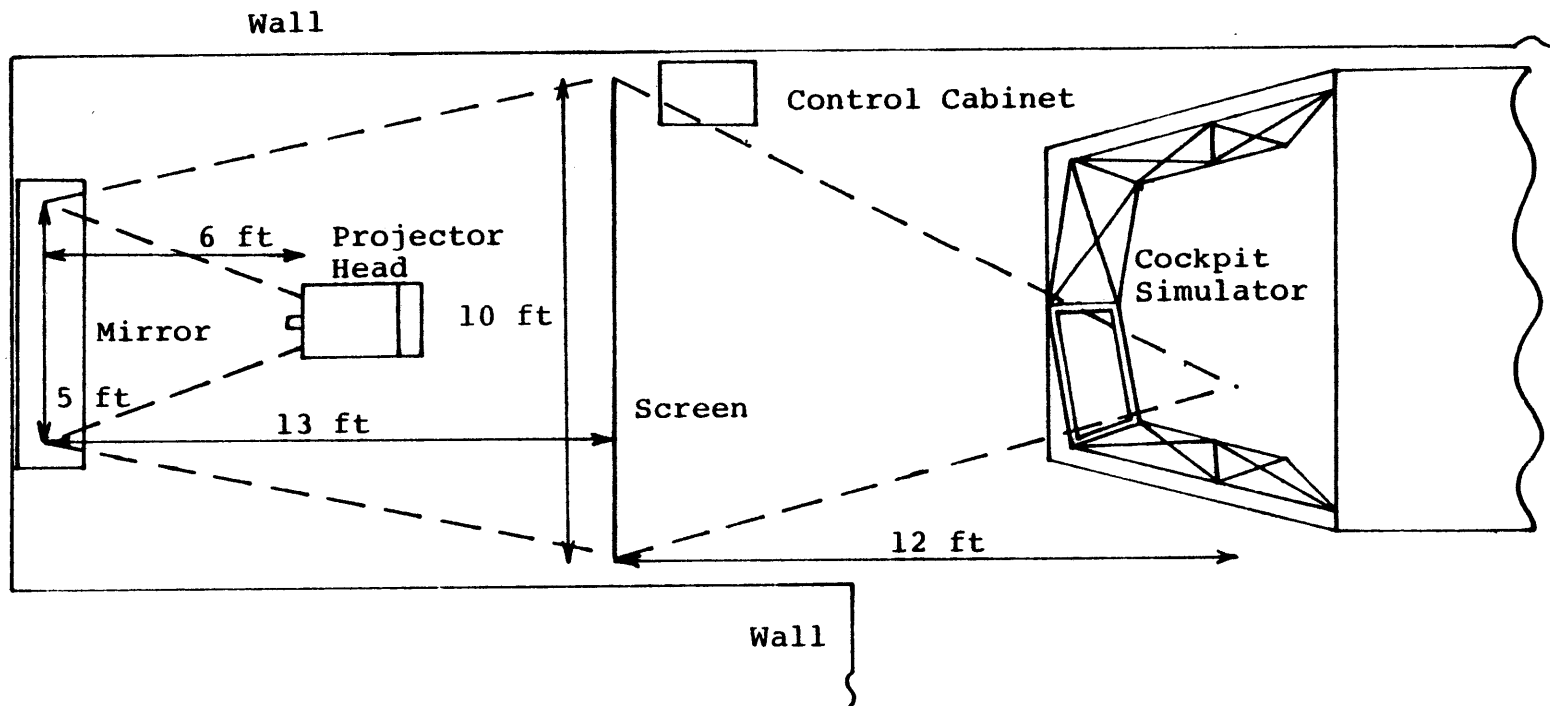


Figure 2.2 Overhead view of simulator and projector layout.
 All cockpit windows except the left forward window are blacked out.

University of Massachusetts Media Center (Columbia Point Campus).

In order to establish the proper viewing angles and optical path lengths, the cockpit was repositioned in the simulator room, and the projector image was reflected through an additional mirror (the extra mirror also helped reduce the keystone effect, or tilt of the image away from the vertical). A total distance of 31 feet - 19 feet projector to screen plus 12 feet screen to pilot - was needed free of optical obstructions. The projector head was mounted near the ceiling, 6 feet from the mirror which reflected the image back underneath the projector to the screen. This reduced the needed outside length enough to allow use of the existing room, while providing a full-size image properly placed to cover the pilot's entire field of view from the left front cockpit window. Electronic control equipment for the cockpit's systems was rewired and remounted to allow an unobstructed view of the screen from the pilot's window. A second pane of glass in this window was removed to reduce optical interference and eliminate dust accumulation between the panes. Photographs of this equipment are shown in Figures 2.3 through 2.6.

2.2 Models versus Computer Image Generation

Television images of the runway needed for the experiment

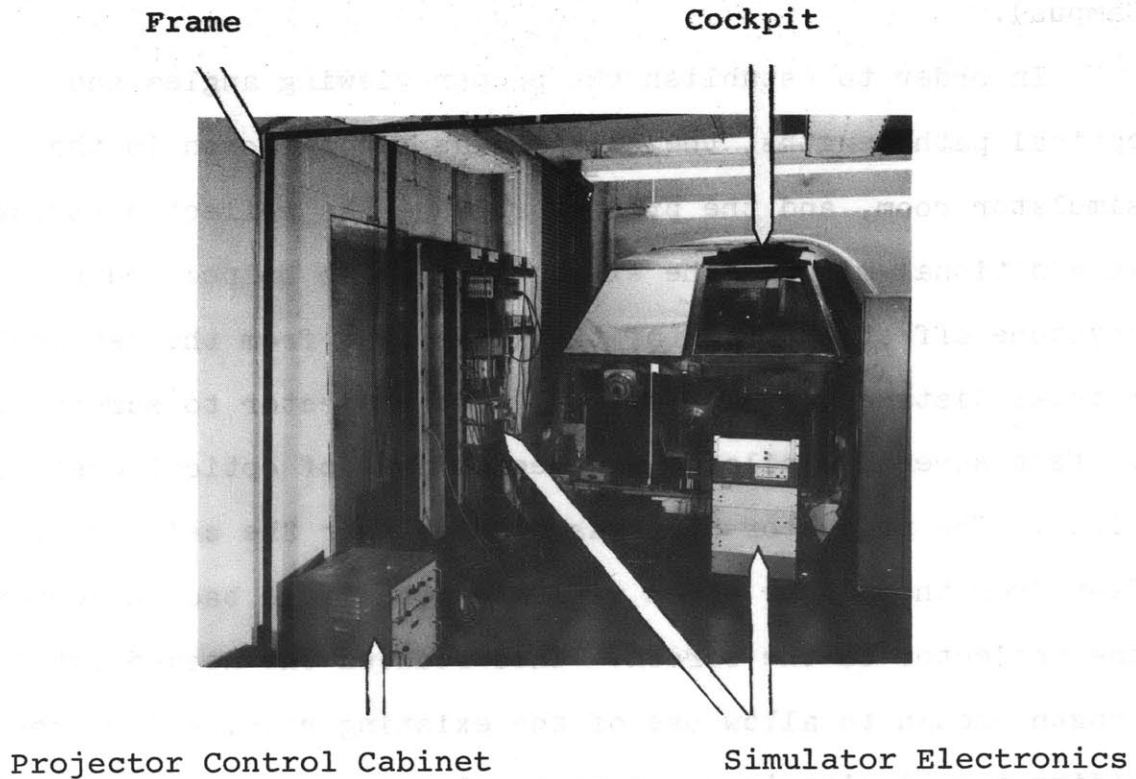
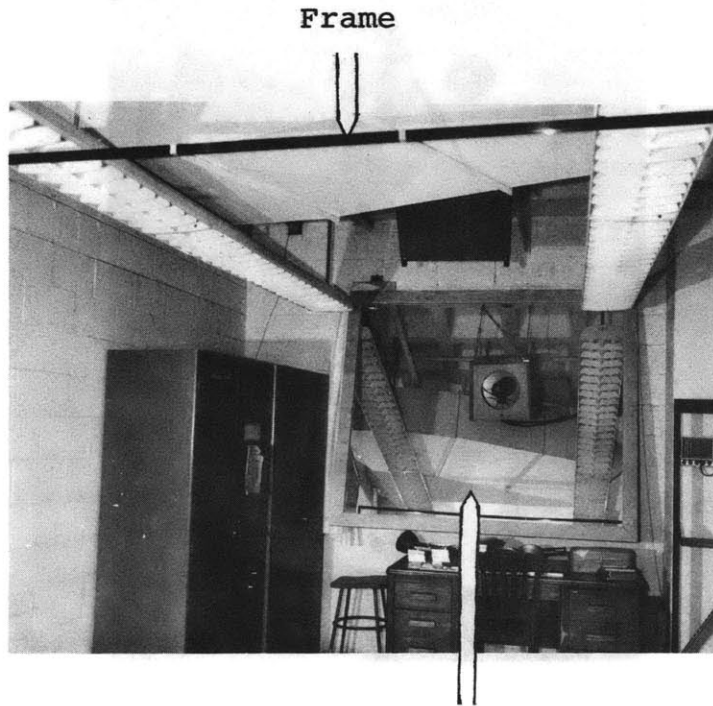


Figure 2.3 Photograph of the cockpit simulator taken from the mirror.

The pilot's forward window is uncovered. The projector's electronic control cabinet is visible at the lower left, next to some of the simulator's electronic equipment which has been moved out of the pilot's field of view. In order to show the equipment clearly, the rear projection screen has been removed from its frame, part of which can be seen at the left and top.



Frame

Mirror

Figure 2.4 Photograph of the mirror taken from the cockpit simulator.

The projector head is visible as a reflection in the mirror. Part of the frame can be seen at the top of the photograph.

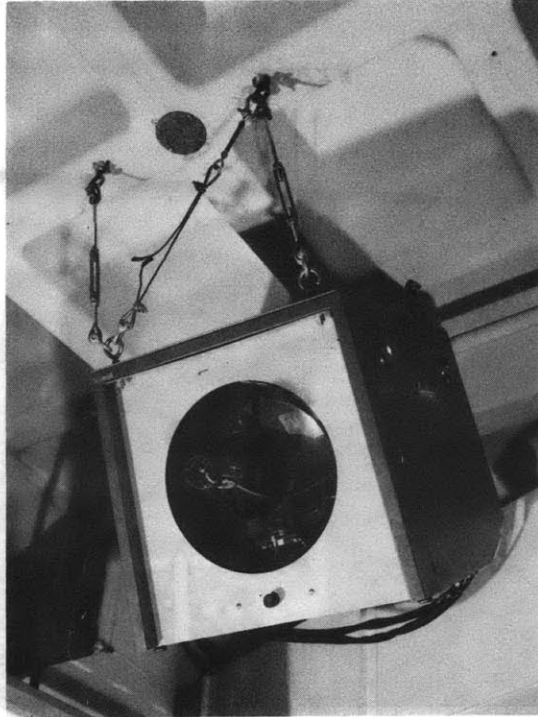


Figure 2.5 Photograph of the projector head taken from beside the mirror.



Figure 2.6 Photograph of the pilot's seat taken from inside the cockpit simulator.

A plastic ball on a wire used as a reference for eye positioning can be seen extending from the cockpit ceiling.

could have been made using either runway models or artificial images generated by a computer. Either method would allow more precision at less cost than taking pictures from a real aircraft. Models have slight theoretical and practical advantages, although doing computer image generation (CIG) at MIT was considered. Video tapes are compatible with both methods, and allow use of facilities outside MIT.

Showing a subject a simple outline figure of a runway does not provide adequate realism; a considerable amount of fine detail of the runway and its surroundings is necessary for accurate interpretation (Barnes, 1978; Dorfel, 1978; and McGregor, 1970). CIG systems are improving in detail capability rapidly, but runway terrain board models were still considered superior when the experiments were planned, even though limitations of their associated camera systems degrade performance (Key et al, 1978; and Welch, 1978). Television cameras are not capable of providing an adequate depth of field in the focused image under all conditions of simulation. Moreover, limitations of the television raster resolution partially negates the advantage of higher model detail (Barnes, 1978; and McGregor, 1970). Though slight, the theoretical advantages of direct compatibility with video tape and relative ease of use of the Langley and Ames Redifon simulators led to the choice of image generation from runway models.

The possibility of using an ADAGE 130 graphics computer at MIT to create the runway images was seriously considered. The limiting problem of CIG is computation speed: The image will flicker if all the drawing operations and supporting numerical calculations are done too slowly. ADAGE computers save time by doing image rotations with analog hardware rather than digital computation, and by using extremely high writing speed displays. But there is no way to access display data after rotation, so perspective calculations must be done using numerical trigonometric calculations. This slows down the process so much that only the crudest of runway images can be made, and even then not without noticeable flicker. Furthermore, the ADAGE is a vector-scan system and is not compatible with line-scan television systems such as the Amphicon projector, and the high writing speed of the ADAGE prevents effective conversion with existing vector-to-line scan converters. Although the ADAGE computer can be programmed to force an image update 40 times a second to avoid flickering, the complete cycle time for all runway image calculations is too slow to keep up. This causes the image to jump discontinuously instead of showing smooth motion. The high persistence CRT displays used on the ADAGE will blur the image if the jumps are too large, as happened during runway programs.

Nevertheless, the ADAGE was tested by pointing a television camera at the ADAGE display and viewing the output on the Amphicon projector. The results, as predicted, were very poor and offered little hope for adequate improvement, so this line of approach was abandoned.

2.3 Video Taping Considerations

Video tapes eliminate the dependence on local image generation facilities (most importantly the ADAGE computer). One set of experiments can be run through and recorded elsewhere, then repeatedly played back to subjects without tying up the scene simulation equipment. Virtually all current aircraft visual simulators use television/CRT systems with either CIG or terrain board models, because they can operate in real time. This naturally favors the use of video recording instead of film for any visual recordings. Video tapes can be edited and copied more easily than film, so that small changes can be made, and all subjects can see exactly the same experiments. The Amphicon was almost immediately compatible with the standard video tape format, but did require modification of the horizontal automatic frequency control feedback circuit to eliminate sideways "shaking" of the picture.

Tapes for the preliminary experiments were recorded at Langley on 1/2 inch open reel videotapes. The quality of these

tapes was barely acceptable, and the available editing facilities at MIT (all 1/2 inch) were unreliable. Since any extensive editing would have to be done outside MIT, there was no need to restrict the taping to the 1/2 inch format. The final set of tapes was recorded at Ames on 3/4 inch cassettes, which provided a noticeably higher quality image and for which excellent editing facilities were available at the University of Massachusetts.

2.4 Projectors versus Collimated Monitors

There are two common ways of presenting visual images to a pilot in a simulator: by projecting the picture on a screen, usually placed well outside the cockpit, or by collimating the picture from a monitor very near the cockpit window. Both methods have their faults. Collimated monitors tend to be subjectively preferred by pilots, but their claimed advantages have not been analytically verified. In the absence of experimentally supported preferences, a projection system already in hand was chosen, and it performed adequately in the preliminary experiments.

When a single video display is used, the total field of view is important. (Multiple overlapping displays can provide arbitrarily wide viewing areas, for a stiff penalty of complexity). Wide fields of view, which generally favor projected

displays, are considered desirable by many researchers (Huff and Nagel, 1975; Kraft and Shaffer, 1978; and McGregor, 1970). Although wide viewing angles are obviously needed for circuit flying (Barnes, 1978), flight tests (Armstrong, 1970) have shown that even a very narrow field of view is adequate for straight-in approaches, as were being simulated here.

Simulating the effect of viewing a runway at a distance is more important than the field of view for straight-in approaches, and more difficult to do. The binocular effect is not important, except perhaps when actually on the runway (Spooner, 1973), but an infinity focus is needed (Barnes, 1978; Dusterberry, 1978; and Kraft and Shaffer, 1978). Projector-and-screen systems cannot provide natural focusing, and it should be noted that perfect collimation also eliminates the focal distance cue.

There are other flaws in collimation systems: High light transmissivity cannot be had without optical distortion, and it is very difficult to get correct focusing over the entire display area using conventional curved-face monitor CRTs (Dusterberry, 1978; and Kraft and Shaffer, 1978). Neither projection nor collimation systems allow normal head movement by the viewer: Collimation lenses restrict the viewing angle (Kraft and Shaffer, 1978), and projectors have a parallax effect (Spooner, 1973).

(It is possible to combine the two methods, thereby reducing many of the disadvantages. Redifon's duoview system projects onto a screen to get an optically flat image, then collimates the image in a large curved mirror, giving a large, bright, and clear image (Spooner, 1973). Unfortunately, such systems are very bulky and expensive, thus impractical for use here.)

Actual system performance of the competing methods were compared by Chase (1971). Collimated monitors versus projectors, and color versus black and white tests were run using a wide variety of pilot performance measurements. Although the pilots' subjective preferences and their performances seemed to favor collimated color monitors, the measured differences were rarely statistically significant. The difference in actual performance was not great enough to justify the time and expense of developing or acquiring a collimated color system for these experiments. The black and white projector performed promisingly enough in early experiments to justify its use.

(Future systems may possibly overcome the current problems of simulator displays. In particular, laser-scanned models and holographic displays look very promising (Driskell, 1978; and Fowler et al, 1970), but they are still in development.)

CHAPTER 3
EXPERIMENTAL DESIGN

3.1 ANOVA and the Statistical Model

To achieve maximum efficiency in the design of the experiment and the analysis of the data, a full-factorial Type I analysis of variance model was chosen. Full-factorial means that all possible combinations of stimuli at every level are presented in the experiment. Type I designs have a finite set of discrete values of the stimuli, chosen in advance. Analysis of variance (ANOVA) is a particular method of statistical analysis of the data (Crow, et al, 1960; Snedecor and Cochran, 1967).

3.1.1 Definition of the Model

A general linear model ANOVA program (UCLA's Biomedical Package BMD01V program) was used. It uses a statistical model of the form

$$Y_{ijk} = \mu + A_i + B_j + C_k + (AB)_{ij} + (AC)_{ik} + (BC)_{jk} + (ABC)_{ijk} + e_{ijk}$$

where Y_{ijk} = response at stimulus levels ijk

μ = overall mean

A_i = effect of 1st factor (range) at stimulus level i

B_j = effect of 2nd factor (glide path) at stimulus level j

- C_k = effect of 3rd factor (aim point) at stimulus level k
 $(AB)_{ij}$ = interaction between 1st and 2nd factors at stimulus levels ij
 $(AC)_{ik}$ = interaction between 1st and 3rd factors at stimulus levels ik
 $(BC)_{jk}$ = interaction between 2nd and 3rd factors at stimulus levels jk
 $(ABC)_{ijk}$ = interaction of 1st, 2nd, and 3rd factors at stimulus levels ijk
 e_{ijk} = effects not accounted for by the model (error).

Deleting the error term, e_{ijk} , gives the model for \hat{Y}_{ijk} , the predicted response.

The model is considered linear because all of its factors are additive, but it does not assume any particular response function. It constructs a best fit model to individual points, rather than an assumed curve, and so avoids distorting the results with prior assumptions. Nonlinear relationships between stimulus and response are thus fully revealed. Of course, if any important combination of variables were to be left out of the model, the results would be questionable (the e's would become very large), but this is true of any type of method of analysis.

Factors are essentially independent variables, or mathematical representations of different types of stimuli. In this experiment, there are three: range, glide path, and flight path angle (or aim point). This is a Type I model,

meaning that each type of stimulus takes on known discrete values, or levels. With each factor is associated an index (i, j, or k), and with each stimulus level of a given factor is associated an index value. Thus, any combination of stimuli is uniquely labeled by a set of index values which also identify the response (Y_{ijk}) to those stimuli.

Changes in the response due to changes in a single index (all others remaining fixed) are called main effects. Changes different from those already predicted by the total of the individual main effects and which can be attributed to particular combinations of index values are called interactions. (It should be noted that the notation (AB), (ABC), etc. does not imply multiplication, but only coincidence. Examples are given below.) The model is thus a linear sum of functions, each function relating changes in the response to changes in the stimuli.

All main effects and interactions are assumed to have zero means:

$$\sum_i A_i = 0, \quad \sum_{i,j} (AB)_{ij} = 0, \text{ etc.}$$

Also, the errors e_{ijk} are assumed to have independent zero mean normal distributions. Any actual response biases are lumped into the overall model mean μ .

The model error e_{ijk} is not the same quantity as the subject's estimate error. If an important effect were left

out of the model, the model error variance would be larger than the subject's estimate error variance. Also, the model is indifferent to whether the subject makes a correct estimate of the stimulus being tested, so subject errors due solely to the influence of other factors included in the model would not contribute to the model error. There are thus three types of error: failure of the subject to estimate the value of the stimulus correctly, deviations from the subject's mean estimate (which itself may be an error), and failure of the model to exactly fit the subject's responses.

A, B, and C are main effects, and are directly controllable independent variables (range, glide path, and aim point stimuli). The effect due to any one of them is the same at all levels of all other main effects. For example, in these experiments, range was frequently a significant main effect, meaning that all estimates of, say, glide path were biased slightly higher at one range than at another, the bias being uniform for all glide path estimates, so that the shape of the response curve was the same at all ranges.

(AB), (AC), and (BC) are interactions, or joint effects, for which the effect due to either member of a pair is different at different levels of the other member. For example, an interaction between range and glide path means that the bias in the glide path estimates due to range effects changes between different levels of glide path, so that the shape of

the response curve changes between ranges. (ABC) is a second-order interaction, for which some particular combinations of all three independent variables have substantially different effects beyond those predicted by the main effects or first-order interactions. With only three main effects - range, glide path, and aim point - there are no other possible combinations of controlled variables, hence no higher interactions.

An example of a possible result would be for mean, range, glide path, and a range-glide path interaction to all have significant effects on glide path estimates (statistical significance is discussed in the next section, ANOVA Techniques and Significance). Thus,

Predicted glide path response = mean + range effect +
glide path effect + range/glide path interaction

or

$$\hat{Y}_{ij} = \mu + A_i + B_j + (AB)_{ij} \quad \text{for all } i \text{ and } j$$

In this thesis, responses predicted by the model are plotted point-by-point against relevant stimuli, with all statistically significant model terms being taken into account. This gives a graphical representation of the important functions within the model.

For maximum generality, the model is indifferent to the physical magnitudes of the stimuli; it merely associates responses with stimulus index labels. This is necessary because the forms of the response functions are not known in advance. The purpose of the method is to determine the relative importance of each function - main effect or interaction - and its relationship to the errors, as calculated by ANOVA techniques.

3.1.2 ANOVA Techniques and Significance

The purpose of ANOVA is to determine the relative importance of the different effects, interactions, and errors in the model. The entire model is reevaluated with each factor eliminated in turn, and the resulting increase in overall error variance calculated. If the model error variance without any effect is not sufficiently greater than the variance with the effect included, then the apparent change in \hat{Y} due to the effect could just as well be due to chance. The probability that the change in \hat{Y} is due to chance is called the significance level, p .

A compromise must be reached between leaving out important effects which happen to be only slightly greater than the noise level (model error) and confusing the analysis with apparent effects actually due to random errors. A 5% significance level (here written $p < 0.05$) was chosen as the cut-off

level. This level is commonly used in experiments such as these with good results. As it turned out, 5% was more than large enough to include all directly estimated effects (glide path and aim point) for the corresponding experiments. Where a factor is significant at some other level L , it is given in parentheses ($p < L$).

Strictly speaking, the presence of a significant interaction invalidates any conclusions about the related main effects in an overall analysis; the analysis must be broken down and re-done at each level of the pertinent main effects separately. The ANOVA algorithm used here was chosen because it computes the effects of all model factors individually, so that the original computations remain valid. Caution is still necessary when deciding which factors to include in any interpretation.

It is common practice to pool non-significant factors with model error to get a better estimate of the overall error variance, on the justification that non-significant effects are likely due to chance anyway. Again caution is necessary, there being a definite chance that an effect judged non-significant really does have an important, but small, effect. Pooling such an effect into the error variance decreases the accuracy of further calculations, and may lead to erroneous conclusions about other factors. For these experiments, the analysis was done iteratively: Only obviously non-significant

factors ($p > 0.25$) were pooled into the error variance on the first iteration; the new error variance was then used to recalculate the significance level of all marginally significant factors. Only then were final judgments made. The ANOVA program was rerun with non-significant factors suppressed as a check on the accuracy of the original analysis; no important changes were noted.

The usual next step in data analysis would be a multivariate regression on significant factors. But this requires that response curves be assumed in advance, which though easy to do for the obviously linear glide path data cannot be done with much assurance for the aim point data. The use of magnitude estimation complicates the problem. It was decided that developing refinements to the experimental technique to get better data would be more useful than trying to force the existing data into a more sophisticated model.

3.2 Magnitude Estimation

Magnitude estimation provides a maximum of usable data from psychophysical experiments of limited length. Perception is difficult to measure: Even those perceptual processes of which an experimental subject is conscious may not be easily described by him. The method of magnitude estimation assigns an arbitrary scale of units to the range of stimuli presented

to the subject, then requires the subject to estimate the level of each separate stimulus in terms of the assigned scale. The resulting data is one step removed mathematically from the actual physical stimuli and perceptions, but it gives numerical data where not otherwise obtainable in a reliable or convenient form. It remains up to the experimenter to correctly interpret the data.

It should be emphasized that magnitude estimation is not a theory of perceptual measurement; it merely provides numerical response data (Anderson, 1974). It does not determine the mechanisms, or their functional models, behind a subject's perceptual or decision-making processes. What it can do is generate a purely numerical model linking responses and stimuli. Getting back to the physical mechanisms underlying the mathematical model is a difficult (and sometimes controversial) interpretive task (Poulton, 1968). But at the very least, magnitude estimation can determine the relative importance of different stimuli on a subject's responses more accurately and reliably than he may be able to describe them himself; this is valuable enough for many psychophysical problems.

Two internal mechanisms can be distinguished for a subject's estimation process: First, a physical stimulus causes a perception, then, second, that perception is mapped onto an internal scale to determine a response. Ideally, the scale is a representation of physical reality based on the same sort

of stimuli and perceptual mechanisms as used in the experiment, but extraneous and uncontrollable influences - such as memory - usually distort the scale. The method of magnitude estimation itself correlates the responses with the stimuli without regard for the internal processes connecting them. It is therefore difficult to determine the subject's physical perceptions free from possible distortions by the judgment scale, which can easily be affected by the particular experimental conditions independently of the physical stimuli themselves.

For an individual subject, distortions and nonlinearities of the internal scale are not very important in a well-designed experiment, as long as the scale is monotonic. As emphasized by S.S. Stevens (1966), the developer of magnitude estimation techniques, the key requirement is that there be a statistically reliable correspondence between stimulus and estimate. However, comparisons between subjects can be confused by different distortions in different scales. It is not always possible to distinguish between differences in perceptions and differences in scales in a single experiment.

These were partition experiments with the judgment scale set by two end-anchors, and with the stimuli and estimates all lying between those two points (Poulton, 1968). End-anchors are stimuli chosen to lie beyond the stimuli used for data collection. This eliminates distortions commonly found at the ends of the judgment scale (Anderson, 1974). Numerical labels

of -10 to +10 were applied to the stimuli extremes, which were shown at the beginning of each experimental session to establish the subject's judgment scale.

A perfect landing approach was also shown at the middle of the stimulus range (numerically zero). Some kinds of nonlinear scale distortions can shift the subject's judgments toward one end of his scale, thus offsetting his judgmental zero from the true stimulus zero, as would happen with a logarithmic scale, for example. Attempting to establish the true numerical zero on the subject's scale with another scaling stimulus would then confuse the subject, so the perfect landing approach was presented separately from the end-anchors for familiarization, not as a scaling stimulus. Also, no data were taken on this stimulus (there were other important reasons for this, discussed in the next section, Practical Limitations of the Experimental Design).

The arbitrariness of the judgment scale does not pose as big a problem as might first appear. Virtually no system of physical units is internally natural to the subject; it is simply a verbal and mathematical convention. Therefore, the judgment scale may be more usefully constructed for experimental convenience than to match a subject's habitual usage. The 21 point (-10 to + 10) range used here is considered simple enough for subjects to readily accept, while giving sufficiently small increments for useful accuracy (Anderson, 1974).

Even when using purely arbitrary scales to which subjects have no previous habitual adjustment, it is often found that subjects will have certain preferred numerical responses. These depend less on the physical magnitudes of the individual stimuli than on the total range of all stimuli presented so far, the magnitudes and range of the scaling stimulus or standards, and possibly on the order of presentation of the stimuli (Poulton, 1968). An obvious example of the problem is the logarithmic nonlinearity of loudness responses in aural perception experiments. Order of presentation effects are easily handled by using random presentations of the stimuli and by ensuring that all subjects are given the same sequence of stimuli (otherwise, comparisons between subjects might not be valid). Poulton (1969) suggests using an iterative technique to control the other problems: The magnitudes of the scaling stimuli and the distribution of the experimental stimulus levels are adjusted after several sessions to place the stimuli near favored numerical responses. The process is repeated until the functional distribution of the most common responses matches that of the physical magnitudes of the stimuli (regardless of whether the magnitude estimates are in fact correct). Although this may appear at first to be juggling the experiment to get good-looking data, it is perfectly valid as long as the subjects do not know the magnitudes of any stimuli in advance (and preferably not even that there are

discrete, fixed values). Unfortunately, such iterative methods are rarely employed in experiments of the complexity of these, and they were not used here for the same reasons: The expected increase in accuracy is not worth the considerable effort and subject time required, and other problems with the experiment dominate.

To determine whether scale distortions had any important effects, the data were linearly adjusted and normalized to make the modelled responses exactly fill a range extending from -1 to +1. This was done separately for each subject and excluded the responses at the stimulus extremes, using only the data intended to be statistically analyzed. This gave approximately equal sensitivity to changes in stimuli to all subjects and preserved curvature in the responses, which helped achieve the best results in statistical calculations. The adjusted data were analyzed in exactly the same manner as the raw data. No differences in statistical significance were noted. This was expected for individual subjects (most statistical parameters do not depend on overall magnitudes, but only on relative differences), but it was also true when several subjects were analyzed together. From this it was concluded that scale distortions had a relatively insignificant effect on the results, or were unusually consistent from subject to subject. Therefore, only the analytical results based on raw data are included in this thesis. (To allow comparison with

other researchers' results, scale compression effects are briefly discussed in Section 4.1, Results for Grouped Data.)

3.3 Practical Limitations of the Experimental Design

Problems with the experimental design can best be illustrated by first considering the following ideal experiment:

Treatment Variable	Levels	Total Runs
Data Runs:		
Range from touchdown	r_1, r_2, r_3	
Glide path deviation	$0, \pm 1\Delta\alpha, \pm 2\Delta\alpha$	$3 \times 5 \times 5 = 75$ (full factorial)
Aim point deviation	$0, \pm 1\Delta\gamma, \pm 2\Delta\gamma$	
Calibration runs:		
Glide path scale calibration	$0, \pm 3\Delta\alpha$	
Aim point scale calibration	$0, \pm 3\Delta\gamma$	$3 \times 3 \times 3 = 27$ (full factorial)
Range	r_1, r_2, r_3	

Although the data runs and calibration runs separately constitute full factorial designs, their combination does not, and a non-singular solution to the linear model program is not

possible for such a combination. So the calibration run data cannot be used in the construction of a statistical model. The purpose of the calibration runs is instead to establish estimation scale "end-anchors" at extreme levels of stimuli.

As many replications as possible are desired for obtaining stable measurements, preferably at least three. Many levels of stimulus deviations allow generating a precise statistical model (assuming the distinctions between levels are not swamped by noise). At least two separate levels in each direction (+ and -) are needed to spot any response behavior beyond gross sensitivity and bias. It is desirable to establish response behavior as a function of distance, and three distance stimuli are indicated as a convenient minimum.

There are thus a total of $75 + 27 = 102$ combinations of stimuli (actually 99, since the three extra nominal approaches in the scale calibration can be eliminated with no loss of useful data). With three replications of each combination, there are 297 separate runs per experiment. As the runs average about 12 seconds long, 60 minutes would be required for just the data runs in one experiment. Both experiments (one for glide path and one for aim point), plus orientation, training, and practice runs, would take a total of over two hours per subject.

This was far too long for a single experimental session. Eyestrain, fatigue from intense concentration, and sheer boredom of the subject forced the length of each session to be drastically shortened. Each run was about as short as practical for subject response already, and breaking up the experiment would have created problems with data consistency and subject scheduling. The only remaining possibility was to reduce the number of data points, requiring the elimination of one range, all zero deviations in glide path and aim point, and interactions in the scale calibration runs. Three replications were still present. The resulting experiment is shown below.

Treatment variable	Levels	Total Runs
Data Runs:		
Range from touchdown	r_1, r_2	2 x 4 x 4 = 32 (full factorial)
Glide path deviation	$\pm 1\Delta\alpha, \pm 2\Delta\alpha$	
Aim point deviation	$\pm 1\Delta\gamma, \pm 2\Delta\gamma$	
Calibration Runs:		
Glide path scale calibration	0, $\pm 3\Delta\alpha$	(3 x 3) x 2 = 12 (no interactions)
Aim point scale calibration	0, $\pm 3\Delta\gamma$	
Range	r_1, r_2	

Note that even with the deletion of the nominal approaches, the data runs still constitute a full factorial design, but now the calibration runs do not, because there are no interactions. The calibration runs continue to function as estimation scale end-anchors.

There were finally $32 + 12 = 44$ stimulus combinations (or 42, eliminating the two redundant nominal approaches in the calibration runs) with three replications, yielding 126 runs per experiment and lasting about 25 minutes. Both experiments plus orientation, training, and practice runs took about one hour of actual experimental time. Set-up time, instructions to the subject, and any debriefing usually caused some runover, but the total time was under 1 1/2 hours and was considered barely acceptable. Some information on distance effects was lost, but this was tolerable for the purposes of these experiments.

3.4 Choice of Stimuli

If the flight path angle was in error, the actual glide path angle would change during an approach. To prevent this effect from influencing the results of these experiments, a limit was placed on the allowable glide path angle change during any given approach segment. For a fixed approach speed, this determined a maximum viewing time for each segment.

Work by other researchers (Gold and Hyman, 1968; Wempe and Plamer, 1970; and Gold, 1973) indicated that pilots could estimate glide path with group standard deviations ranging from 0.25° to 0.6° , with experienced naval carrier pilots giving the most accurate estimates. It was therefore decided to limit the maximum possible change in glide path stimuli to twice this value, or 0.5° . The aircraft simulated here was a light single-engined Cessna approaching at 80 miles per hour, which set time limits of 4.66 seconds at the 3000 ft range and 9.33 seconds at the 6000 ft range, thus keeping angular glide path changes for a given stimulus approximately the same at each range. The time was the same for all runs at the same range to avoid giving the subjects an artificial cue for the glide path and aim point estimates.

Standard deviations of estimates of aim point of 0.25° of flight path angle were also reported (Gold and Hyman, 1968; and Gold, 1973), so the aim point stimulus increments were also set at twice this, or 0.5° . The other experimenters used different visual systems, experimental protocols, and statistical analysis methods, but the results were considered to provide useful guidelines.

The shorter time of 4.66 seconds was more than enough for stabilized estimates of aim point. Palmer (1969) reports little change in error scores for estimates of aim point in an artificial visual field above 1.5 seconds viewing time.

All of the subjects here wanted more time, but it was not likely that their estimates would have actually improved. A gap of five seconds between each run gave the subjects time to make their estimates.

The first set of preliminary experiments used for refining experimental procedures had ranges of 1000 ft, 3000 ft and 10,000 ft. The closest range was too close to the runway (the optical probe on the terrain board sometimes hit its mechanical altitude stops), and the greatest range caused the subjects severe difficulties in making the aim point estimates. Since the revised experiments had only two ranges, the closest range was eliminated and the greatest reduced to twice the 3000 ft range, or 6000 ft.

Numerical values for the stimuli in the modified experimental design were as follows:

Glide path:	+ 3 deg nominal ± 0.5 deg and ± 1.0 deg deviations (± 1.5 deg scale calibration end-anchors)
Flight path:	nominal = total glide path ± 0.5 deg and ± 1.0 deg deviations (± 1.5 deg scale calibration end-anchors)
Range:	3000 ft and 6000 ft

As the glide path changes at a fixed range, different portions of the extreme foreground come into the subject's view. This potentially presents another artificial glide path clue besides run time (e.g., if the top of a tree is visible, the glide path is low). To compensate for this effect, small variations were included in the ranges. The variations were just large enough to make a foreground object 10% as high as the total vertical field of view disappear between two views at different variations. This required variations of $\pm 3\%$ and $\pm 6\%$ of range, depending on glide path. One of the three replications of each stimulus at each range was given a positive variation, one a negative variation, and one no variation, all chosen at random so that any effects due solely to the variations would average out in the data analysis. Starting points were chosen to make the views coincide with the desired range exactly at the middle of the run.

3.5 Experimental Procedure

Each subject was led through the already darkened simulator room to the cockpit simulator, which was dimly lighted inside. The subject was asked to sit in the pilot's seat on the left side of the cockpit and to review the Instructions to Subjects (included in Appendix B). A small ball on a wire

extending down from the cockpit ceiling served as a gauge to adjust the subject's eye position to a known point. This ensured that all subjects had the same view of the screen, which completely filled the field of view out the pilot's forward window.

Slightly behind the subject and on the right side of the cockpit sat the test monitor (the author), who controlled the video recorder and wrote down the estimates. A small television monitor, not visible to the subject, allowed the test monitor to watch the tapes to check that they were working properly and to make sure the subject did not fall behind in his estimates. Announcements of each run by number were recorded on the soundtrack immediately before each run, giving the test monitor another check on the experiment's progress, and providing the subject with a timing cue for making his estimate should he delay too long.

Each subject was shown the same two video tapes in the same order: glide path runs first, then aim point runs. Both tapes began with similar sequences of orientation, training, and practice runs. A long orientation run began at the nominal touchdown point and ran backwards up the glideslope to 10,000 ft; a second approach segment covered the same range going down normally from 10,000 ft to 0 ft. Four training runs then showed the two most extreme stimuli, positive and negative, at 3000 and 6000 ft ranges. These were

the end-anchors to which the subjects were told to assign estimate magnitudes of +10 and -10. A short orientation run (7500 ft to 1500 ft) followed, then several practice runs (eight for the glide path tape, seven for the aim point tape). Any problems with understanding the experiment or getting adjusted to the experiment were worked out with the subject here. One last short orientation run came after the practice runs. The experiment was always halted here to give the subject a chance to ask any questions before proceeding with the data runs. Orientation, training, and practice runs all took about seven minutes on each tape.

126 data runs were next shown to the subject. (Their configuration is described above in Section 3.3, Practical Limitations of the Experimental Design, and 3.4, Choice of Stimuli). The data runs were identical copies for both glide path and aim point tapes. These took 25 minutes for each tape.

The subject estimated glide path as being "high" or "low" with respect to the nominal touchdown point, basing his magnitude estimates on the +10 to -10 scale. To keep the subject interested and to raise a competitive spirit, scores of the subject's performances were kept and revealed to him during the experiment. The score was the number of estimates in the correct direction, and was announced by the test monitor after every ten runs.

CHAPTER 4

RESULTS

The raw data were lumped together and analyzed in two groups: all subjects together and the five high-time (over 1000 hours flying experience) subjects together. The comments in this section refer to the grouped subjects' estimates of glide path and aim point, as modelled by the ANOVA program (discussed in Section 3.1, ANOVA and the Statistical Model). Results for each individual subject are given in Appendix C. All interpretations are based on results from the two groups of combined data.

The overall average, or mean, of all estimates of either glide path or aim point for all subjects in each group is given when significant with each of Figures 4.1 through 4.4. Significant effects other than the mean are given with their significance levels. Also given is the model RMS error, a statistical estimate of the overall standard deviation for all subject estimates included in the group model. The model RMS error includes errors due to mis-fitting of the model curves and is therefore slightly larger than the true standard deviation. (See the end of Appendix C for plots of means and standard deviations calculated directly from the raw data.) Note: The lines connecting data points on the plots serve only to illustrate the patterns in the responses. The statistical modelling method used here does not predict responses between data points.

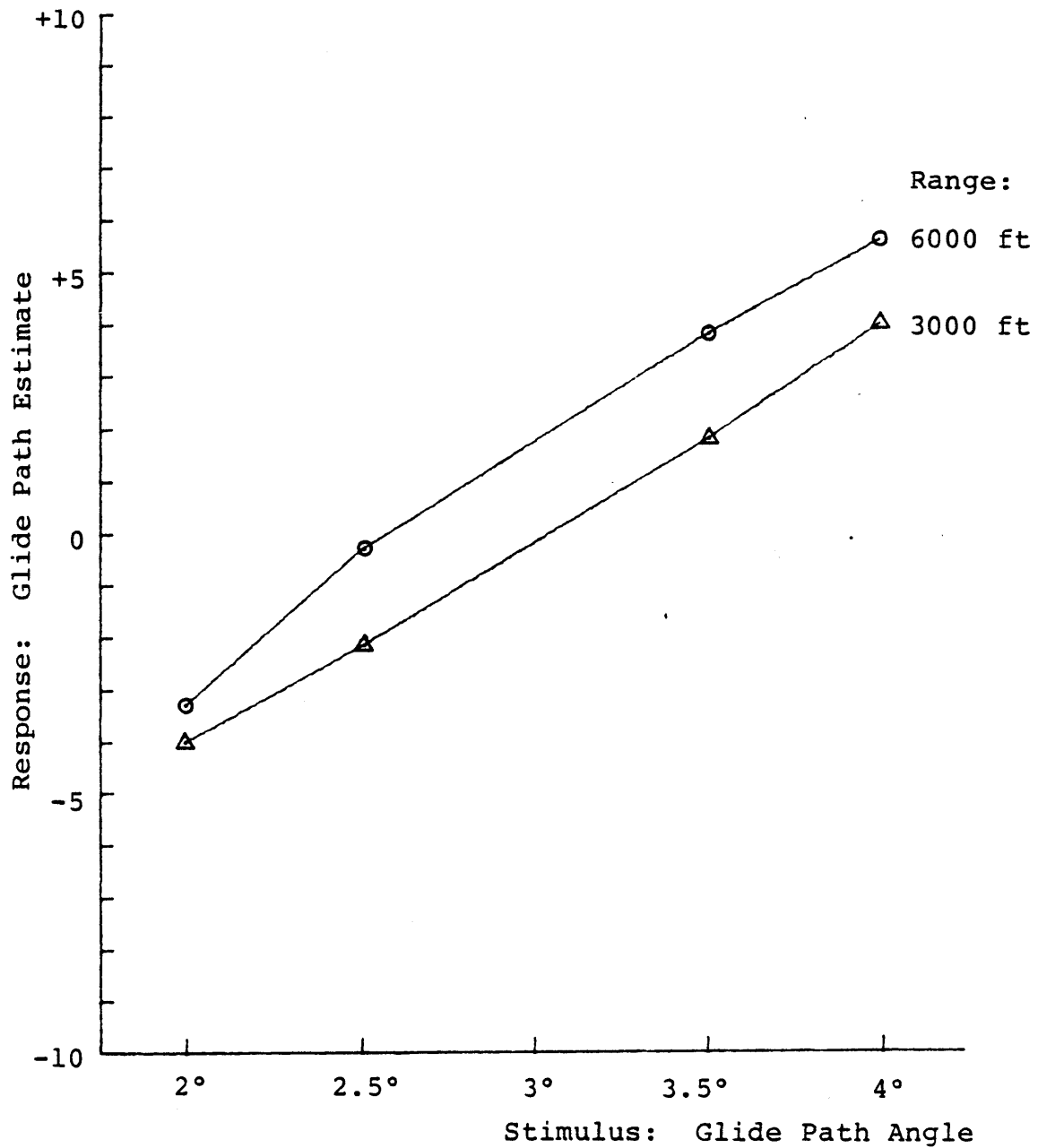


Figure 4.1 Glide path estimate model for all subjects combined.

Significant effects: glide path and range ($p < .005$)
range/glide path interaction ($p < .05$)

Mean = 0.71

Model RMS error = 2.50

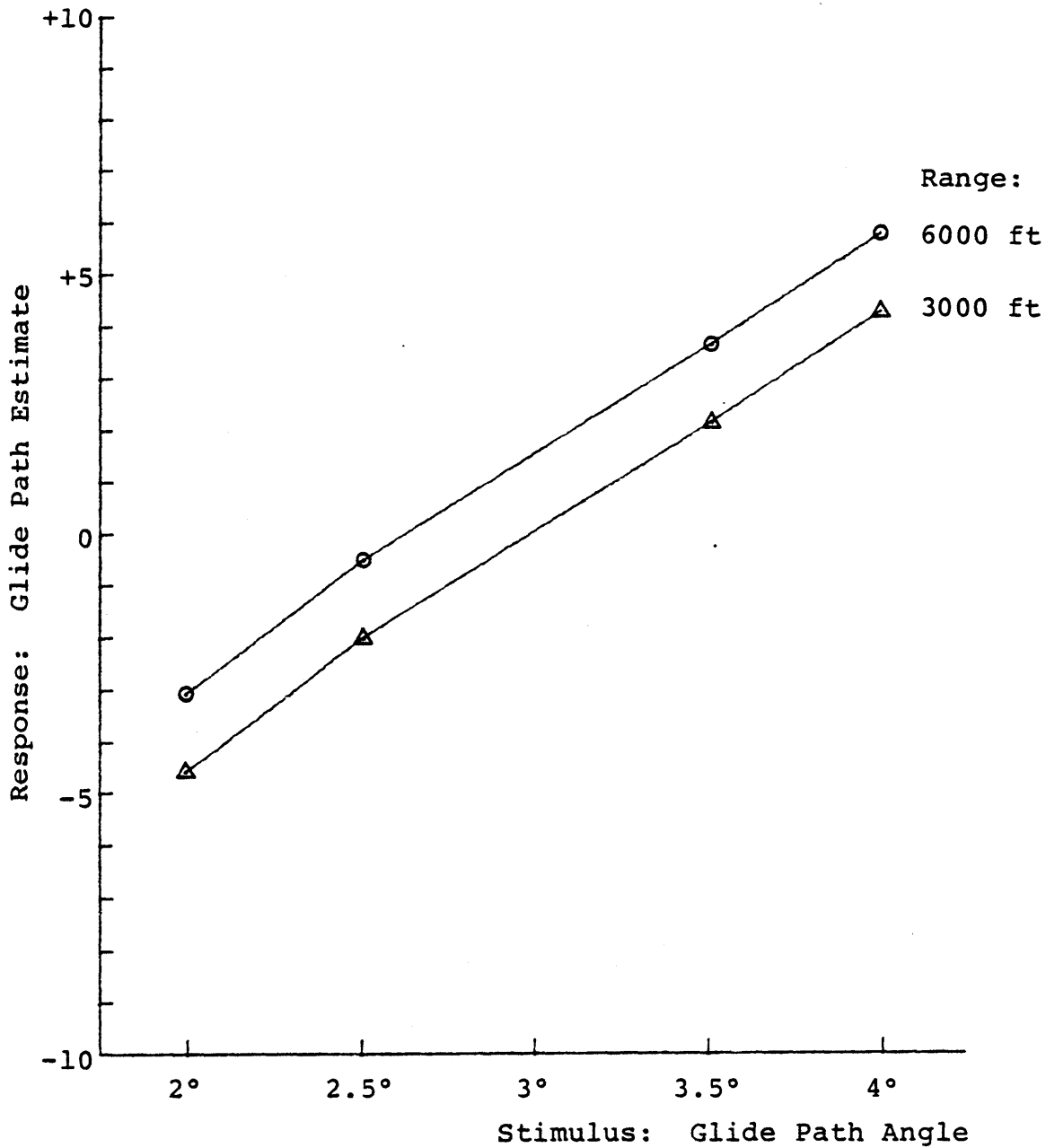


Figure 4.2 Glide path estimate model for high-time pilots.

Significant effects: glide path and range ($p < .005$)

Mean = 0.73

Model RMS error = 2.24

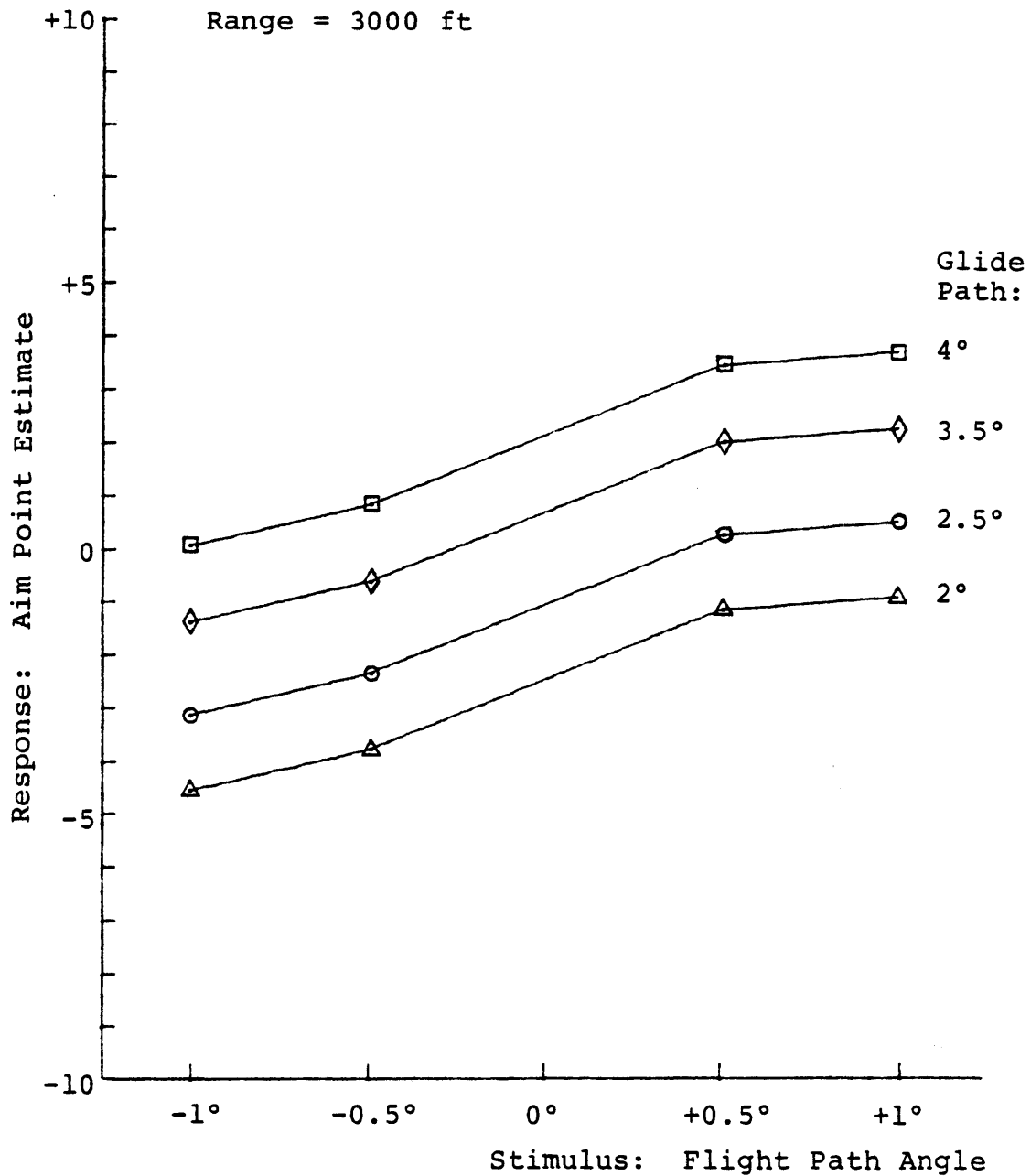


Figure 4.3a Aim point estimate model for all subjects combined (at 3000 ft range).

Significant effects: aim point, range, glide path, and range/glide path interaction ($p < .005$)

Mean = 0.41

Model RMS error = 3.08

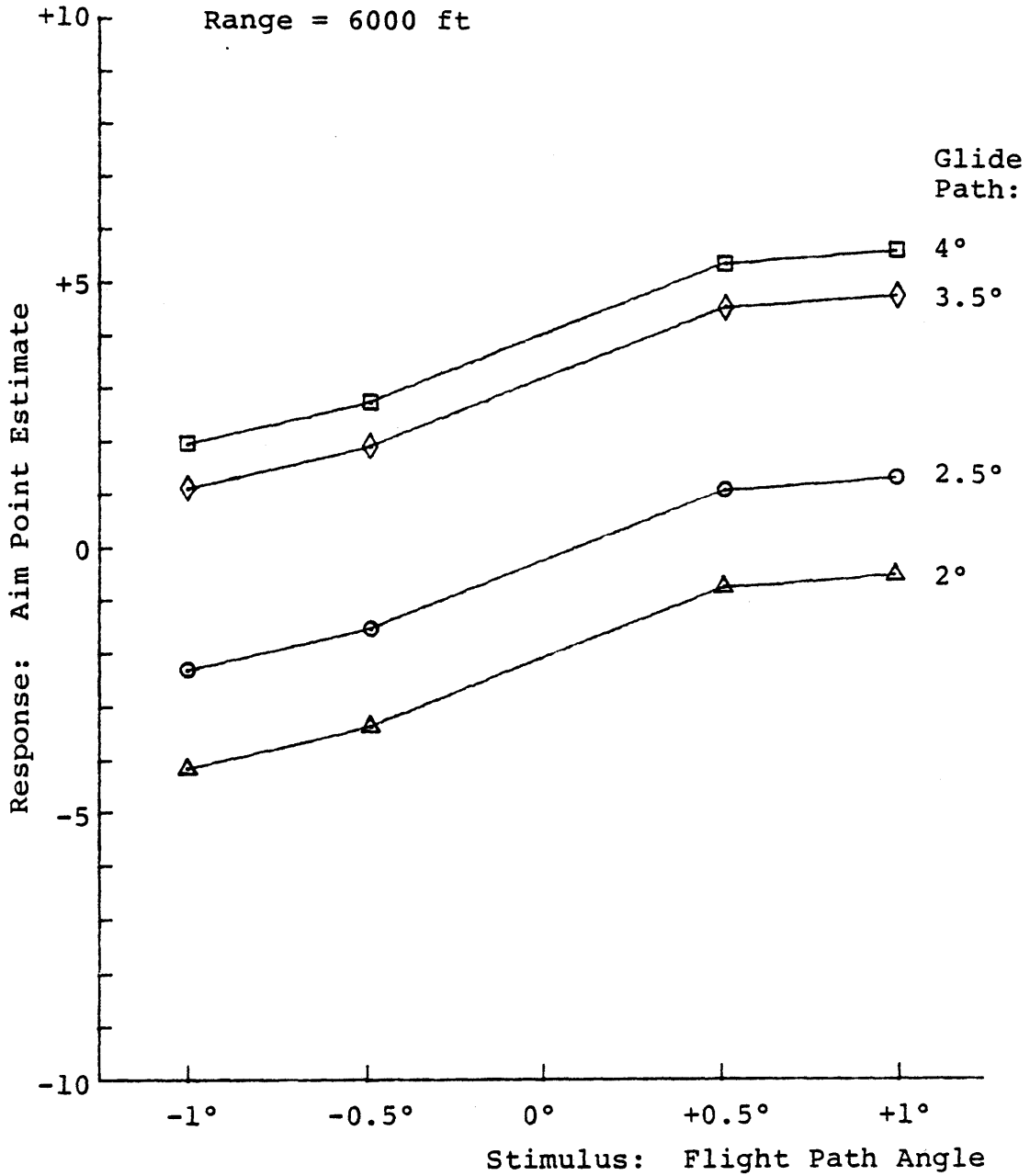


Figure 4.3b Aim point estimate model for all subjects combined (at 6000 ft range).

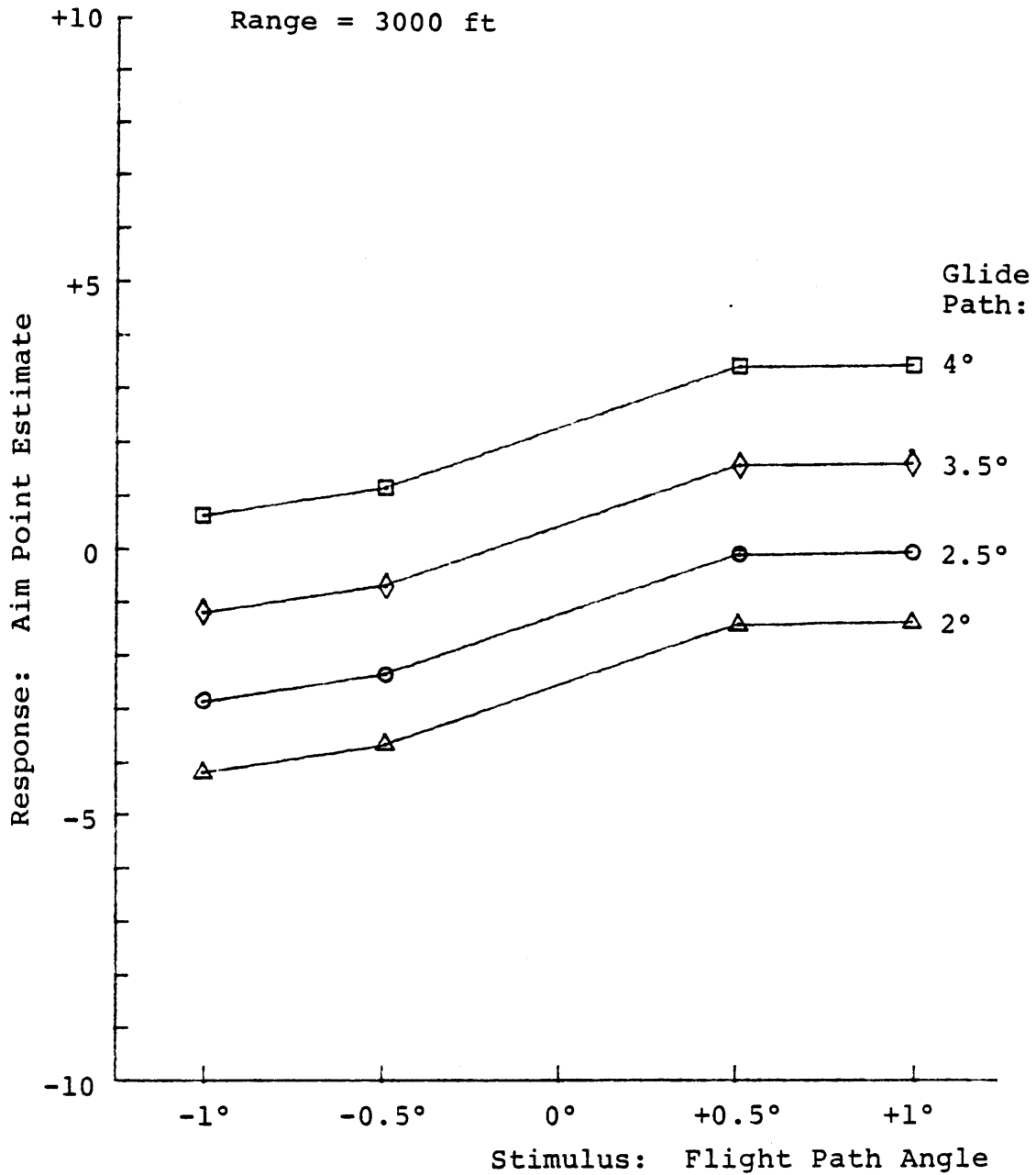


Figure 4.4a Aim point estimate model for high-time pilots (at 3000 ft range).

Significant effects: aim point, range, and glide path ($p < .005$)
range/glide path interaction ($p < .05$)

(Mean not significant)

Model RMS error = 2.75

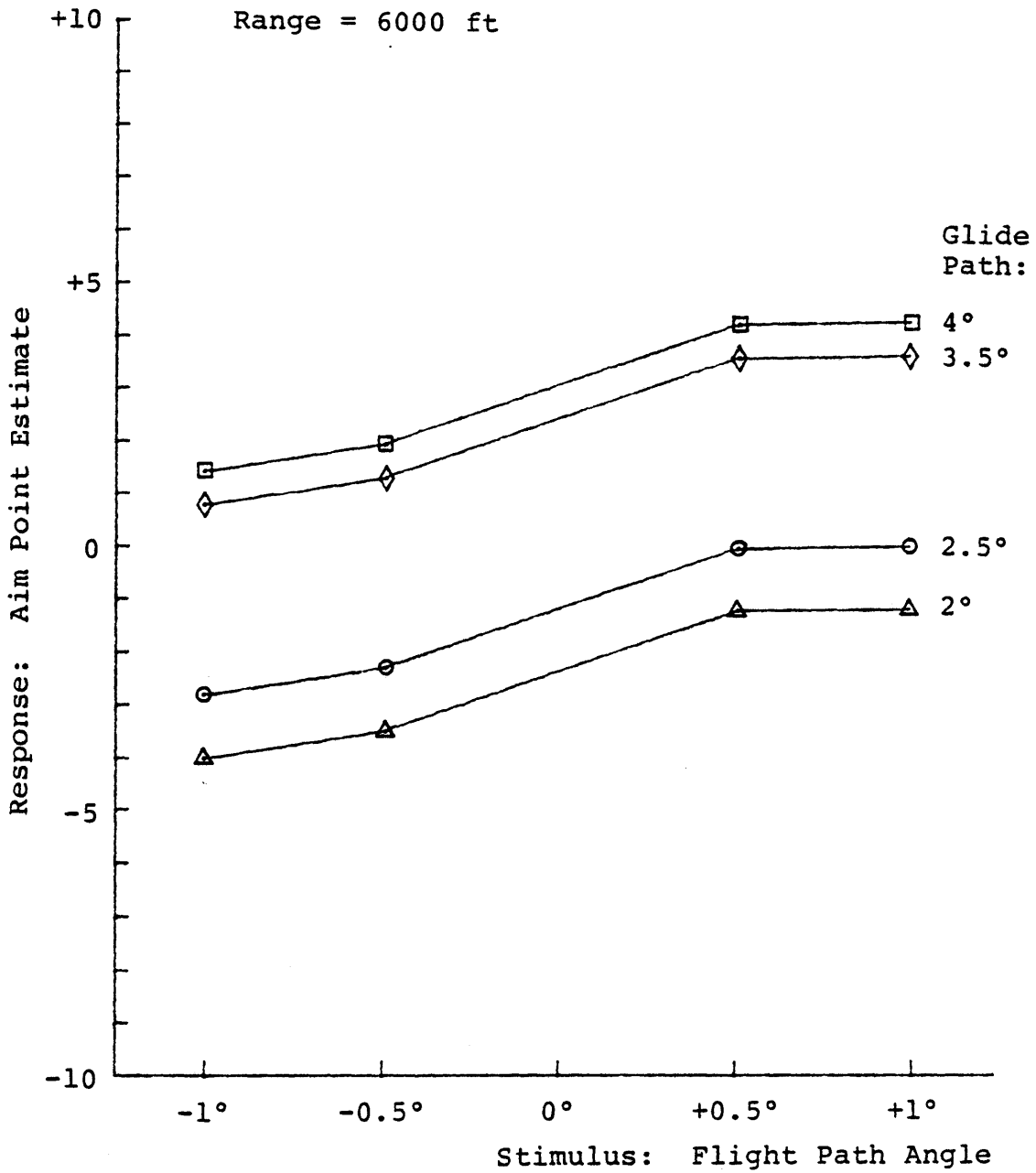


Figure 4.4b Aim point estimate model for high-time pilots
(at 6000 ft range)

4.1 Results for Grouped Data

The glide path estimates were nearly linear, especially for the five high-time pilots (see Figures 4.1 and 4.2). It is convenient to define sensitivity as the total change in the estimates relative to the defined estimate scale (-10 to +10), divided by the total change in stimulus magnitudes relative to the maximum range (set by the stimulus end-anchors). This gives a non-dimensional number which may also be described as the ratio of percentage change in response to the percentage change in stimulus. A perfect subject would have a sensitivity equal to 1.0. The group of all subjects combined had a sensitivity of 0.63; the group of high-time pilots had an average sensitivity of 0.66. This reduced sensitivity, represented by a reduced slope of the plotted response curves, is referred to here as "scale compression". The effect was also seen by other researchers, who reported glide path sensitivities of 0.88 (Gold and Hyman, 1968) and 0.82 (Gold, 1973).

The aim point estimates for both groups were less linear, with noticeably reduced sensitivities at the extremes of the stimuli (see Figures 4.3 and 4.4). Scale compression was more severe than for the glide path estimates. The group of all subjects had an average aim point sensitivity of 0.30; that of the high-time pilots was 0.24. This rather low sensitivity to aim point stimuli was also noticed by other

researchers, who calculated sensitivities of 0.39 (Gold and Hyman, 1968) and 0.41 (Gold, 1973).

These results are reasonably consistent with the other reported experiments. In all cases, the aim point sensitivities were less than half those for the glide path. It should be noted here that scale compression is not necessarily due to perceptual processes. Failure to follow an artificial scale is common in many types of psychophysical experiments, particularly in magnitude estimation experiments where no attempt is made to match stimuli to favored sets of response values (Poulton, 1968). The scale compression exhibited here probably results from the experimental method used, at least for glide path estimates. (The problem of magnitude estimation experiments are discussed in Section 3.2, Magnitude Estimation).

As for statistical results, range had a significant effect ($p < 0.005$) on the glide path estimates of both groups. This is seen (in Figures 4.1 and 4.2) as an increase in the glide path estimates at the greater range by about $1/3^\circ$ (after correction for scale compression). Although slightly less than the model RMS errors, which were 2.50 for all subjects combined and 2.24 for the high-time pilots, the increase was consistent and therefore significant. (Figures C.17 and C.18 show means and standard deviations calculated from the raw data at each combination of significant stimuli.) There was also a significant interaction between range and glide path ($p < 0.05$) for the group of all subjects. The effect of this interaction was

a slight increase in sensitivity to low glide path stimuli at the greater range (Figure 4.1). (See Section 3.1, ANOVA and the Statistical Model, for a discussion of interactions and the statistical method.)

For both groups - all subjects and the high-time pilots - range and glide path each had significant effects ($p < 0.005$) on the aim point estimates, and the interactions between range and glide path was also significant ($p < 0.05$). Aim point estimates increased slightly with increasing range, and markedly with increasing glide path, due to the range and glide path effects respectively (Figures 4.3 and 4.4). Differences between adjacent glide paths were usually less than the model RMS errors, which were 3.08 for all subjects combined and 2.75 for the high-time pilots; but the total differences between the highest and lowest glide paths were twice the model errors, and all differences were sufficiently consistent to be significant. (Figures C.19 and C.20 show means and standard deviations calculated independently of the model at each stimulus combination.) Since flight path is measured from the current glide path, the glide path effect on aim point measurement is a real perception effect and not an artifact of geometry (see Figure 1.2). The range/glide path interaction caused the curves to be less evenly distributed at the greater range, grouping into pairs at high and low levels of glide path stimuli (Figures 4.3b and 4.4b; compare with Figures 4.3a and 4.4a). (Note that glide path stimuli themselves are

not evenly distributed: there is no stimulus at 3° .) In all cases, the slopes of the curves (sensitivities) decreased at the extreme stimulus levels.

4.2 Interpretation

4.2.1 Glide Path Estimates

It appears from Figures 4.1 and 4.2 that pilots can estimate glide path in a consistent linear fashion. Scale compression was noticeable. As discussed immediately above and in Section 3.2, Magnitude Estimation, it is most likely an artifact of the experimental method.

The increase in glide slope estimates at greater ranges is more difficult to explain, especially as it contradicts the range-independent results obtained by Gold and Hyman (1968). The answer may lie in the visual system used. Textural details are thought to be important in judgments of altitude, at least at close range (Barnes, 1978). Supposedly, Tiger Moth pilots used to flare when they could see individual blades of grass. While this technique would not work for a large transport or military fighter pilot, or even a modern light aircraft pilot operating on paved runways, it makes plausible the hypothesis that the presence of textural detail is an important cue signalling closeness to the ground. It is consistent with duck-

under during night approaches, where this cue would not be present when expected. The low resolution of the video system used in these experiments may not allow a natural transition of textural quality in the displayed image, and the discontinuous nature of the stimuli would exaggerate apparent differences in textural details between the two ranges. This would increase the relative effect of any textural cues at the closer range, leading to lower glide path estimates than at the greater range.

4.2.2 Aim Point Estimates

Looking at the modelled aim point responses for the two combined groups of subjects, several things are immediately noticeable:

- (1) A very small change in total estimated aim point magnitude with range. For the high-time pilots, the change in total range of estimates - lowest to highest - was only 10% between the two ranges; the change for the group of all subjects was about twice as much.
- (2) Compression of the estimation scale relative to the ideal.

- (3) Reduced sensitivity (slope) at the extremes of the stimuli.
- (4) Clustering of the curves into distinct pairs at high and low values of glide path at the greater range.
- (5) Greater sensitivity of aim point estimates to changes in glide path than to changes in aim point itself. The high-time pilots were about twice as sensitive to glide path stimuli as to aim point.

It is apparent that the subjects were not estimating ground distance, but were, in fact, responding primarily to changes in flight path angle, as was desired. This is not to say that absolute ground distance had no effect at all; indeed, the estimates did increase slightly with increasing range from the runway, but by an order of magnitude less than the actual change in absolute aim point distance. At double the range, the aim point ground distance associated with a given flight path angular error should also double, but the estimates increased a maximum of 25% (compare Figures 4.3a with 4.3b), and usually changed much less. Comparing these curves with the plot of Estimate of Ground Distance for the Ideal Pilot (Figure 4.5), it can

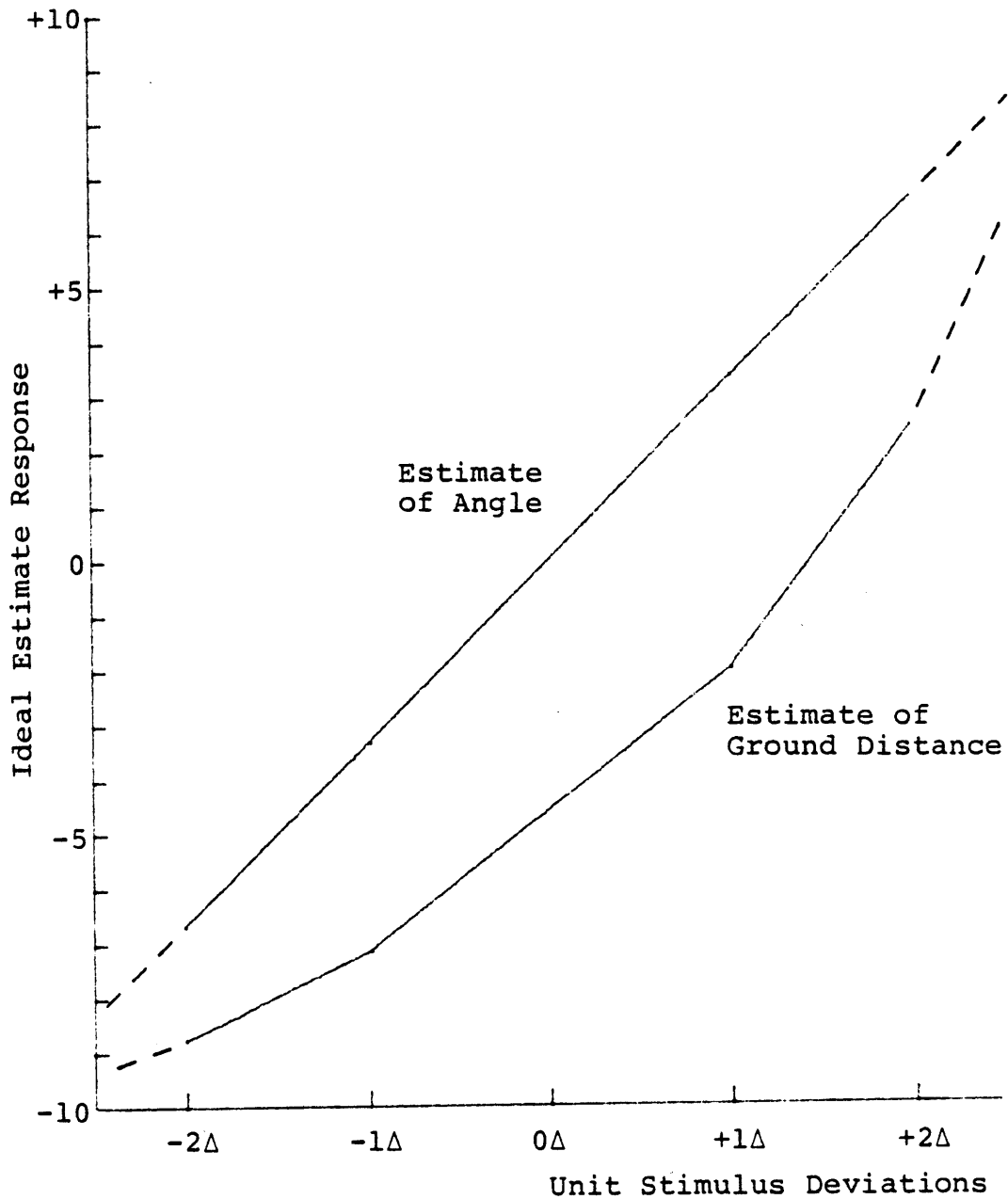


Figure 4.5 Estimates of aim point angle and ground distance by an ideal pilot.

The lower curve shows the effect of trigonometric nonlinearities on aim point ground distance, given linear angular stimuli.

be seen that they haven't the same shape at all, the modelled estimate curves being approximately antisymmetric instead of continuously increasing in slope. So the pilots were able to distinguish very well between ground distance to the aim point and flight path angle, although their estimates of magnitude were of poor accuracy. Other minor ground distance effects were evident in an interaction between range and glide path effects, discussed below.

The scale compression is again entirely reasonable here. The reduction in sensitivity at the extremes of the stimuli may be due to the practice of trying only to null out any errors during actual approaches, rather than somehow using exact magnitudes to determine aircraft control inputs, especially in this case where the errors in aim point are difficult to estimate relative to others such as altitude, lateral alignment, etc. (Naish, 1971). If ground distance is used at all in the perception process, even if mentally normalized with respect to total range from the runway, the high nonlinearity of the stimuli (ground distance versus flight path angle) could be confusing, so that on the average the subjects could distinguish well between different directions only, but not between different magnitudes.

The tendency to fly approaches using a nulling technique - that is, mentally defining a proper approach as one that takes the aircraft back to the nominal glideslope - was

mentioned by several subjects as feeling natural, even though the proper approach is then variable and depends on the initial error. Flying is a dynamic process where any error, once perceived, immediately becomes less important than the action needed to correct it. Nulling behavior helps explain the increase in aim point estimates with increasing glide path stimuli. As the size of the glide path error increases, so does the size of the flight path angle needed for correction, but in the opposite direction (e.g., the higher a pilot is on the glide path, the steeper he must dive to get back to the nominal glideslope). But the nominal glide path, as defined in this experiment, merely took the aircraft to the touchdown point, and did not intersect the glideslope until touchdown; hence the nominal flight path angle was smaller than desired by the subjects. (See Figure 4.6.) This made nominal flight paths (and their associated aim points) look too long at high glide paths and too short at low glide paths.

Airspeed, angle of attack, power, and altitude are of more immediate concern to a pilot than aim point, at least in conventional aircraft flying standard landing approaches. Avoiding a stall or loss of more height than can be regained in time are of primary importance. Even in aircraft carrier landings, where achieving an exact touchdown point is essential, pilots are trained to de-emphasize aim point cues and concentrate instead on staying on the glideslope and keeping

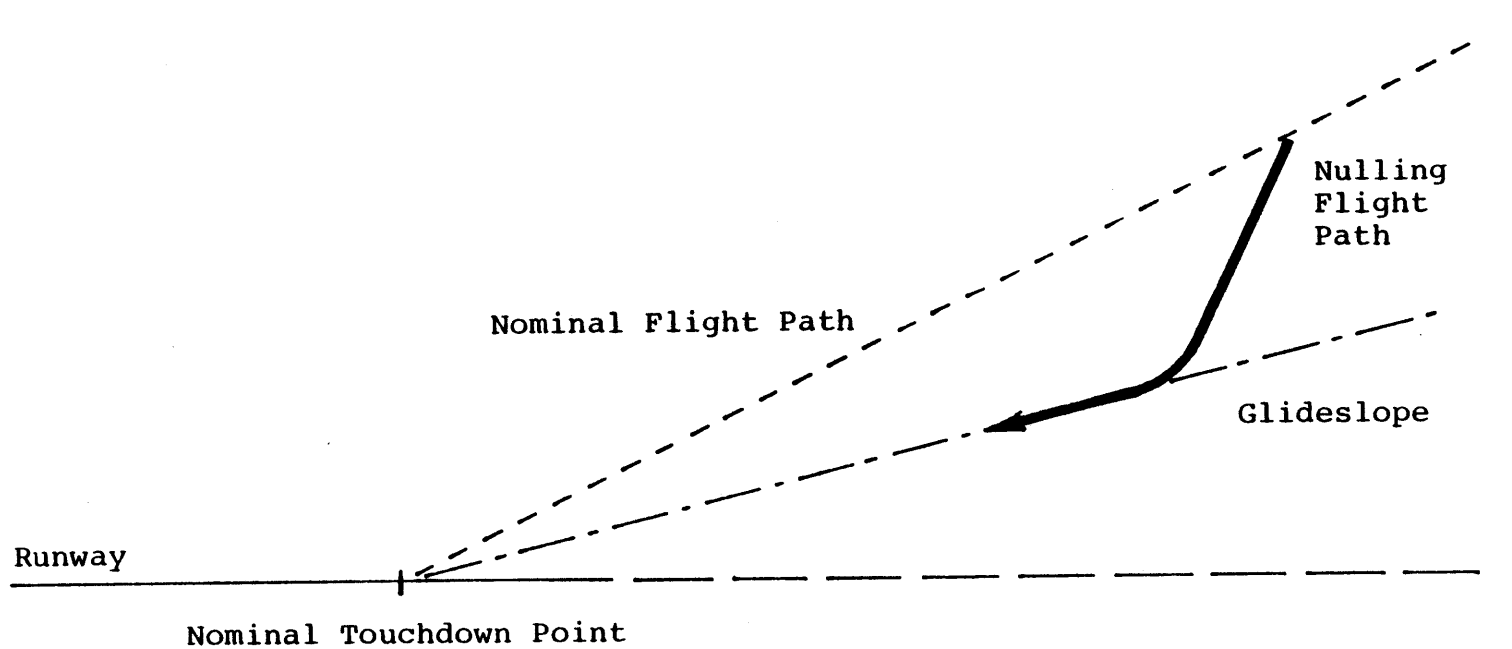


Figure 4.6 Illustration of a nulling approach. (Angles are exaggerated.)

within their aircraft's performance limits. The trained behavior of getting back on the nominal glideslope as quickly as possible and then following it to touchdown would cause the response behavior seen here.

Another possible explanation for the effect of glide path on aim point estimates is that subjects could not accurately detect absolute aim point errors - either in ground distance or in flight path angle - but could detect only relative changes. In the experiments, all flight path angles were referenced to the current glide path, not the nominal glideslope. If the subjects could not properly determine the actual flight path, they might base their estimates on the expected or average flight path, due in part to confusion over what was a proper flight path or aim point. The responses would then be dominated by the statistically average stimulus, which had an absolute flight path angle of 3° below the horizon. This would again result in longer aim point estimates at higher glide paths and shorter estimates at low ones.

In principle, a judgment of aim point can be made directly by using the "expansion" or "streamer" effect. All points on the ground will appear to expand outward from the aircraft's actual aim point. In conjunction with a reference such as a windshield frame, this provides a cue for determining the aim point; it is sometimes called the "gunsight method" (Hasbrook, 1975). But its theoretical accuracy is very low, and nearly

useless until almost over the runway itself (Naish, 1971). This is supported by other experimental work which showed poor accuracy of aim point judgment (Palmer, 1969). Although they can be clearly seen by most pilots, expansion cues would provide no more than a coarse indication of aim point, at best.

The term "streamer effect" is sometimes also used to describe peripheral vision motion cues (Hasbrook, 1975). There were none in these experiments, as the subjects only had a direct forward view out of the cockpit. It does not appear that such cues would be of much aid in determining aim point until very close to the ground anyway. Expansion and streamer cues would probably be of most use in making final corrections for the flare.

It was thought that pilots may estimate flight path angle by looking for changes in glide path angle. But the experiments were set up so that the maximum change in glide path angle seen during any landing approach run was only $1/4^\circ$, or the expected standard deviation of pilot's estimates of glide path (Gold and Hyman, 1969). These changes would therefore be marginally detectable, but the pilots could in fact detect changes in aim point. Furthermore, trigonometric nonlinearities cause faster rates of change of glide path for any given change in flight path as the average glide path angle increases. This should cause increased sensitivity at high glide paths and statistical interactions between aim point

and glide path, but neither of these effects was seen in the grouped data.

The possibility that pilots rely on perception of vertical velocity (rate of descent or sink rate) during their estimates of aim point was considered. In these experiments, the downwards vertical velocity was on average higher at high glide paths, so that aim point estimates should have been lower at high glide paths according to the hypothesis. But exactly the opposite results were obtained. There should also have been interactions between aim point and glide path due to trigonometric nonlinearities, but, as mentioned above, none were found in the data.

The magnitude and the range of the aim point estimates changed little with range from the touchdown point, but the distribution across glide path stimuli changed appreciably. The clustering at the greater range, with high glide path curves closer together than low glide path curves, is consistent with the nonlinearities expected in absolute ground distance aim point estimation. It is possible that pilots may look for absolute ground distance errors, and correct them for total range from the runway to get angular estimates. The low accuracy of such information makes its usefulness questionable, and this is almost certainly not the most important perceptual mechanism.

Pilots can estimate aim point errors reliably, if not very accurately. However, these experiments do not allow a

distinction between the two possibilities: a preference for nulled approaches, and poor accuracy of aim point perception. Moreover, the two are not mutually exclusive, and depend upon the pilot's training and experience. If pilots cannot rely on their own perceptions of flight path angle, they may be strongly dependent upon simply being on the right glide path, trusting it to eventually take them to the correct aim point. Aim point itself may be of little concern until the flare, which was not included here. Nevertheless, there was still a definite and consistent sensitivity to changes in flight path angle (hence aim point).

CHAPTER 5

CONCLUSIONS AND SUGGESTIONS FOR FURTHER RESEARCH

5.1 Pilot Performance

Glide path presented no major difficulty in estimation to pilots. The estimates made by the pilot subjects in these experiments were usually quite linear. A change in the magnitude of the glide path angle stimulus was usually met by a smaller change in the subject's responses, and glide path estimates were higher at the greater range. However, these two response phenomena may have been due respectively to the method of magnitude estimation used for measuring the subjects' perceptions and to an imperfect visual system.

Aim point errors were relatively hard to estimate. Although the subjects could reliably detect large changes in flight path angle, their sensitivity to small deviations was very low. Glide path strongly influenced the magnitudes of the aim point estimates, which were longer at high glide paths, but the sensitivity to changes in flight path angle was not affected by glide path. Range from the runway had little effect. There was some evidence that aim point ground distance along the runway played a minor role in the estimation process. The data did not support the possibility that pilots use either vertical velocity or changes in glide path to estimate aim point.

While these results for aim point estimates could be due in part to the same problems with magnitude estimation as were noted for glide path estimates, the poorer performance indicates that other factors must be involved. One obvious possibility is simply that pilots have low accuracy of aim point perception, but another is that pilots prefer to null their approaches to the correct glideslope, and do not worry about aim point until near the flare.

If pilots do prefer to use a nulling technique to stay on a pre-chosen approach path or glideslope, and this is due to difficulties with perceiving aim point and not just to training conventions, then pilots flying difficult approaches will need cockpit aids to determine their true aim point. STOL and multi-segment approaches would present the worst problems, especially without increased aircraft performance margins. Pilots cannot control aim point nearly as easily as glide path, which can be chosen to intersect a desired aim point, then simply followed down to it. Any artificial aim point display should be integrated with a glide path display to be readily usable, as the aim point helps define the desired glide path.

The presumed preference for nulled approaches, however, needs to be better established first. A fairly simple modification to the experiments would be to have the subjects estimate the rate at which they are converging on or diverging from the

glideslope, making no mention of aim point. The stimuli should be recalibrated to make a zero-deviation flight path parallel to the 3° glideslope regardless of altitude, instead of converging on the nominal touchdown point. A parallel flight path moves neither closer nor further from the glideslope, so any such flight path may serve as an equivalent angular reference. Greater accuracy of perception, linearity of response, or even just higher sensitivity to flight path angle changes in such experiments would support the nulling hypothesis.

5.2 Experimental Procedure

Aside from the type of stimuli presented, the method of presentation should be modified in future experiments. The experiments seemed very long: half-hour sessions are barely acceptable. More data points are, as usual, desirable. Breaking up the experiment into several shorter sessions, each session with only one replication of the stimuli, would help make a longer experiment tolerable.

Data filling in the gap near zero would be useful for determining the true shape of the response curves before attempting a regression analysis. Either adding zero deviation data runs or pulling the smaller magnitude stimuli together to get more uniform data point intervals would work. It is probably more important in a magnitude estimation experiment

to make sure that each stimulus level falls on a whole number of the estimation scale than to have exactly equal intervals between all stimulus levels.

Adding a third range could also be done without extending a single replication much beyond 15 minutes, if scale calibration end-anchors were used sparingly. Alternating between glide path and aim point estimates for each subsequent session would help relieve the boredom, as would a forced break in the experiment every two or three sessions. Six 15-to-20 minute sessions - three replications for both glide path and aim point - could be fit into three hours, including plenty of time for instructions and rest breaks.

The long orientation runs at the beginning of the first tape were appreciated by the subjects, but were boring when repeated on later tapes. More practice runs on the first glide path and aim point sessions are desirable, but are probably unnecessary for any later sessions. The subjects would have preferred to see the calibration runs repeated after the practice runs, and the calibration runs should also be shown before each replication in a multi-session experiment.

The video tape/television projector system is useful for working out flaws in the experimental design and procedure, but difficulties with the equipment make its usefulness for getting final hard data questionable. Preliminary experiments

should definitely be run on this system with a few subjects to ensure that any experiment will actually run smoothly and produce the desired data. It would then probably be more efficient to run definitive experiments on a more sophisticated simulation system at a NASA facility. This would at least verify the performance - good or bad - of the simpler simulation.

APPENDIX A
SUBJECT DATA

SUBJECT	AGE	VISION	RATING(S) AND FLIGHT EXPERIENCE
1	24	20/20 (corrected)	Student, 31 hours light civil
2	47	20/20	Flight instructor, instrument, commercial instrument multi, 1200 hours military jet, 1500 multi, 1000 light civil
3	29	20/20 (corrected)	Private single VFR, 400 hours light civil, 1000 other
4	29	20/20	Flight instructor multi instrument, 1000 hours military transport, 500 military jet, 2000 light civil, 500 helicopter, 250 other
5	43	20/15	Commercial multi instrument, 800 hours military jet, 400 light civil
6	34	20/20 (corrected)	Commercial single instrument, 530 hours light civil
7	28	20/20	Commercial single instrument, 850 hours military jet, 200 light civil
8	30	20/15	Commercial multi instrument, 1400 hours military jet, 200 light civil

APPENDIX B
INSTRUCTIONS TO SUBJECTS
FOR LANDING APPROACH ESTIMATION EXPERIMENT

The purpose of this experiment is to determine your ability to detect errors in glide path and aim point during aircraft landing approaches. The experiment has two sets of video-taped landing approach runs. To save time, only a short segment of each run is shown. During each set, you will be asked to estimate either glide path or aim point errors for each run. Both kinds of errors may occur simultaneously, but you should estimate only the one asked for. Tell the test monitor your estimate at the end of each run. He will write it down for you so that you can concentrate on watching the approaches. Since altitude along the glide path and aim point miss distance depend on initial distance from the runway, you should base your estimates on the angles of the glide path and aim point vector errors. (See the descriptions below and the figures on the next two pages.)

[Same as Figures 1.1 and 1.2.]

Each set of runs begins with two orientation runs to show you the touchdown point and a correct approach. Four scaling runs follow to show you the largest errors in that set for either glide path or aim point. You should call the maximum positive and negative glide path errors "10 high"

"10 low" respectively, and estimate all glide path errors in terms of this scale. For example, a positive glide path error half as large as the maximum should be estimated "5 high". Similarly, aim point errors should be called "long" or "short". Except for orientation runs, there are no normal approaches (with error equal to 0).

The "glide path" is the path through space that would take you to the runway touchdown point. The correct glide path is the "glideslope", which here makes a 3° angle to the horizontal. For any given glide path error, the difference in altitude will change with distance from the runway, so you should estimate angular error of the glide path (the glide path error angle). If you are above the glideslope, call the error "high", and if you are below the glideslope, call it "low", with the appropriate magnitude.

The "flight vector" is the direction you are moving through space. The "aim point" is the place on the ground that you will reach if you continue along your present flight vector. The correct aim point is simply the runway touchdown point; to reach it, the flight vector must be exactly aligned with the glide path. In an actual aircraft, only the instantaneous flight vector angle can be controlled directly, not the ultimate aim point, and this experiment is set up accordingly. For any given flight vector angle error, the ultimate touchdown point depends on the initial distance from the runway.

Also, the absolute size of the aim point error is not symmetrical for initial symmetrical flight vector angle errors. So you should estimate the error of the flight vector angle, rather than the ground distance of the resulting aim point. Estimate the directions of the error as being "long" or "short" with respect to the touchdown point.

Note that it is possible to reach the correct touchdown point even if the glide path is incorrect, and that the aim point can be in error even if you start out on the proper glideslope. If the flight vector is not aligned with the glide path, you may notice a slight change in the glide path during the run. If so, simply estimate the average glide path, or that at the middle of the run.

A score of your performance during the test will be kept. You will not be scored on correctly estimating the exact size of the error, just the right direction (high/low, long/short). Your score is simply the total number of estimates in the right direction. Your score does not represent your actual abilities as a pilot in a real aircraft and will be kept confidential.

The runs average about 7 seconds long each (5 to 9 seconds), with three seconds between runs, so you should make your estimates quickly. You will have several practice runs, and you may repeat the scaling and orientation runs, if you wish.

APPENDIX C
INDIVIDUAL RESULTS

C.1 Individual Estimate Models

The comments in this section refer to the subjects' individual estimates of glide path and aim point, as modelled by the ANOVA general linear model program. (Analyses were also performed on the data combined into two groups: all subjects together, and the pilots with over 1000 hours flying time together. Those results are discussed in Section 4.1, Results for Grouped Data. See also Section 4.2, Interpretation.) All figures are given in the order of their subject numbers.

All eight subjects exhibited scale compression in their glide path estimates. That is, a percentage change in glide path stimulus produced a smaller percentage change in estimate response, so that the slopes of the plotted curves are less than that for an ideal pilot.

Range had a significant effect ($p < .05$) on the glide path estimates for seven subjects, all but # 6, including all five high-time pilots. Glide path was always estimated to be higher at the greater range, usually by about $1/4^{\circ}$. Only one subject, low-time pilot #3, had any significant interaction between range and glide path ($p < .01$). The interaction yielded atypical behavior at the greater range: The slope of the curves changed an unusually large amount between low and

high values of glide path stimuli, but only at this range.

Subject #3 also had a significant interaction between range and aim point ($p < .025$) on his glide path estimates, although the aim point main effect itself was not significant ($p > .25$). One high-time pilot, #8, also had a significant interaction between range and aim point ($p < .05$) without a significant aim point main effect. Although there is no consistent trend of effects of aim point on the two subjects' glide path estimates, the interaction caused a general reversal of aim point effect between the two ranges for both subjects. (The presence of an interaction between range and aim point precludes drawing any firm conclusions about the significance of either range or aim point, so both of those main effects are included in the models for subjects #3 and #8).

For the aim point estimates, scale compression - reduction in slope from the ideal - was usually more severe than for glide path estimates. However, for several subjects glide path effects spread out the aim point estimates over a greater portion of the estimation range, sometimes making greater contributions to the estimates than did the aim point effects themselves.

Several subjects appeared to be at least as sensitive to glide path stimuli as to aim point stimuli when making their aim point estimates. Indeed, the sum of squares of aim point

estimates due to aim point stimuli was less than (or about equal to) that due to glide path stimuli for half of the subjects - #2, #3, #7, and #8- all but #3 high-time pilots. There was a consistent trend to make longer aim point estimates at higher glide paths, and a tendency to make longer estimates at the greater distance.

Glide path stimuli had a significant effect ($p < .005$) on the aim point estimates of all eight subjects. The estimates always increased with increasing levels of glide path stimuli. Range was significant for five subjects: #1, #2, #3, #4, and #6. (The three subjects for whom it was not significant were all high-time pilots.) For these five subjects, aim point estimates were always higher at the greater range. Interaction between range and aim point was significant for only one subject, low-time pilot #6. His behavior was different at each range: At low levels of aim point stimuli, the slopes of his estimates increased between the smaller and greater ranges, but at high levels of aim point stimuli, they decreased, actually becoming negative at the greater range. Three subjects - #2, #6, and #8 - had significant interactions between range and glide path ($p < .05$). For subjects #2 and #6, this caused the aim point estimate curves to spread out at the greater range, meaning that the glide path stimuli had stronger effects there. Subject #8, however, showed the reverse effect, with his curves

spread out at the smaller range. Note that subject #8 did not have a significant range main effect, but the interaction between range and glide path required it to be included in his aim point estimate model.

A fairly regular pattern of reduced slope of the aim point estimate curves at the extremes of the aim point stimuli was present. Only one subject, high-time pilot #2, failed to exhibit this behavior. In fact, for two high-time pilots - #7 and #8 - the slopes of the curves was negative at one or both extremes. Subject #6 behaved anomalously at the greater range only, with slopes which constantly decreased with increasing aim point stimulus levels.

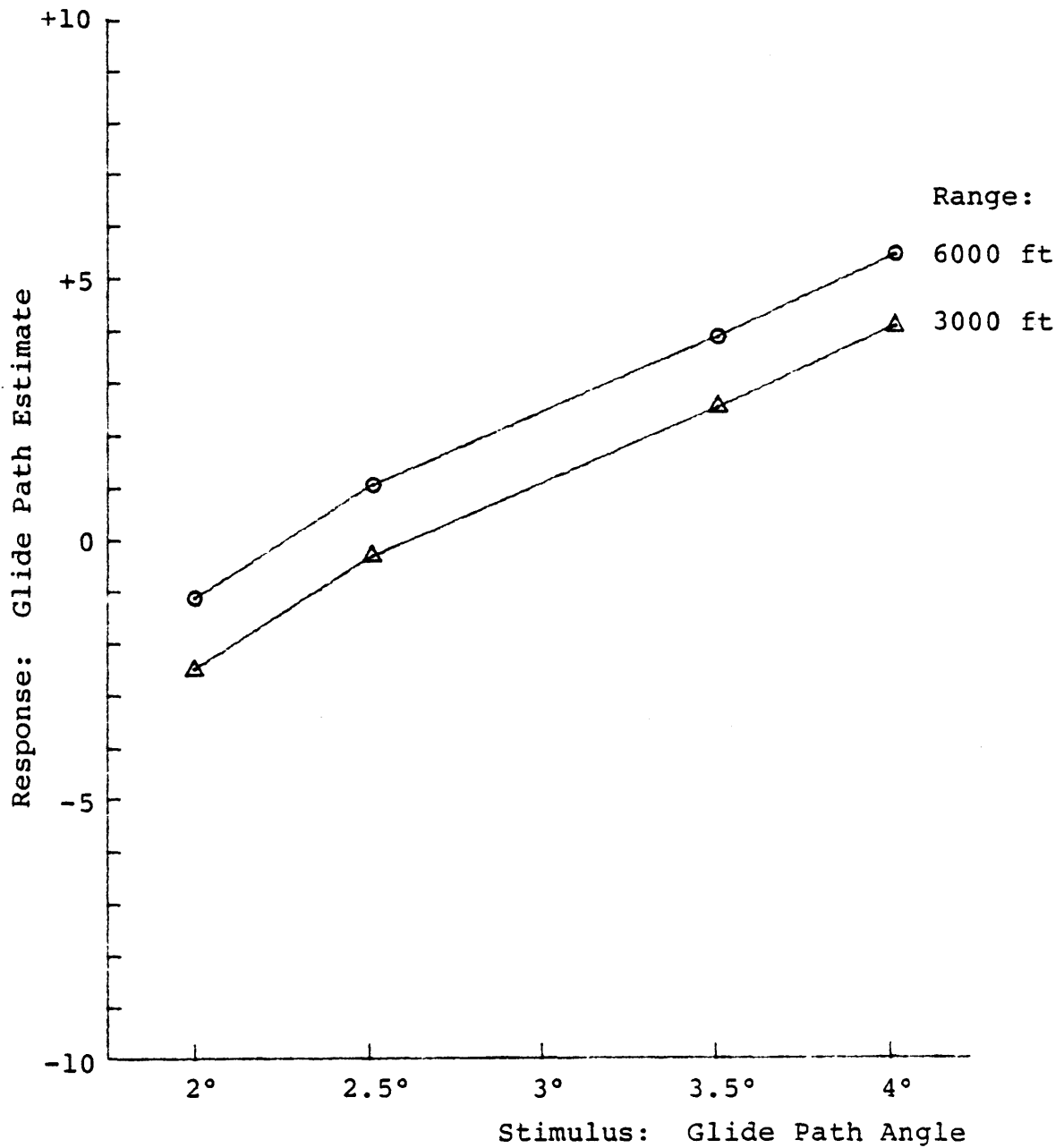


Figure C.1 Glide path estimate model for subject #1

Significant effects: glide path and range ($p < .005$)

Mean = 1.67

Model RMS error = 1.62

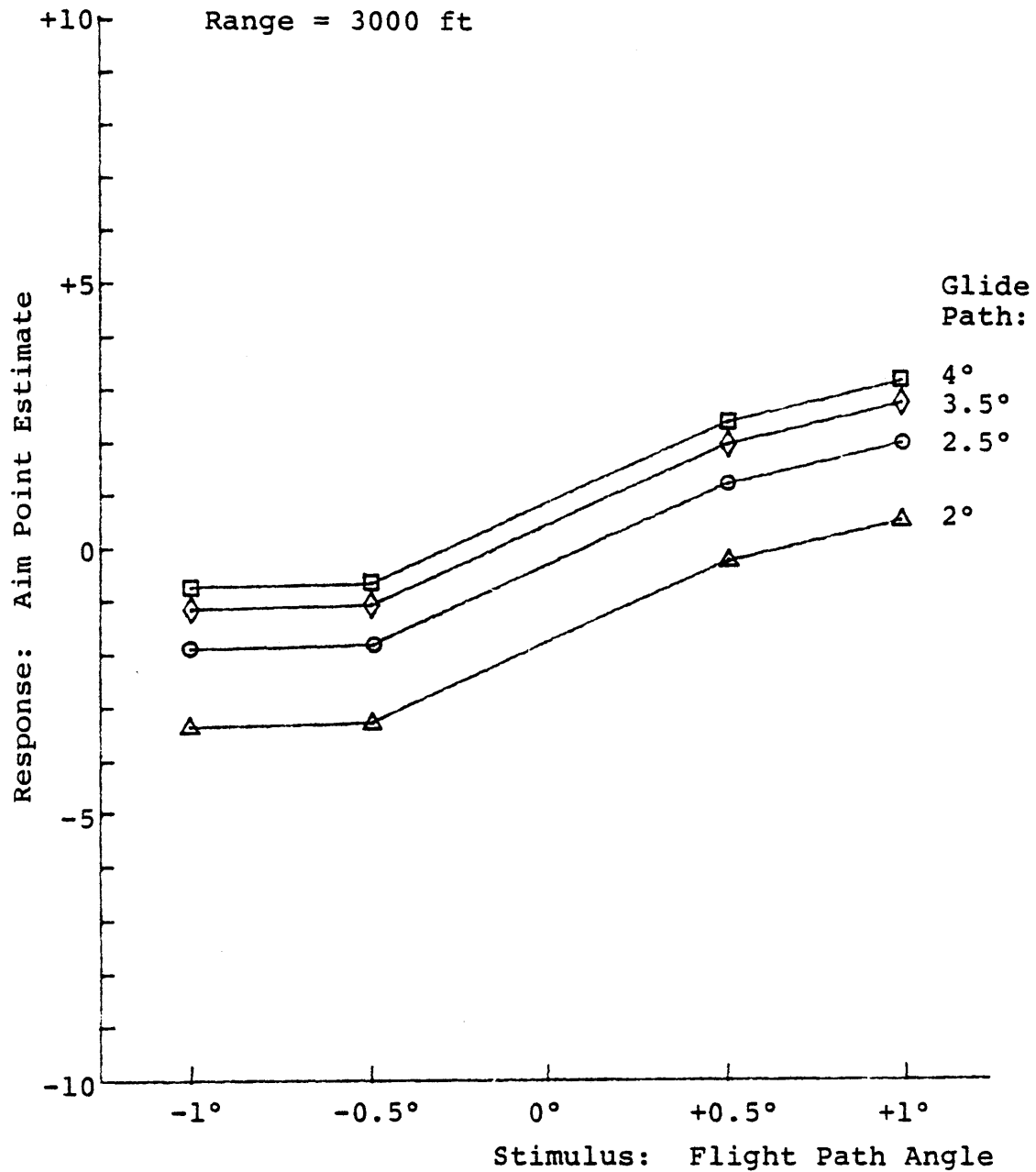


Figure C.2a Aim point estimate model for subject #1 (at 3000 ft range).

Significant effects: aim point and glide path ($p < .005$)
range ($p < .05$)

Mean = 0.45

Model RMS error = 1.97

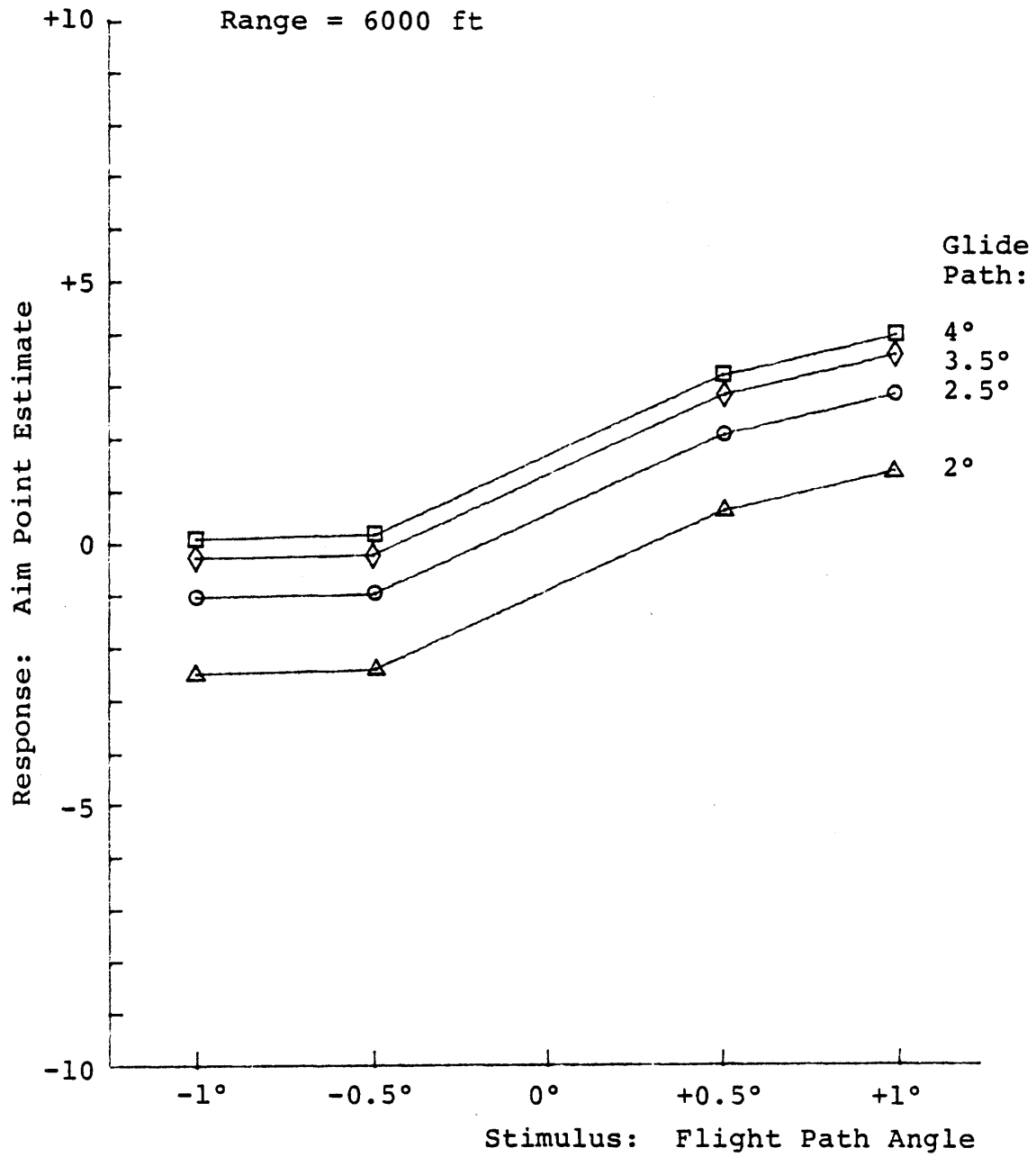


Figure C.2b Aim point estimate model for subject #1 (at 6000 ft range).

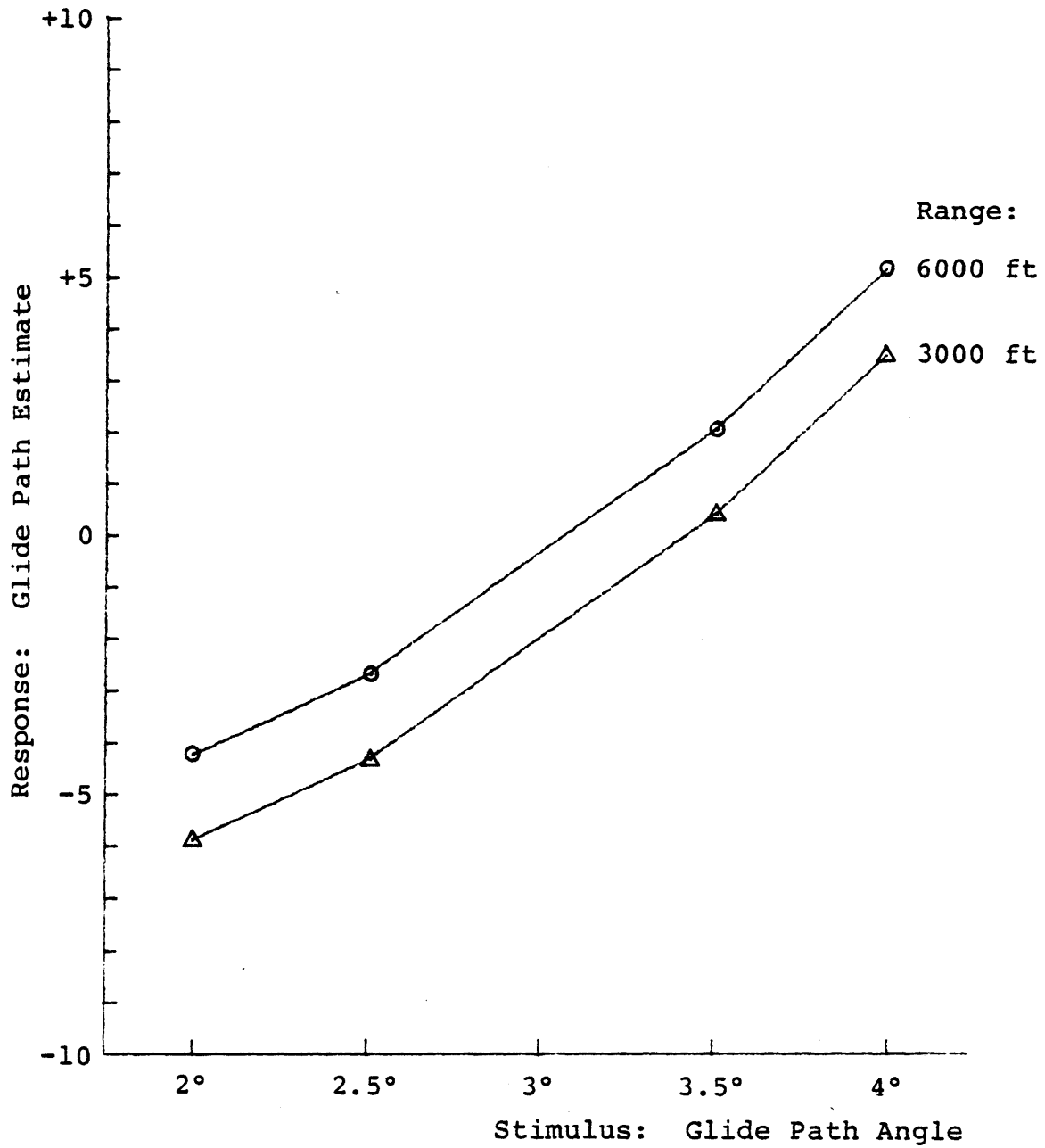


Figure C.3 Glide path estimate model for subject #2

Significant effects: glide path and range ($p < .005$)

Mean = -0.72

Model RMS error = 1.67

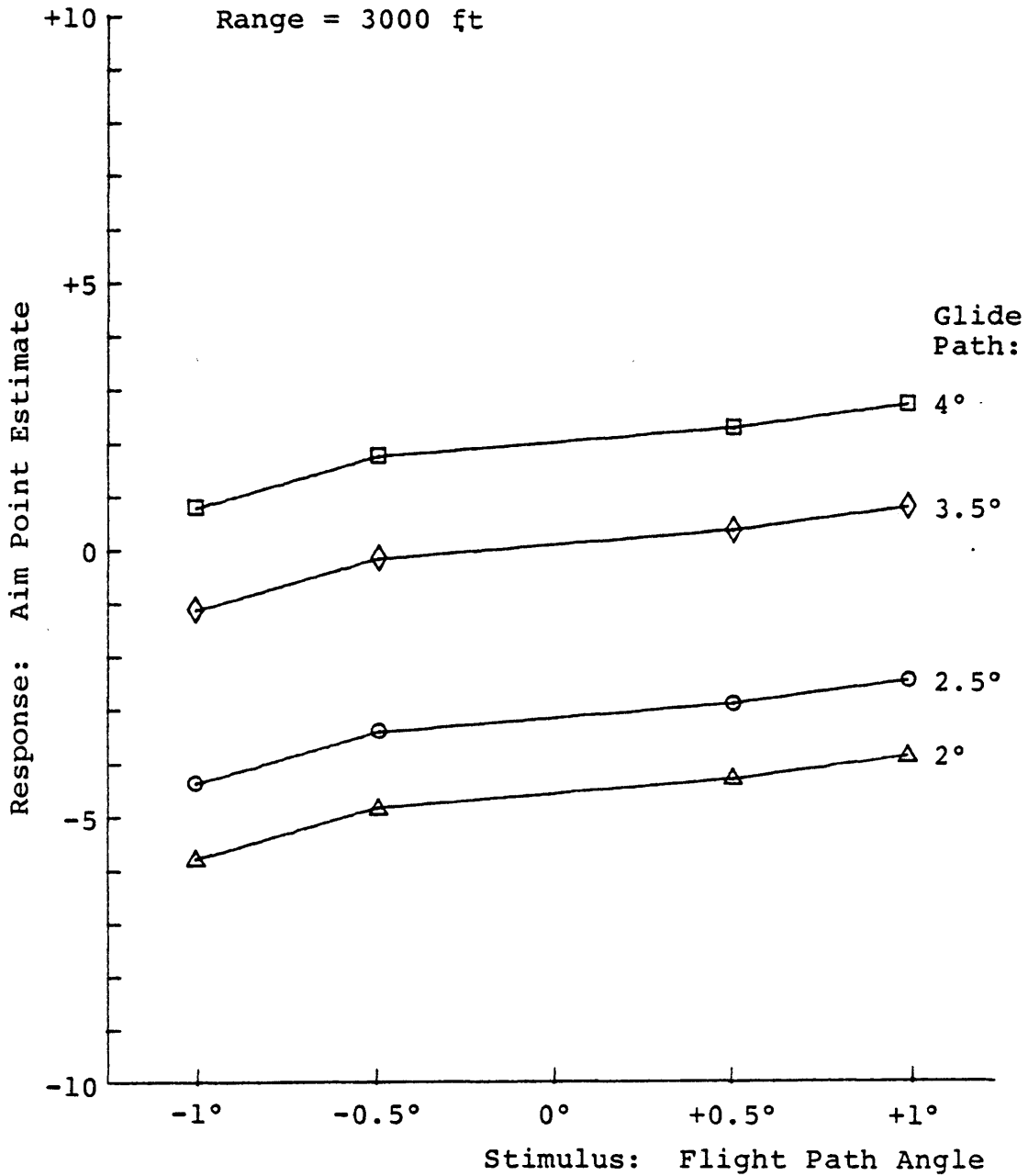


Figure C.4a Aim point estimate model for subject #2 (at 3000 ft range).

Significant effects: aim point, range, glide path, and range/glide path interaction ($p < .005$)

Mean = -0.73

Model RMS error = 1.91

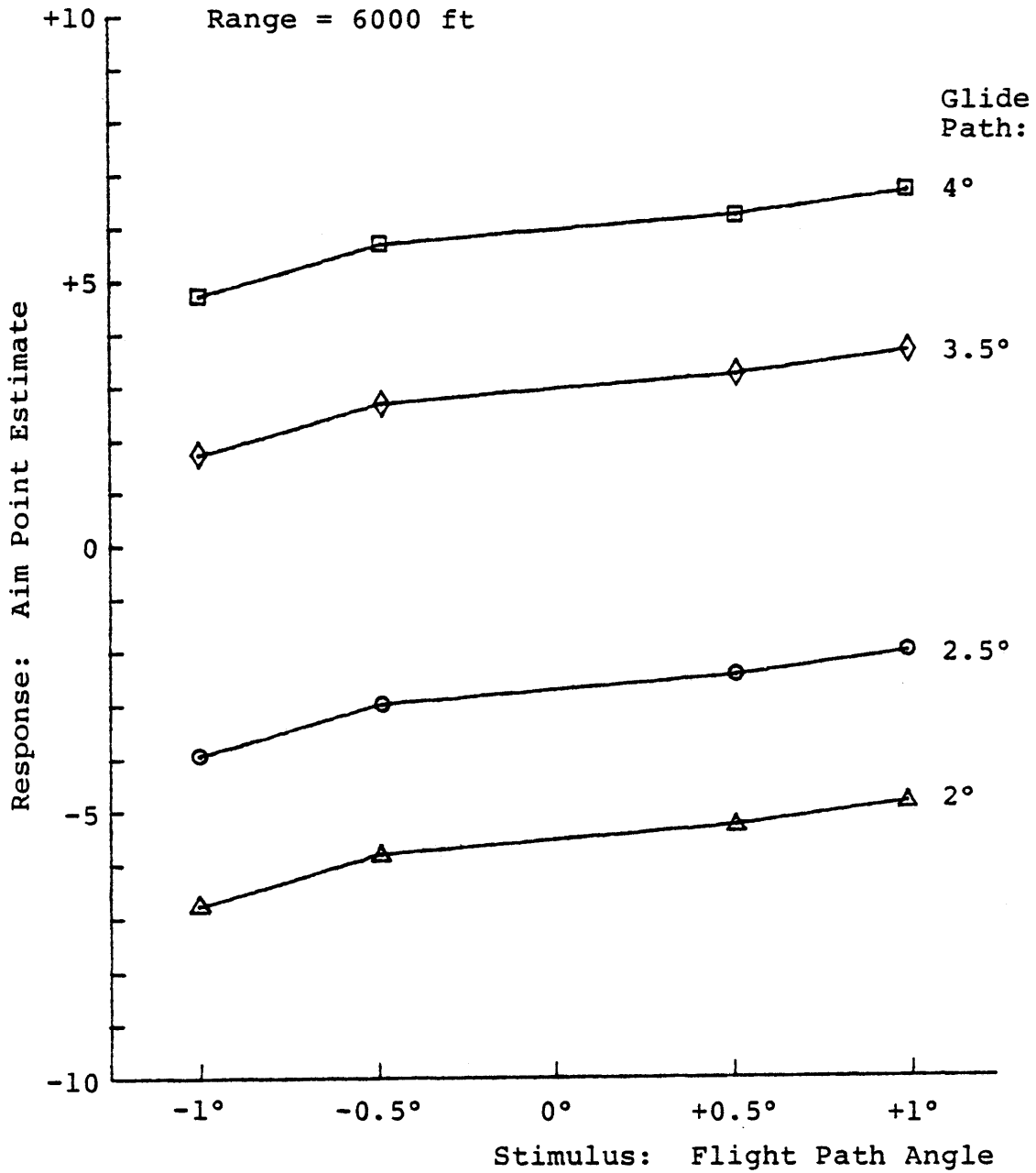


Figure C.4b Aim point estimate model for subject #2 (at 6000 ft range).

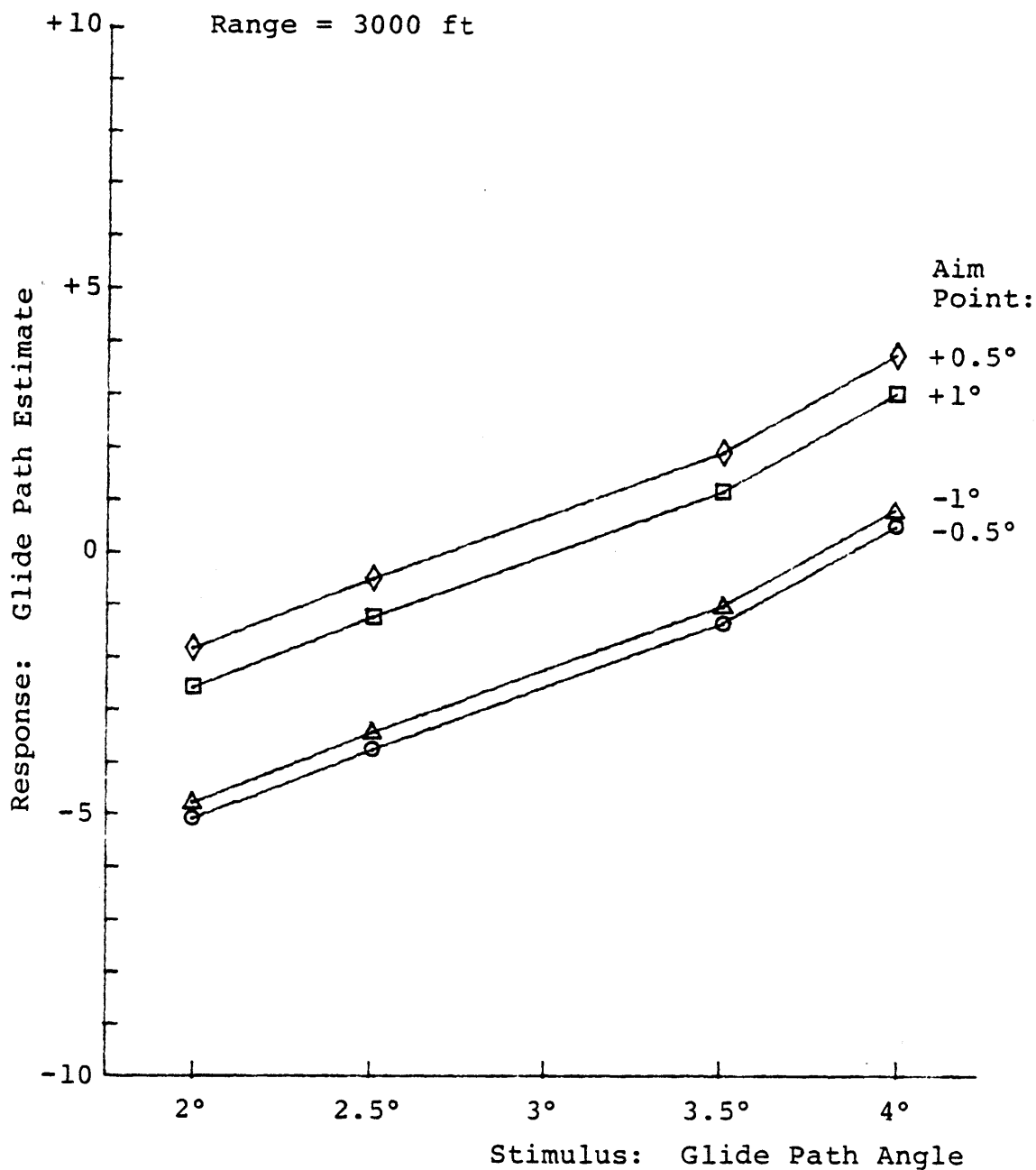


Figure C.5a Glide path estimate model for subject #3 (at 3000 ft range).

Significant effects: glide path and range ($p < .005$)
 range/glide path interaction ($p < .01$)
 range/aim point interaction ($p < .025$)

Mean = 1.09

Model RMS error = 2.63

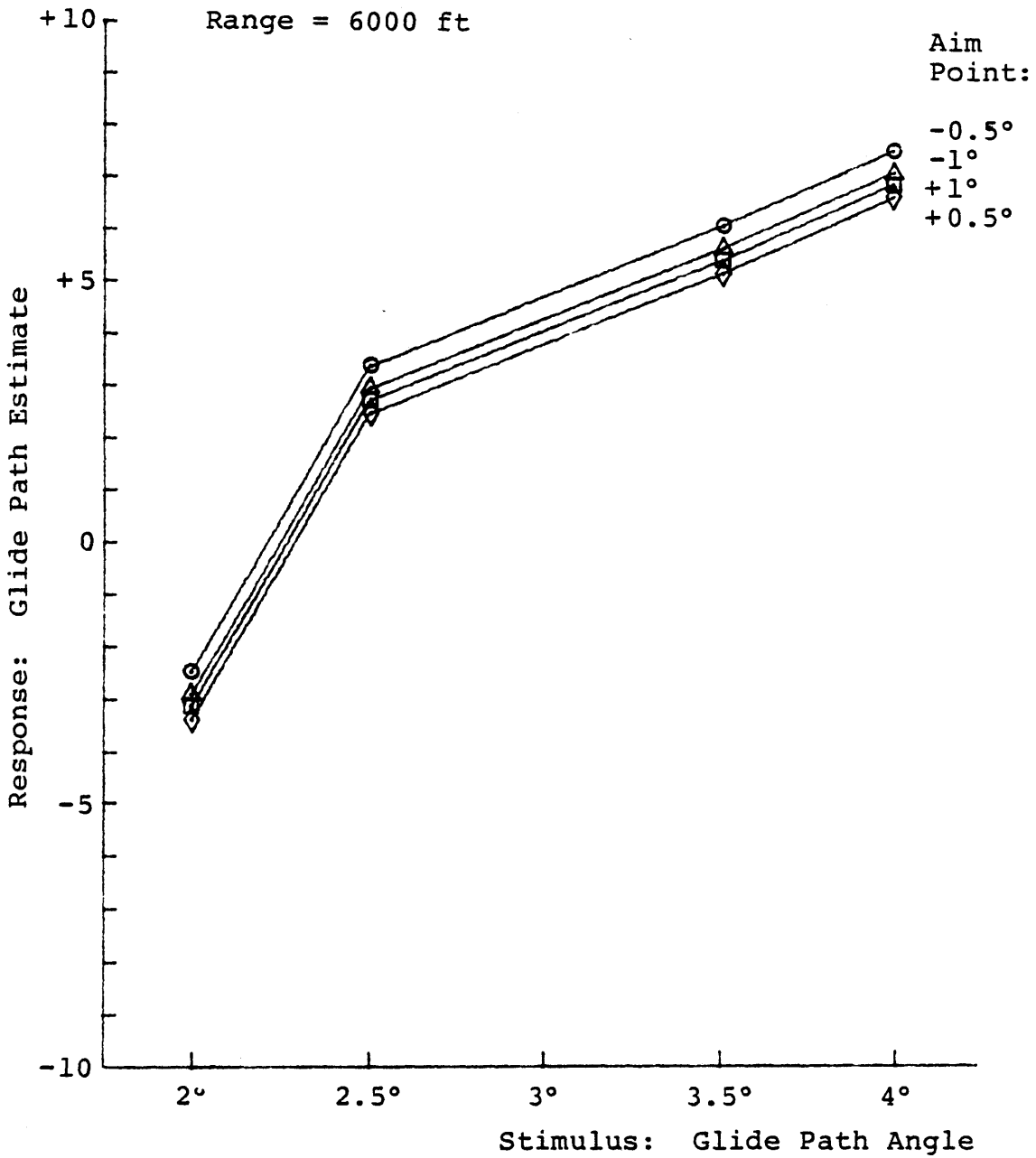


Figure C.5b Glide path estimate model for subject #3 (at 6000 ft range).

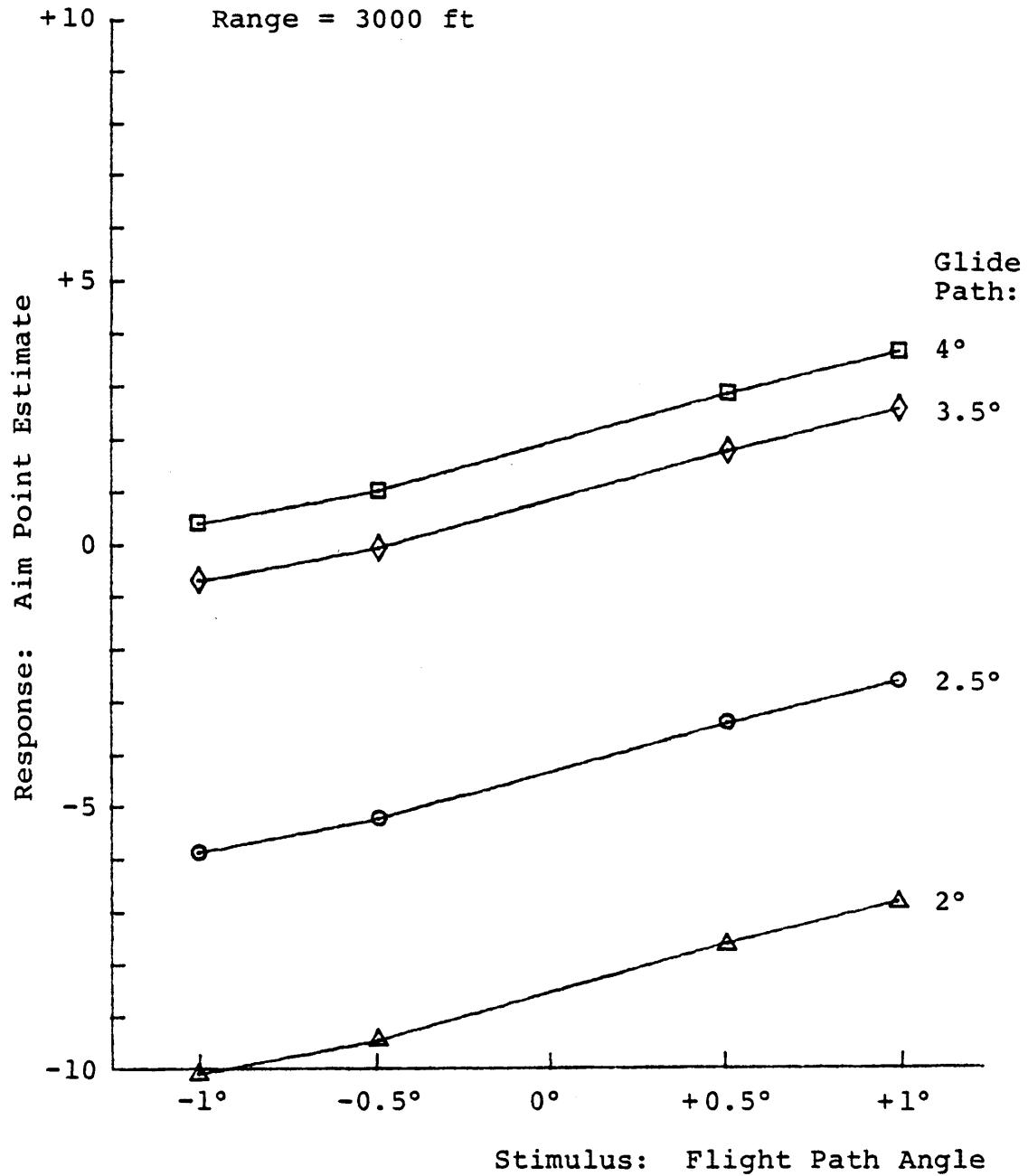


Figure C.6a Aim point estimate model for subject #3 (at 3000 ft range).

Significant effects: aim point ($p < .01$)

glide path and range ($p < .005$)

(Mean not significant)

Model RMS error = 3.74

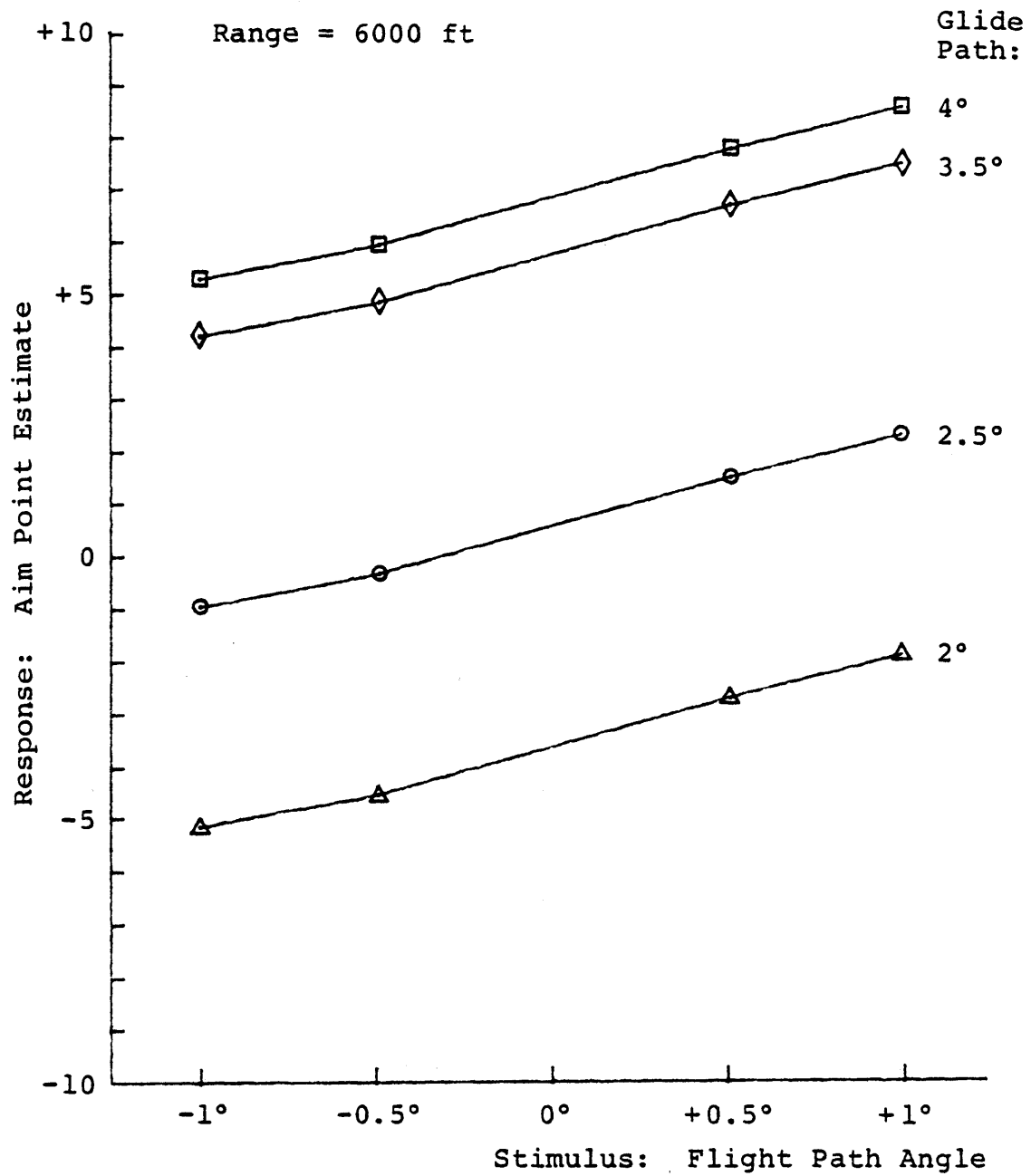


Figure C.6b Aim point estimate model for subject #3 (at 6000 ft range).

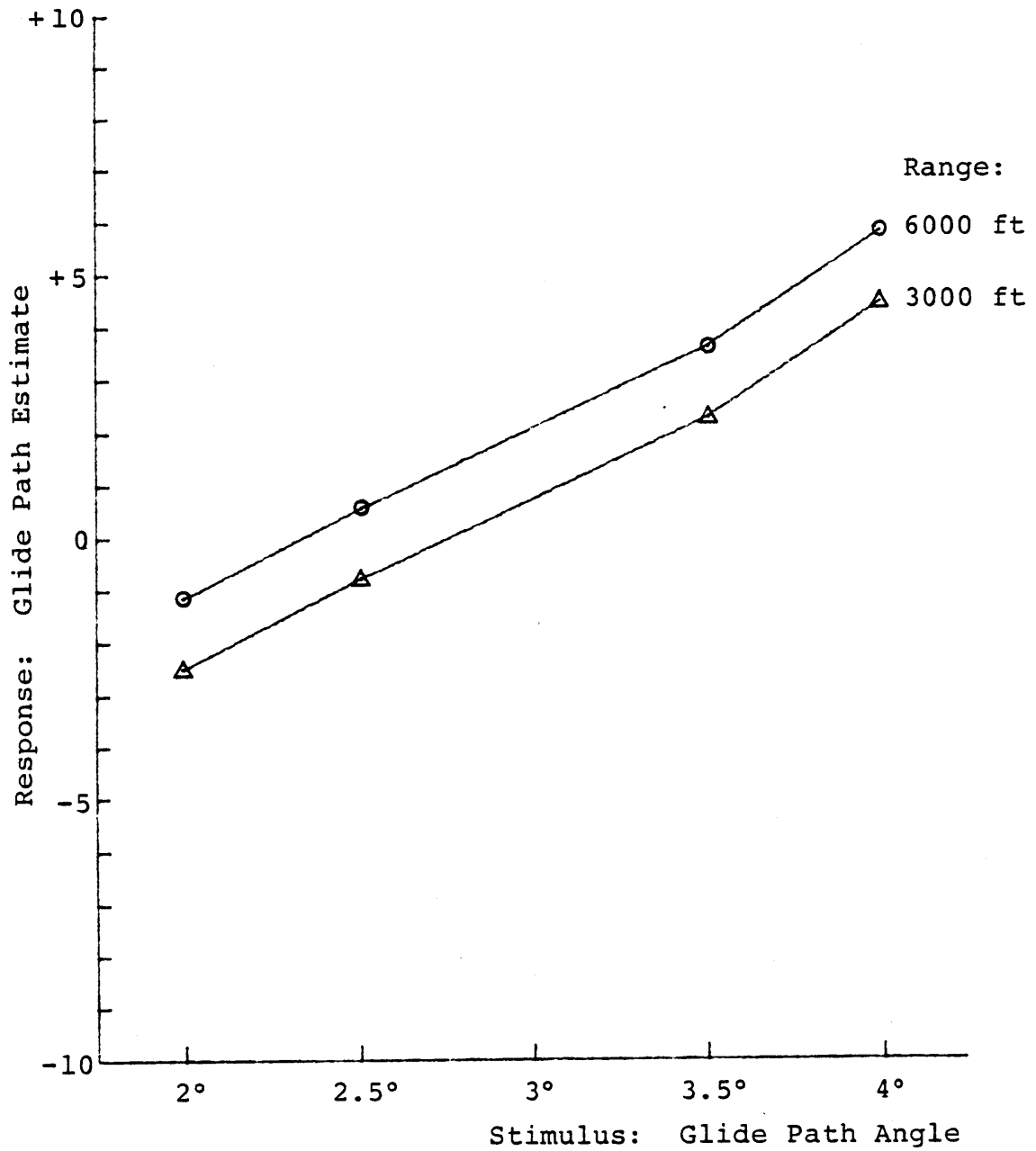


Figure C.7 Glide path estimate model for subject #4

Significant effects: glide path and range ($p < .005$)

Mean = 1.55

Model RMS error = 1.76

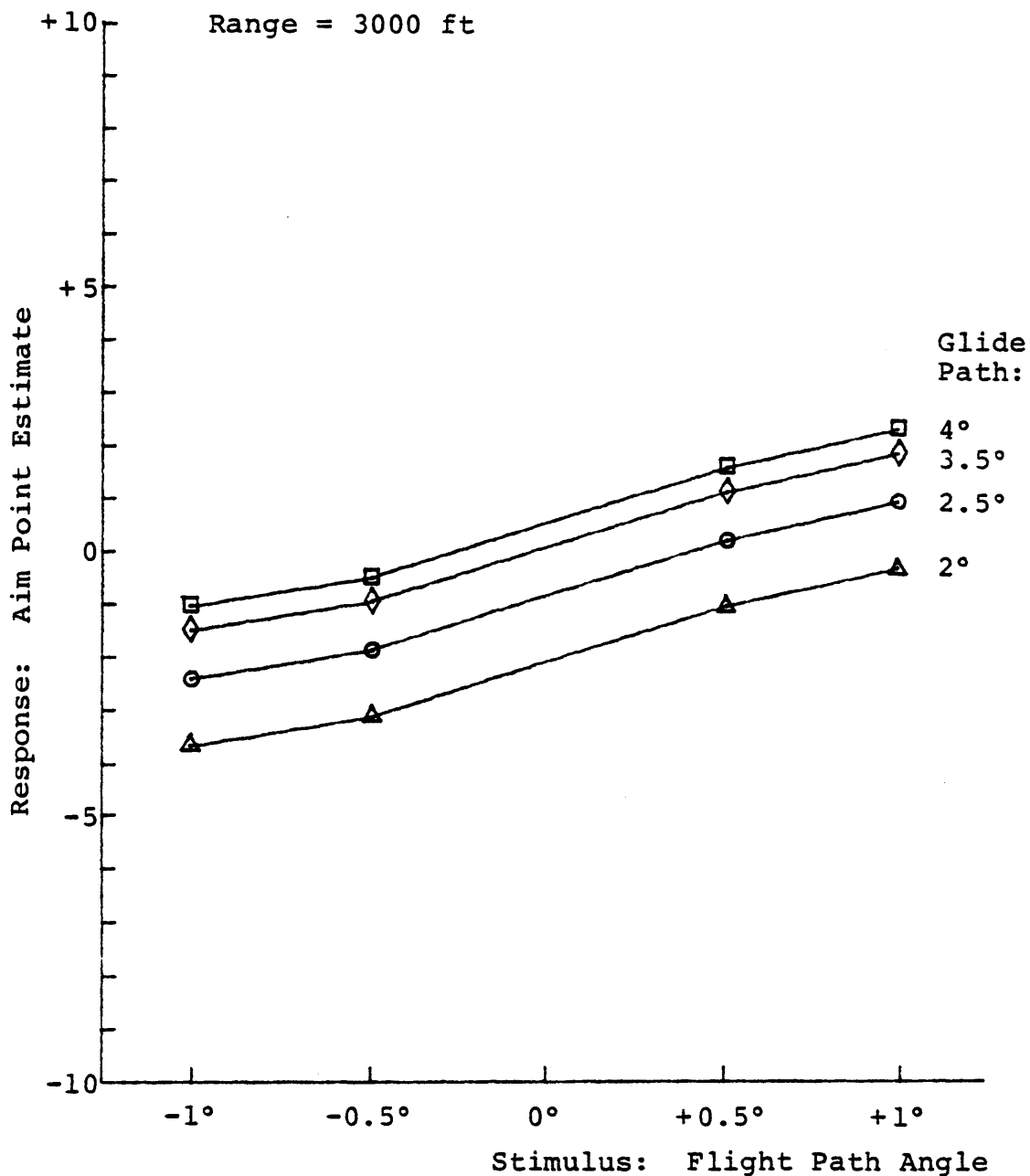


Figure C.8a Aim point estimate model for subject #4 (at 3000 ft range).

Significant effects: aim point, range, and glide path
($p < .005$)

Mean = 0.34

Model RMS error = 1.60

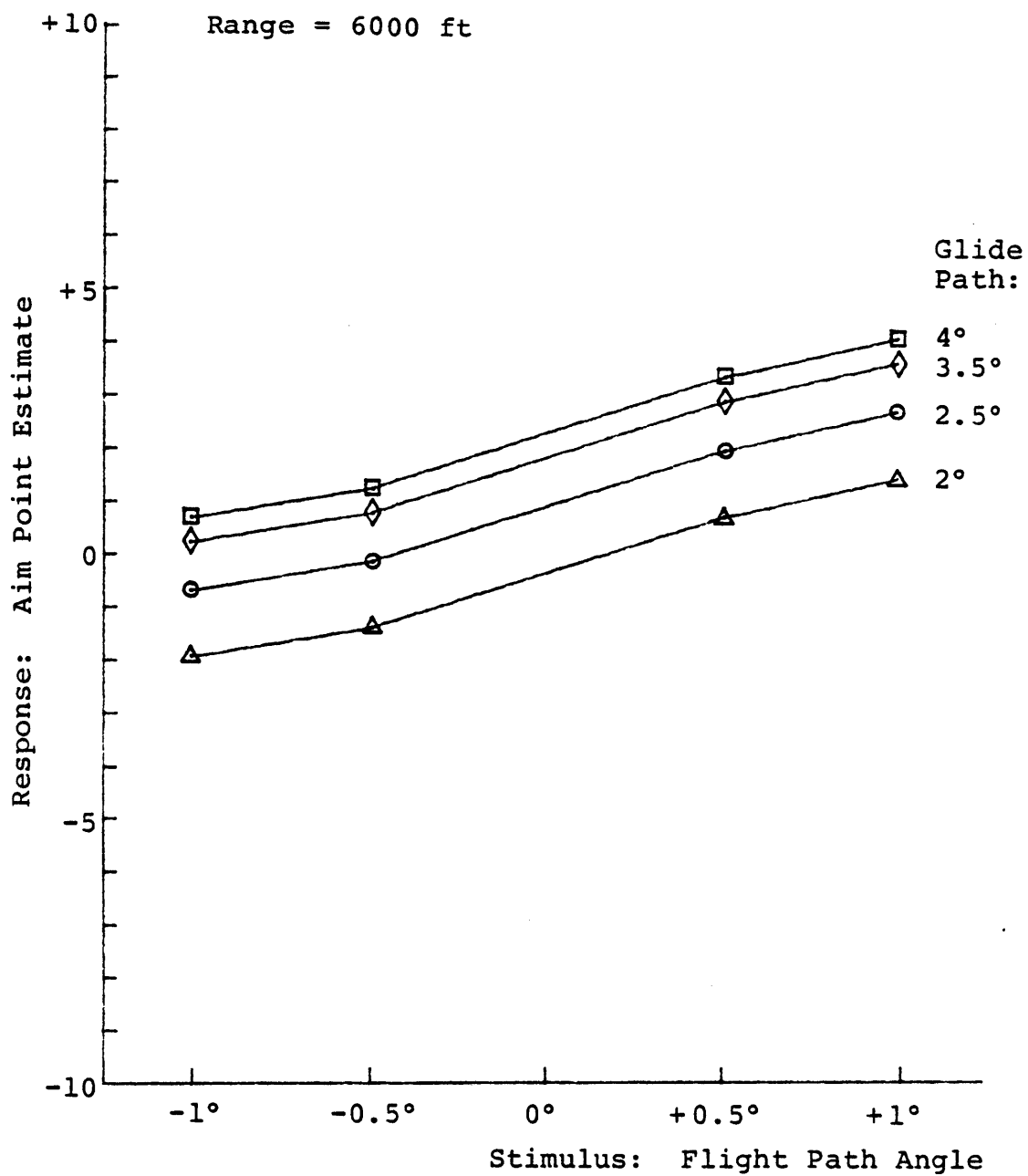


Figure C.8b Aim point estimate model for subject #4 (at 6000 ft range).

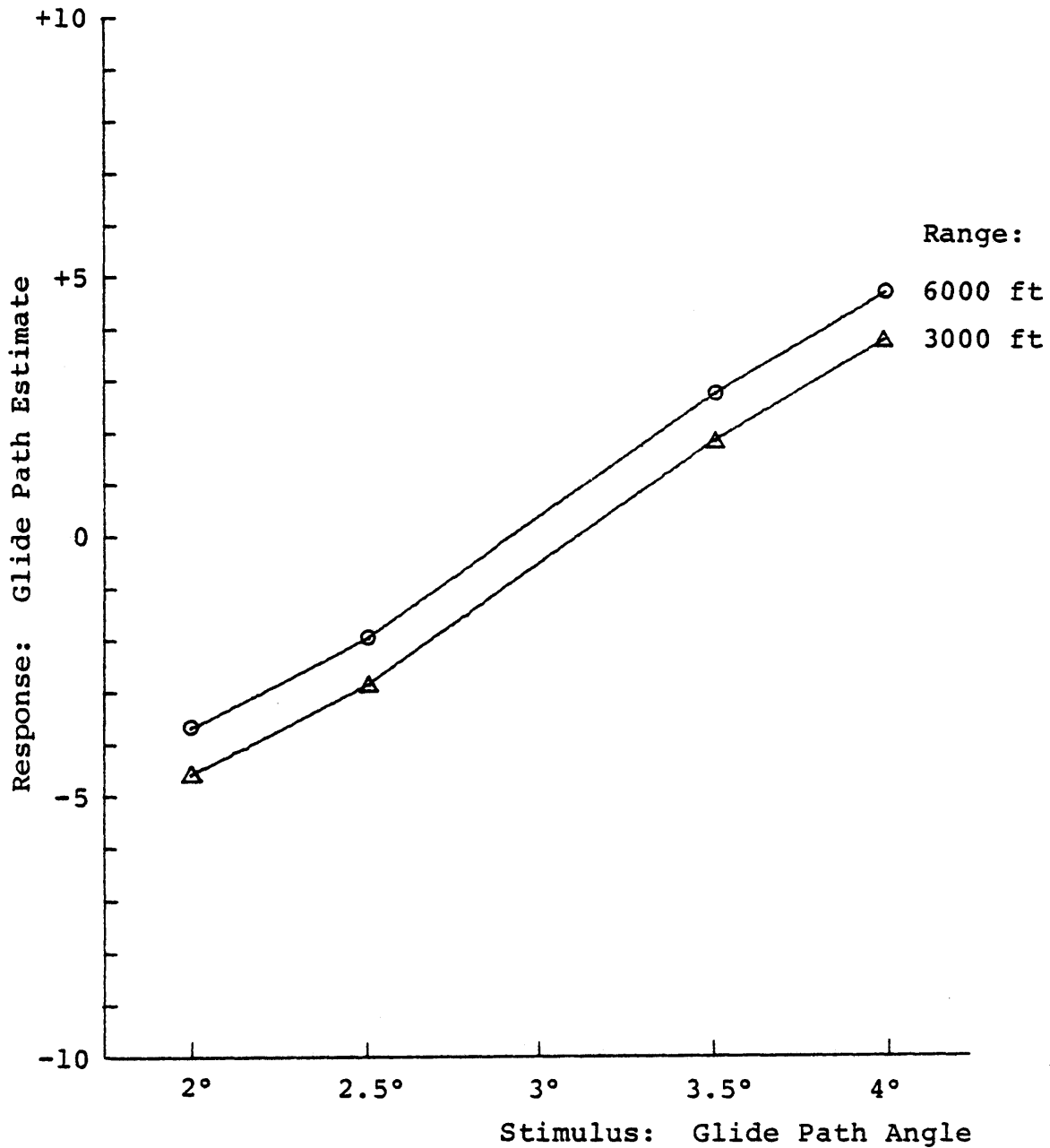


Figure C.9 Glide path estimate model for subject #5

Significant effects: glide path ($p < .005$)

range ($p < .05$)

(Mean not significant)

Model RMS error = 2.16

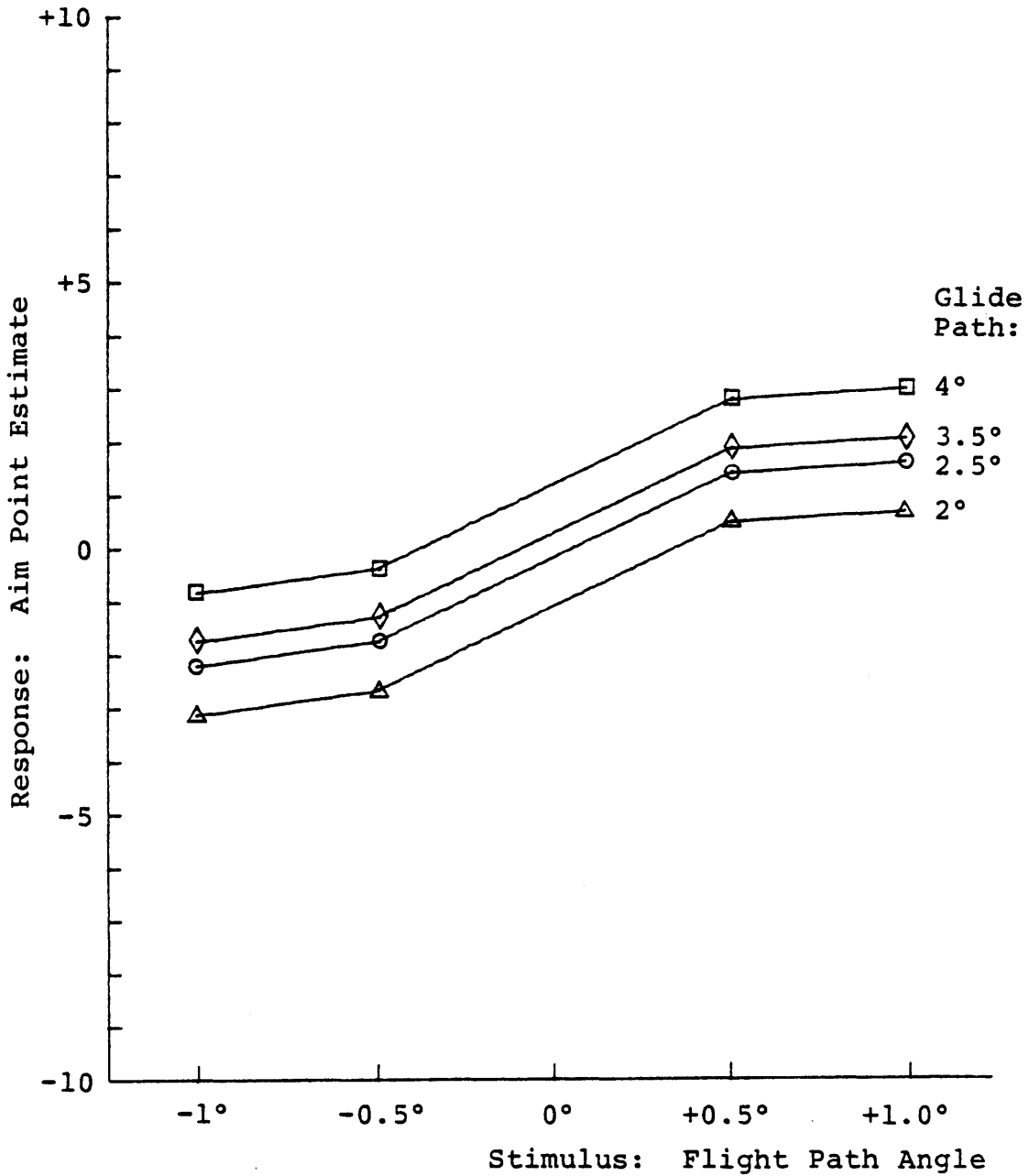


Figure C.10 Aim point estimate model for subject #5

Significant effects: aim point ($p < .005$)

glide path ($p < .025$)

(Mean not significant)

Model RMS error = 2.40

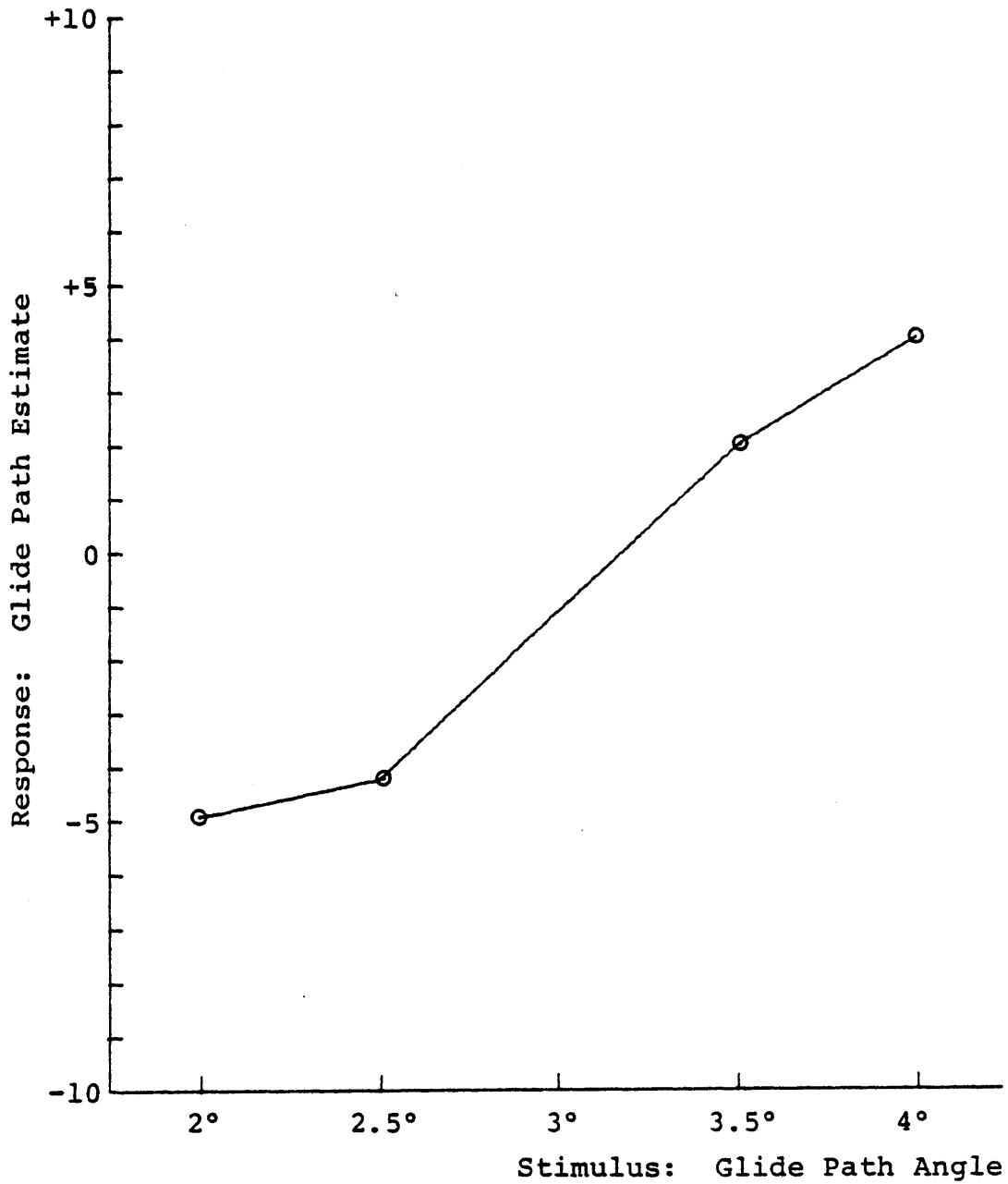


Figure C.11 Glide path estimate model for subject #6

Significant effect: glide path ($p < .005$)

Mean = -0.75

Model RMS error = 2.71

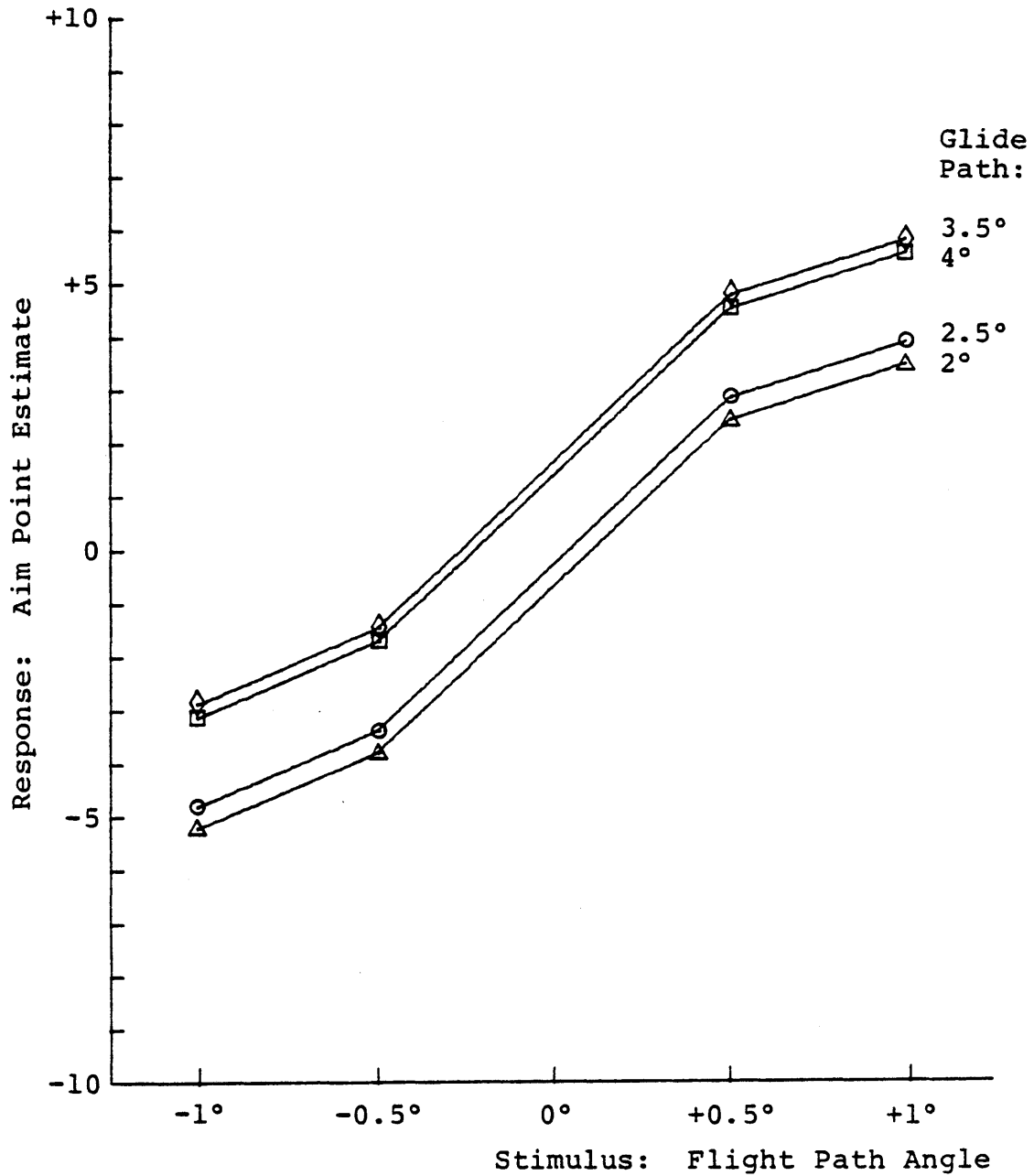


Figure C.12a Aim point estimate model for subject #6 (at 3000 ft range).

Significant effects: aim point, range, and glide path ($p < .005$)
 range/aim point interaction ($p < .025$)
 range/glide path interaction ($p < .05$)

Mean = 1.31

Model RMS error = 2.62

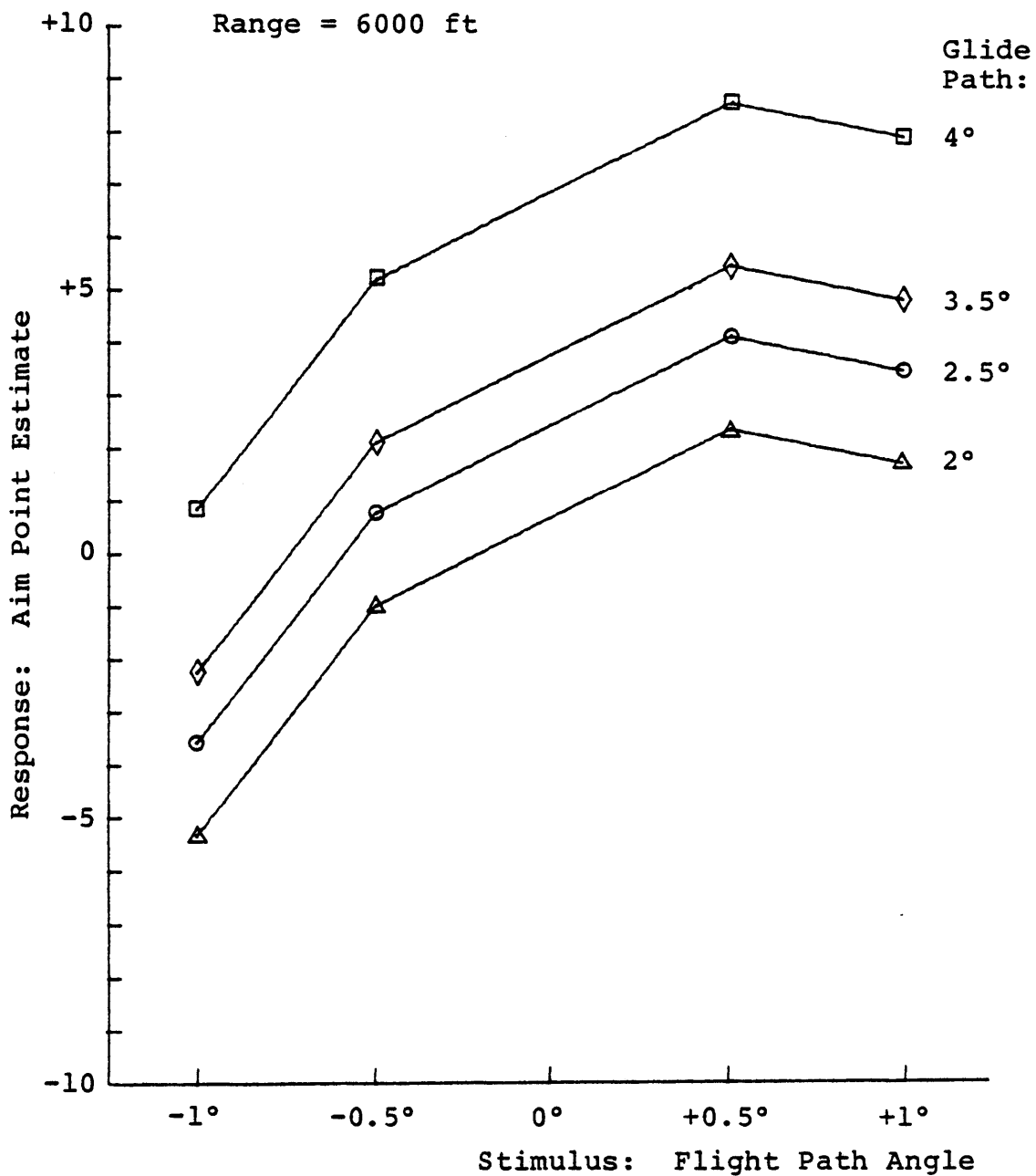


Figure C.12b Aim point estimate model for subject #6 (at 6000 ft range).

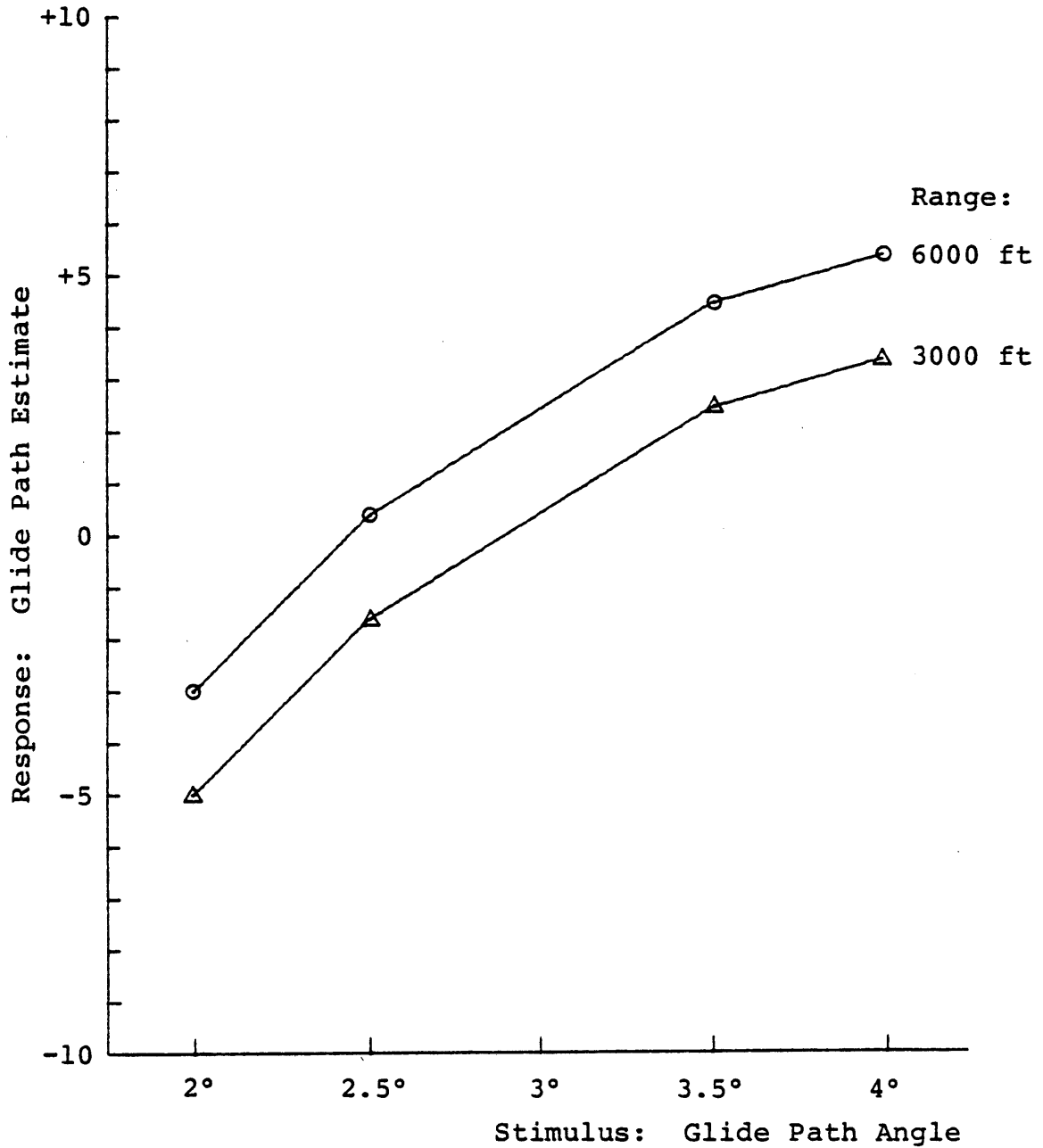


Figure C.13 Glide path estimate model for subject #7

Significant effects: glide path and range ($p < .005$)

Mean = 0.81

Model RMS error = 1.57

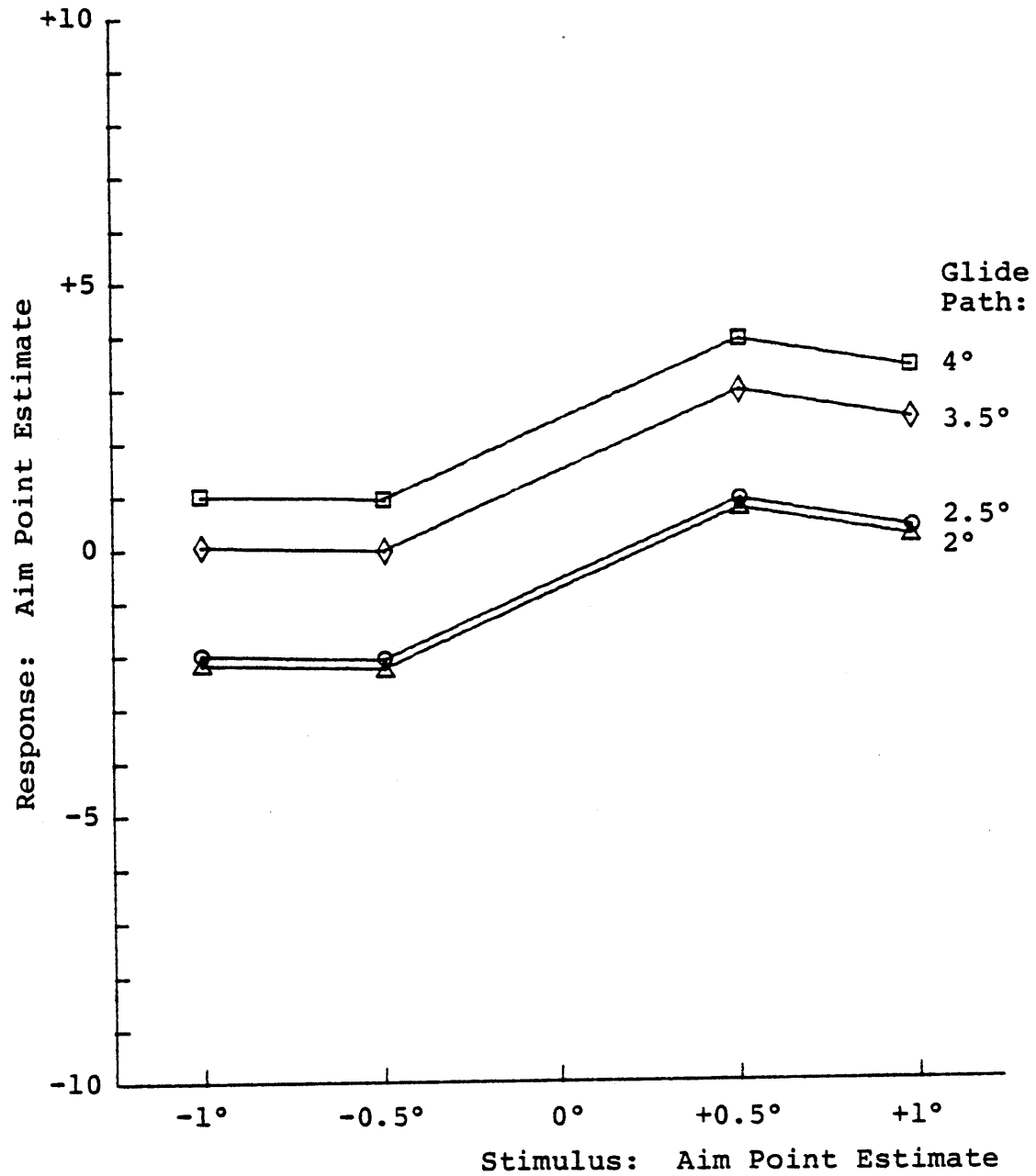


Figure C.14 Aim point estimate model for subject #7

Significant effects: aim point and glide path ($p < .005$)

Mean = 0.51

Model RMS error = 1.87

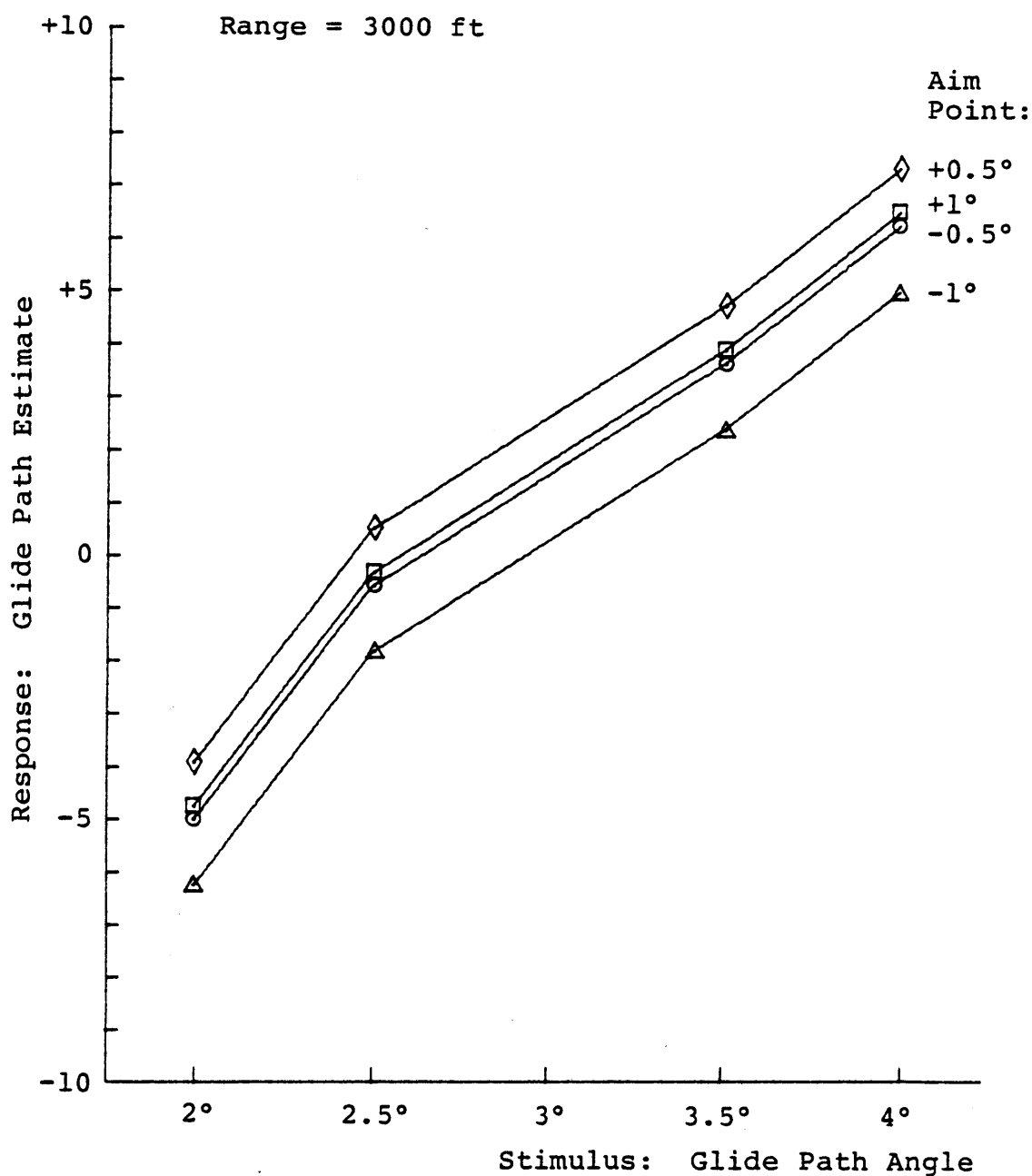


Figure C.15a Glide path estimate model for subject #8 (at 3000 ft range).

Significant effects: glide path and range ($p < .005$)
range/glide path interaction ($p < .05$)

Mean = 1.92

Model RMS error = 2.31

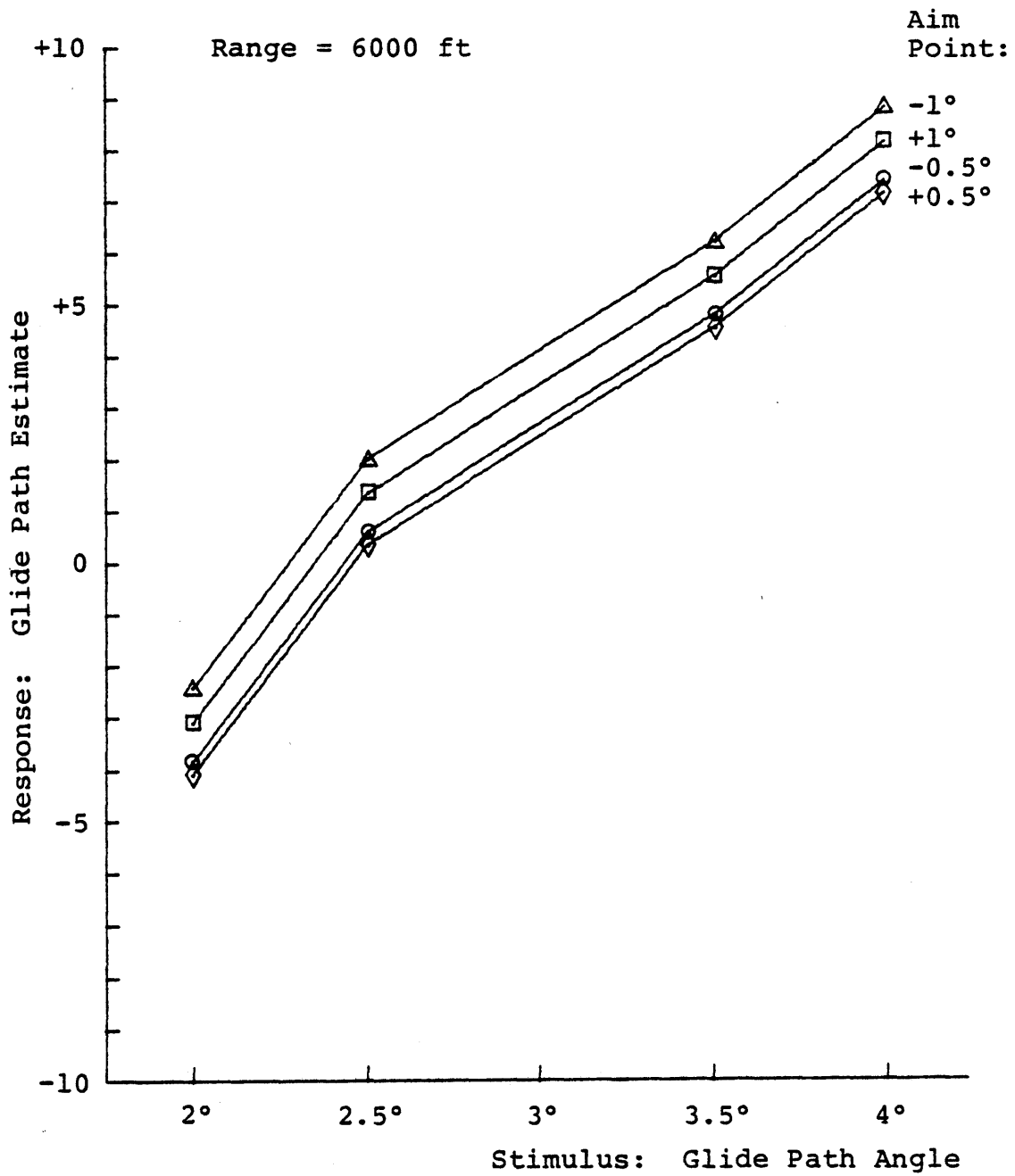


Figure C.15b Glide path estimate model for subject #8 (at 6000 ft range).

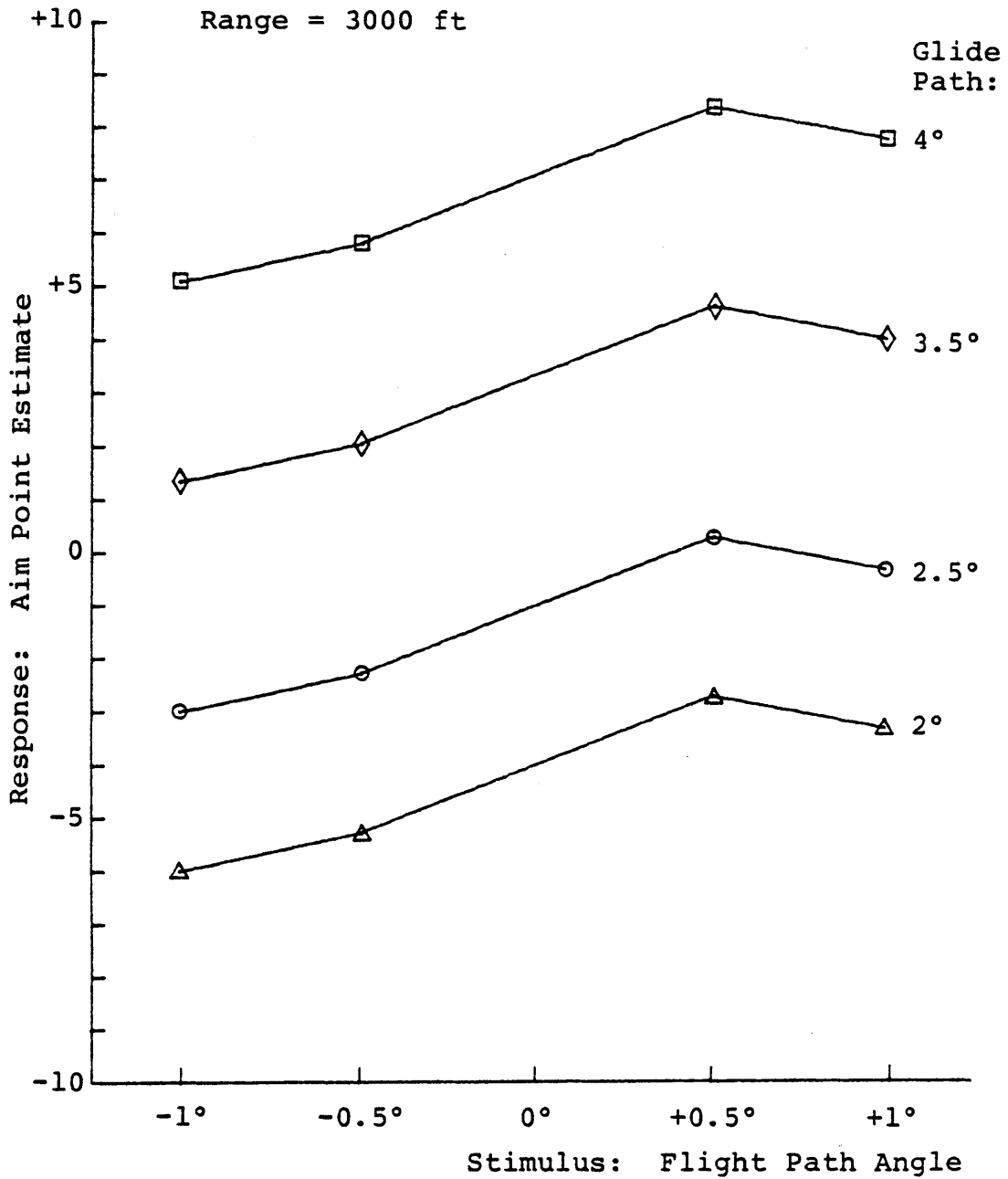


Figure C.16a Aim point estimate model for subject #8 (at 3000 ft range).

Significant effects: aim point and glide path ($p < .005$)
 range/glide path interaction ($p < .05$)

Mean = 1.21

Model RMS error = 3.28

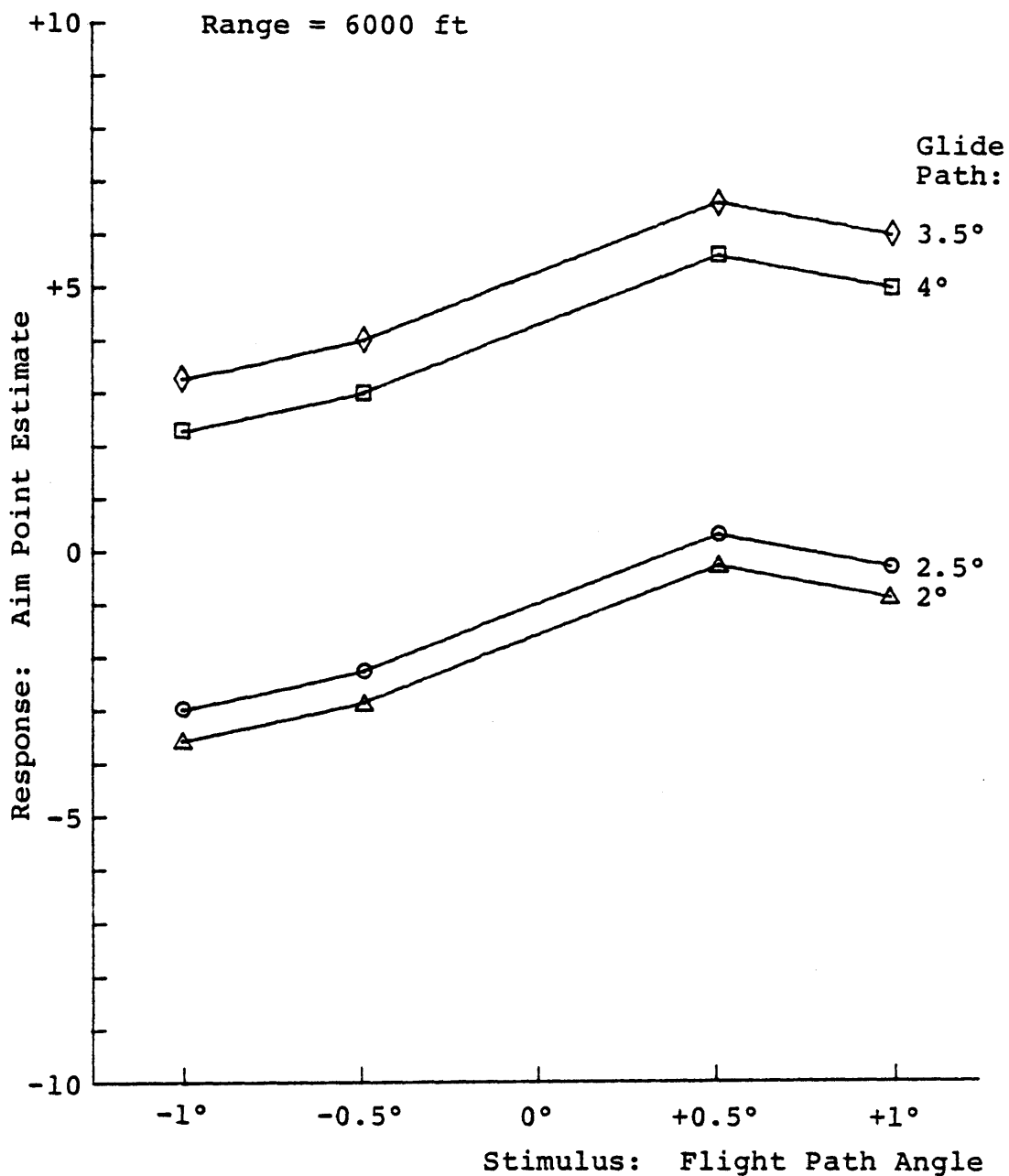


Figure C.16b Aim point estimate model for subject #8 (at 6000 ft range).

C.2 Raw Data Statistics

As a check on the reasonableness of the statistical models, means and standard deviations of the estimates at significant combinations of stimuli were calculated for the two groups of combined subjects. These are plotted in Figures C.17 through C.20. Although the ± 1 standard deviation error bars are large, the biases and trends are clear and consistent, supporting the validity of the statistical model results.

Statistical models based on the same data are plotted in Section 4.1, Results for Grouped Data. Figures 4.1 and 4.2, the glide path estimate models, correspond respectively to Figures C.17 and C.18, the glide path raw data statistics for all subjects combined and high-time pilots. Similarly for aim point, Figures 4.3 and 4.4 correspond to Figures C.19 and C.20.

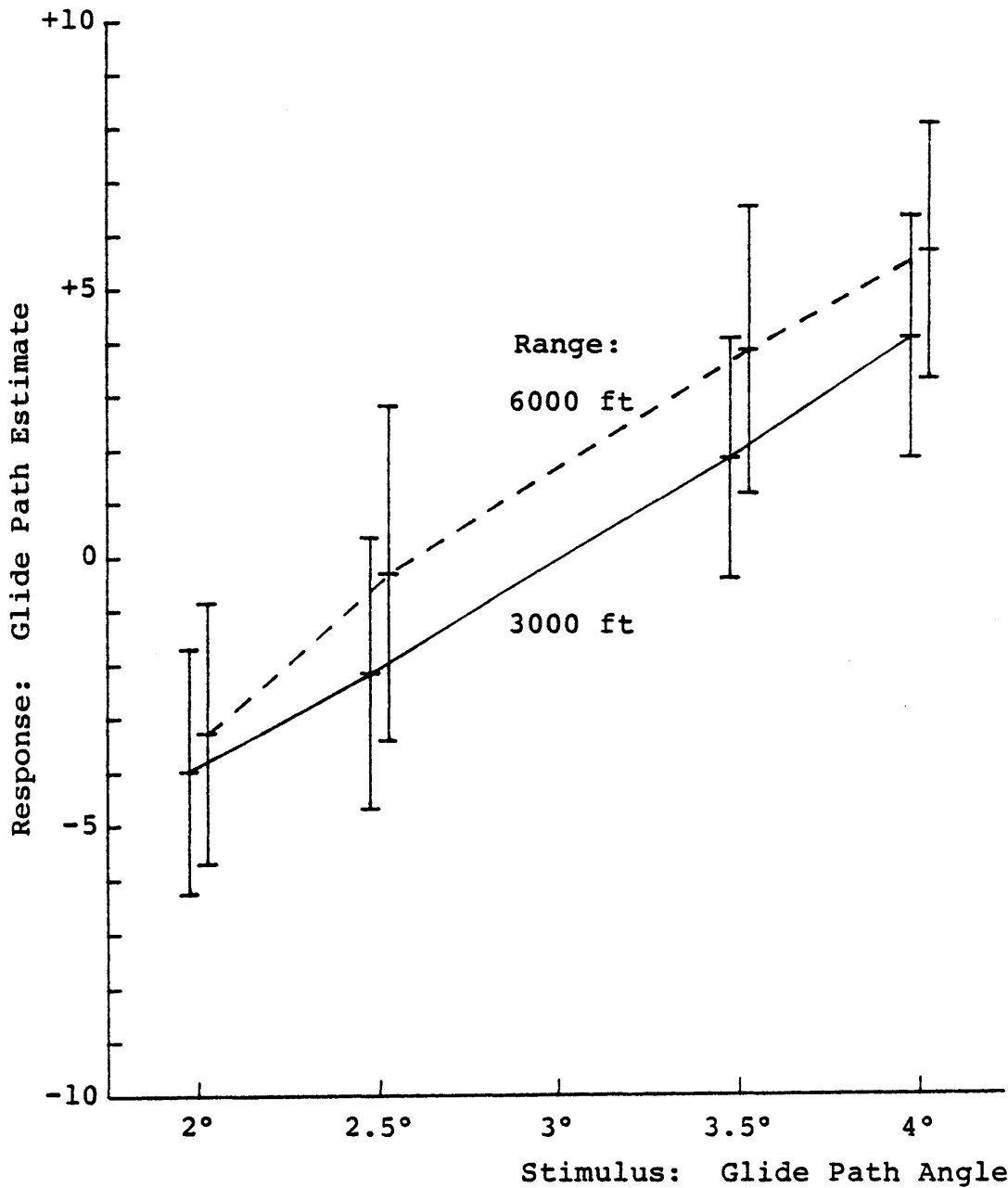


Figure C.17 Glide path data statistics for all subjects combined.

Means \pm 1 standard deviation are shown at each glide path and range.

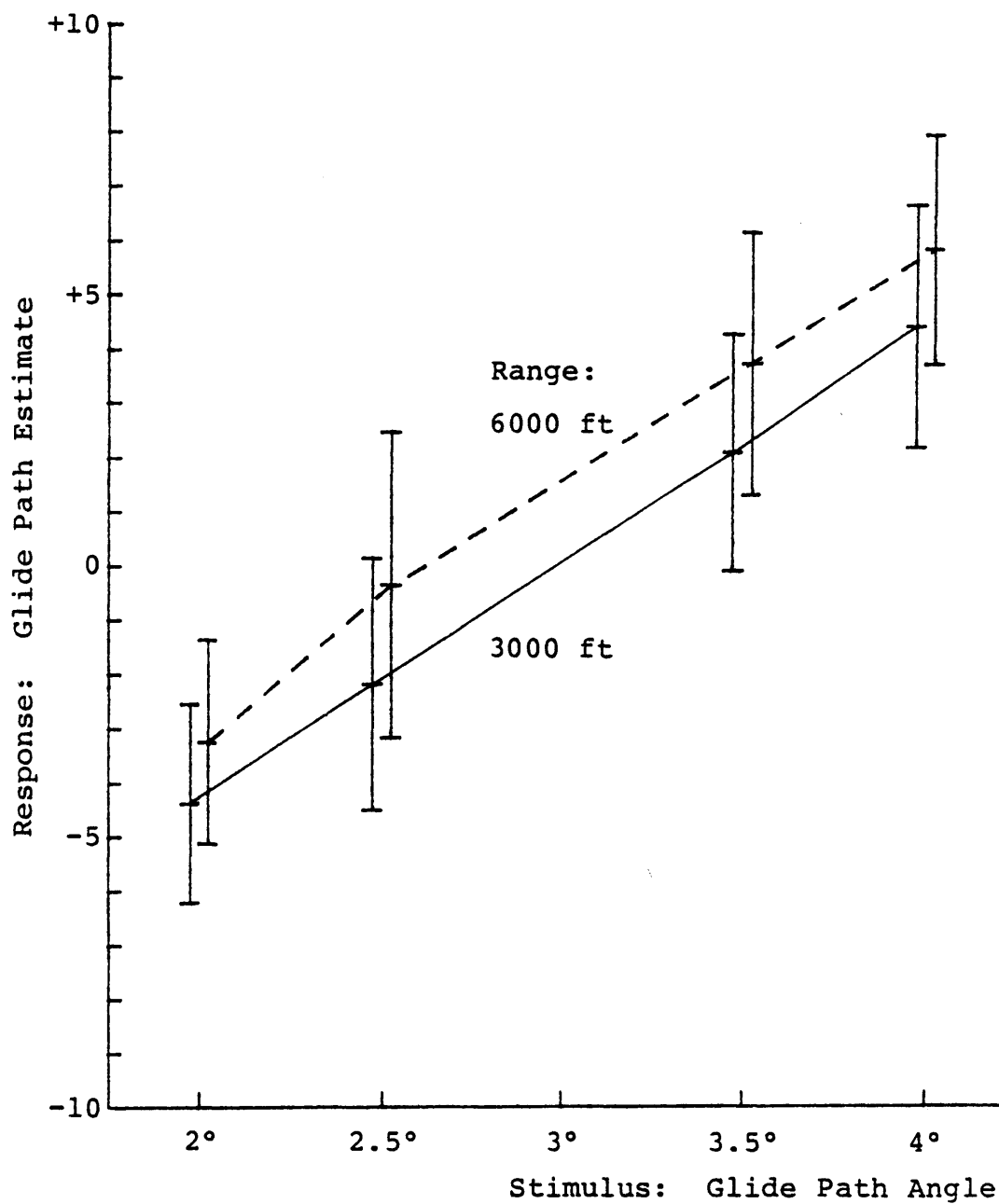


Figure C.18 Glide path data statistics for high-time pilots.

Means ± 1 standard deviation are shown at each glide path and range.

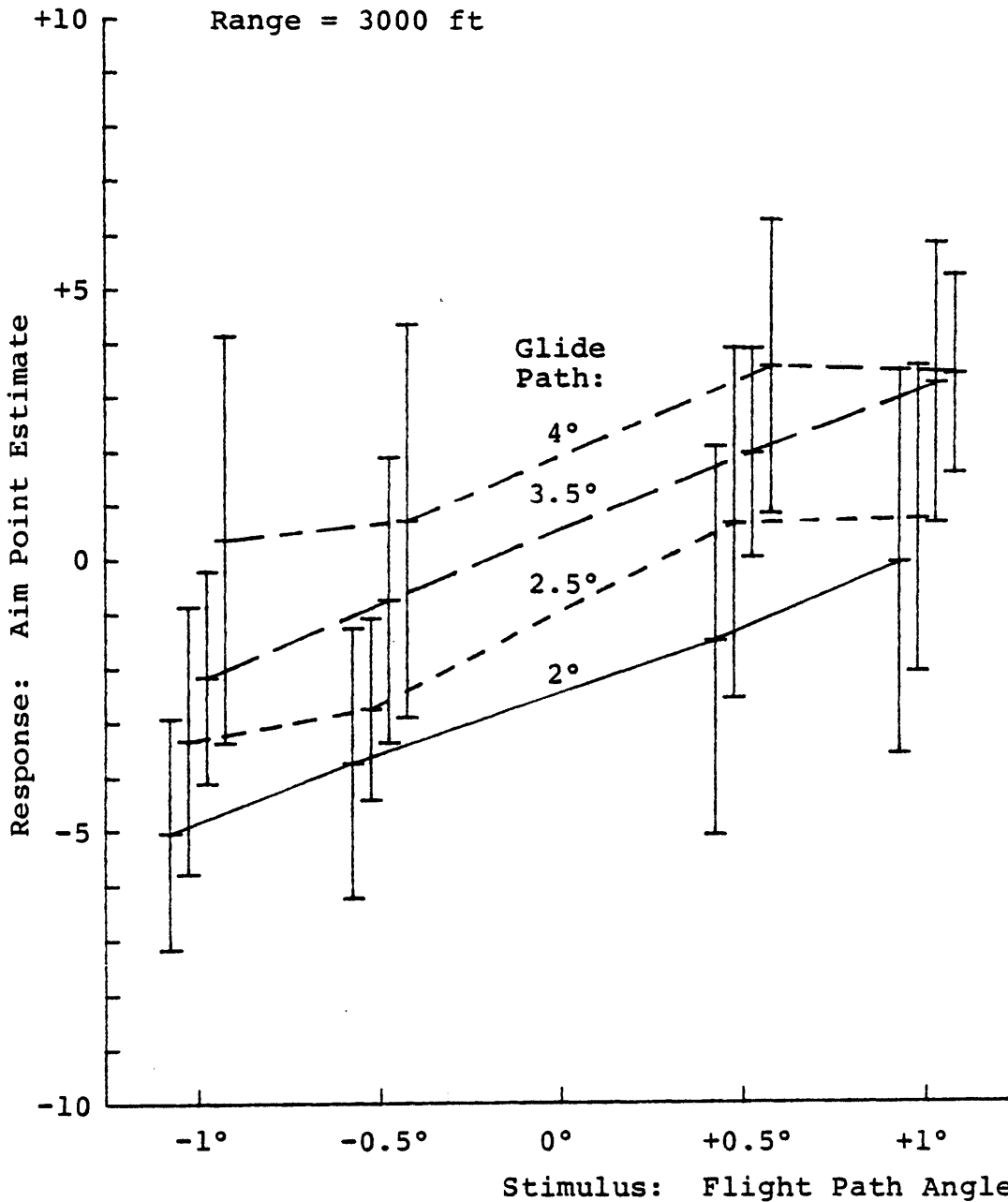


Figure C.19a Aim point data statistics for all subjects combined (at 3000 ft range).

Means \pm 1 standard deviation are shown at each aim point and glide path.

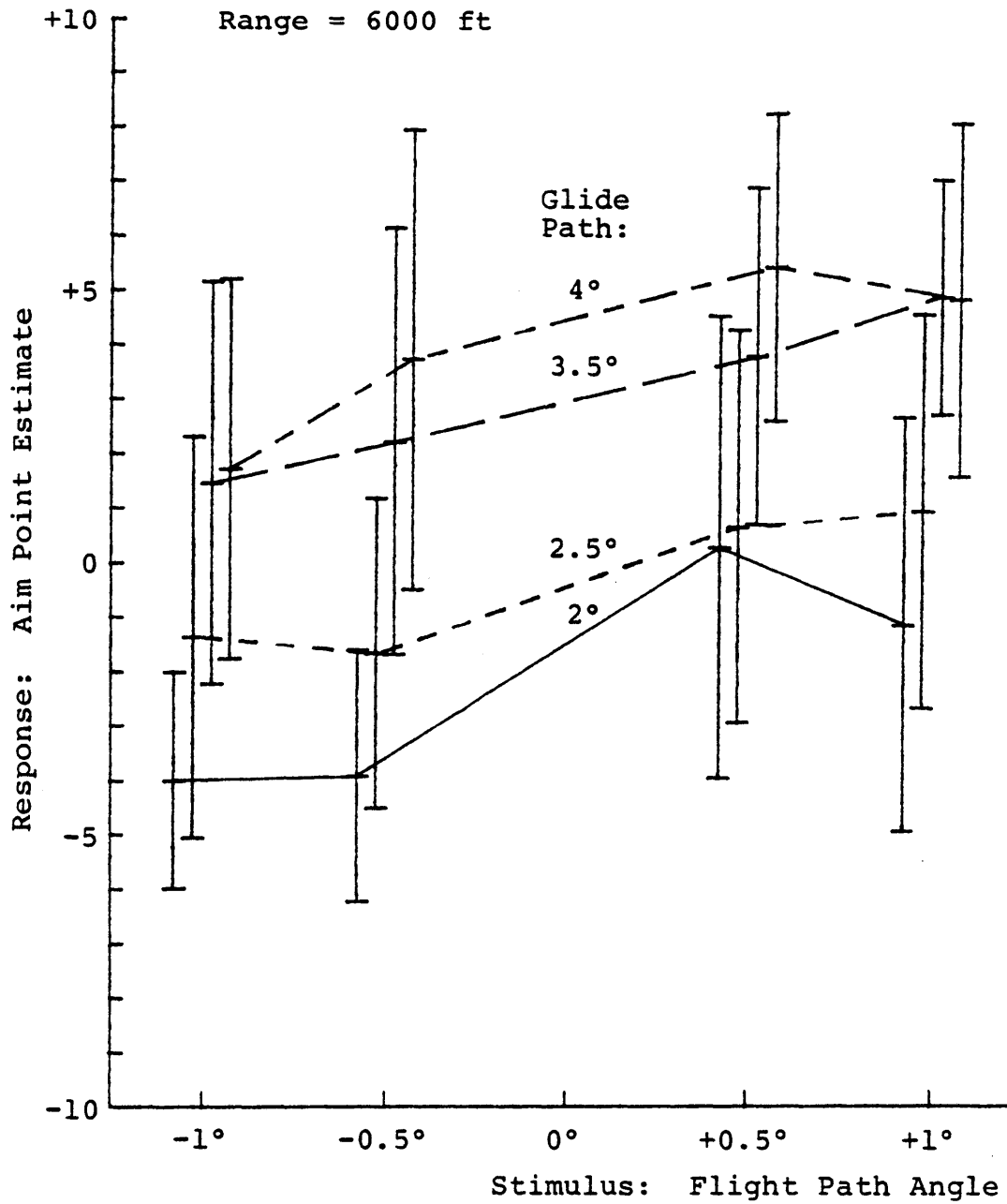


Figure C.19b Aim point data statistics for all subjects combined (at 6000 ft range).

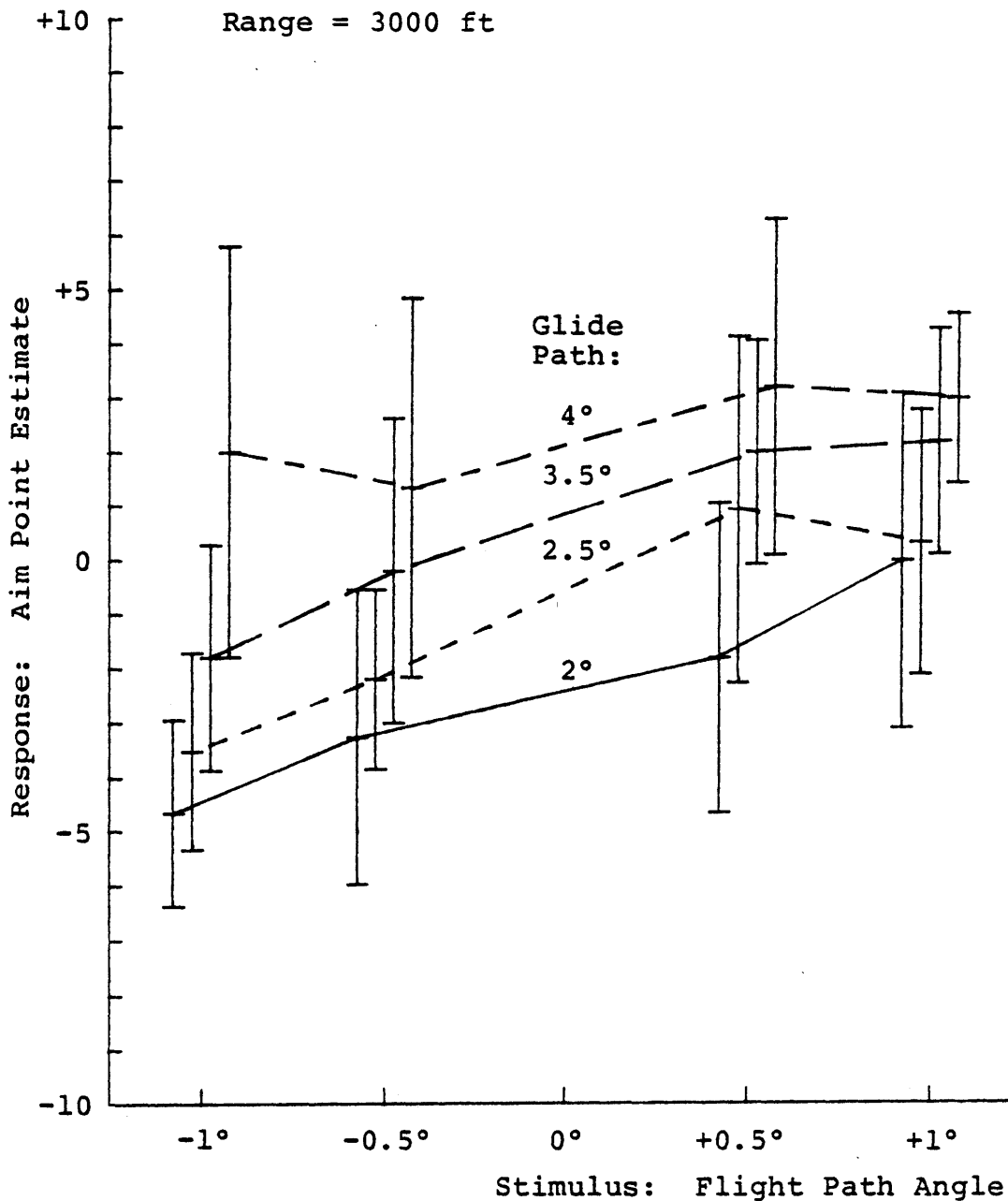


Figure C.20a Aim point data statistics for high-time pilots (at 3000 ft range).

Means \pm 1 standard deviation are shown at each aim point and glide path.

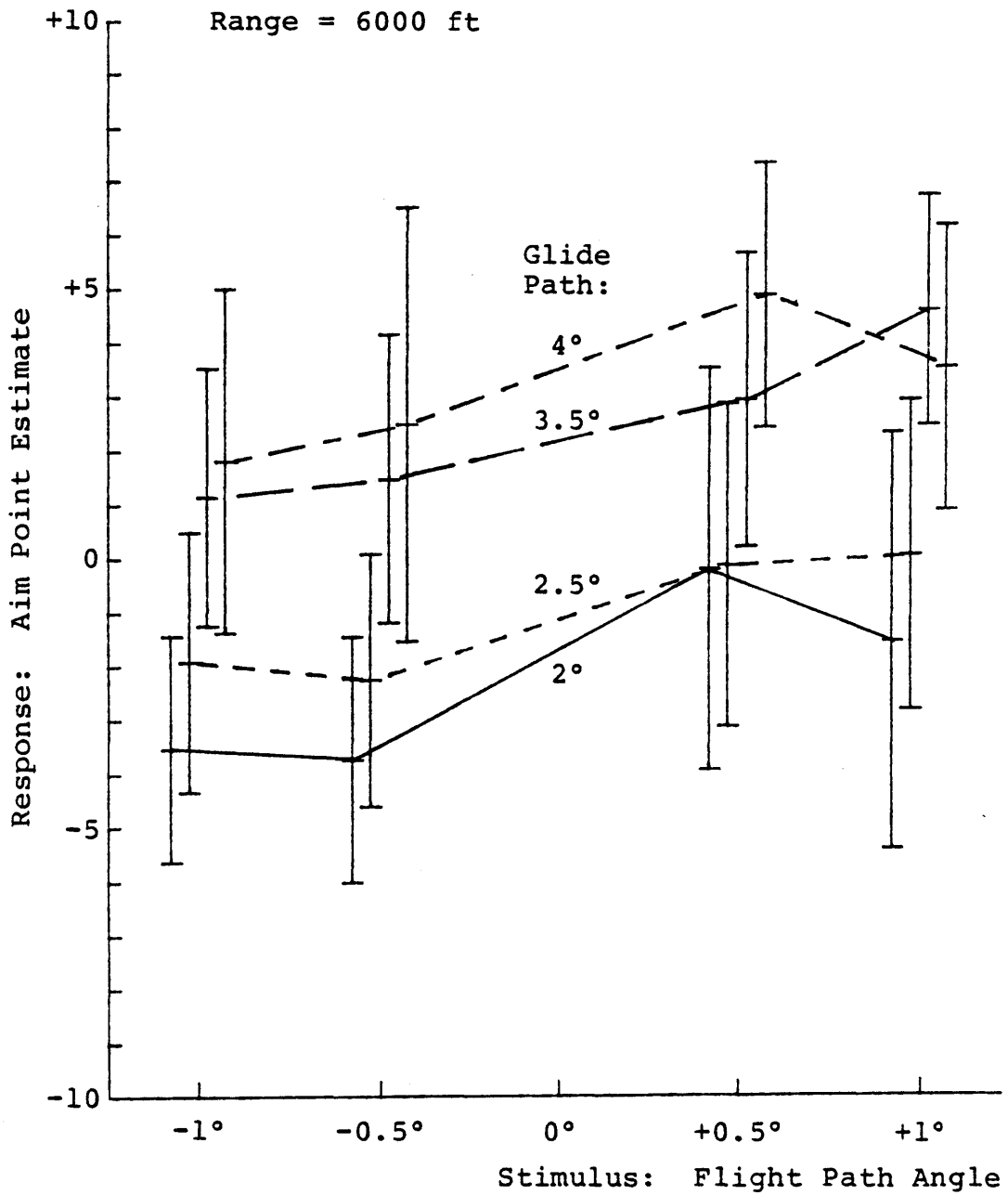


Figure C.20b Aim point data statistics for high-time pilots (at 6000 ft range).

APPENDIX D
INSTRUCTIONS FOR THE OPERATION OF THE EQUIPMENT

This appendix covers operation of the video equipment and use of the cockpit simulator in MIT Room 35-220. A description of the experimental equipment, including illustrations of the layout, is given in Section 2.1.2, Description of the Equipment. Some suggestions for future use of video recorders and editing considerations are discussed below.

D.1 Operation of the Projector

The Amphicon 260 projector is fully described in its operation manual (Amphicon Systems, 1967), which should be referred to before using the projector. A summary of the operation of the projector and suggestions for most effective use follow.

The projector is turned on and off with the brightness control on the front of the electronic cabinet. Do not turn up the brightness past the detent for at least 30 seconds to allow for warmup (there are time-delay relays to prevent damage from too-rapid on/off cycling, but it is better not to provoke problems). With a video signal applied to the upper BNC input inside the rear of the cabinet, turn up the brightness for a picture. If the brightness is turned up

too far past the mid-range on the meter, the picture will "bloom", or wash out. The extreme size of the image causes poor resolution, brightness, and contrast, especially when looking at it from close up. A better picture may sometimes be gotten by turning down the brightness and increasing the contrast instead.

Focusing is quite difficult because there are two focus controls which interact optically. The mechanical control on the projector head has considerable backlash, and electronic control on the electronic cabinet has noticeable time delay. Make small, careful adjustments, alternating between the two controls. When the raster lines are visible on the screen, the unit is very nearly in perfect focus.

If getting a proper horizontal hold proves difficult, carefully adjust the horizontal preset. It functions as a sort of coarse horizontal hold and is extremely sensitive. The tracking control on video players may also need adjustment to get a stable vertical hold.

Some modifications have been made to the Amphicon's electronic unit. Switches S1 and S2 (scan rate) have been brought out to the front panel, along with the horizontal present potentiometer. Capacitors C33 and C91 were both changed to 0.0022 μf to speed up the automatic frequency control circuit (sync discriminator feedback) to allow the projector to maintain electronic stabilization with video players, which are often much less stable than the direct

camera systems for which the projector was designed. A probe socket and relay plus override switch were added by previous users to allow remote cutoff of the projector head video signal without having to shut down the entire projector. This is not needed for simulator work and the probe override switch may be left on. The remote relay and some minor modifications to the horizontal switch circuit are diagrammed in the back of the manual.

If it becomes necessary to disconnect the cables between the projector head and the electronic cabinet, do so very cautiously after the unit has been turned off and unplugged from the power outlet for several minutes. There may be residual high voltages on some of the cables. Touch each cable to the ground as it is disconnected to discharge any high voltages.

D.2 Use of the Simulator Room and Screen

Other people will be using the cockpit simulator, so check with the supervisor of the computer room well in advance to avoid schedule conflicts. From time to time, modifications are made to the simulator, and equipment racks may be placed in the pilot's field of view. These are intended to be easily removable, but, again, check with the supervisor before moving them.

The room must be well darkened to get a good image on the screen. The hallway doors should be blacked out, and both sets of window blinds closed. To minimize reflections, black drapes (stored with the screen) should be hung from the plastic pins on the walls immediately behind the screen.

Assemble the rear projection screen frame by matching the colored tapes on the frame segments. Very low vertical clearance required some modifications to the frame. The leg mounts are offset to the outside of the vertical frame sections. Don't overtighten the nuts or disassemble the leg mounts: The frame may become permanently bent locally, and the leg mounts are quite difficult to reassemble properly. Note that the top row of snaps faces the opposite direction from the side and bottom, so that the top edge of the screen must be wrapped over the top of the frame.

The screen material is not especially strong and should not be allowed to get dirty. Make sure that it is rolled up or unrolled on a clean surface; tissue paper on the floor works well. The slick side sticks to itself, so fold the screen once, rough side in, then roll it up with paper between layers in the roll. In theory, the screen shows a better picture with the slick side facing the viewer, but experiments should be performed to see how it looks with particular tapes.

D.3 Suggestions for Video Taping

Good video recording and editing procedures are essential: The Amphicon projector will not tolerate mediocre video signals from poor tapes or players. Video-tape Recording by J.F. Robinson (1975) is a good reference as it concentrates on terminology, standards and technical procedures, rather than circuit details or theatrical productions.

Video recorders mechanically sweep their heads across the tape, thus necessitating very accurate mechanical and electronic alignment of the heads (called "head switching"). Since mechanical drives are not quite stable enough to alone maintain alignment, a series of reference pulses - the "control track" - is recorded along one edge of the tape. The playback video tape deck locks on to this control track to keep the head aligned. The control track and head switching must be synchronized with the video sync signal for proper operation. This is what causes the most difficulty in recording and editing. Mechanical splicing cannot be used because of both possible head damage and inability to maintain synchronization over the splice, so all video editing is done by electronic re-recording of the video signals.

If the control track is lost even momentarily, the playing deck may take several seconds to restabilize. The

Amphicon projector has a poor vertical sync lock, and may take several more seconds after the deck is running smoothly again to regain vertical hold. So tapes played on the Amphicon must have good control tracks and smooth transitions at edit points to avoid loss of vertical hold. Furthermore, many video recorders use the control track for edit control, so if it is missing or poorly recorded a proper edit may not be possible.

Both sets of preliminary experiments (described in this thesis) were made on a very simple 1/2 inch open reel recording deck. It was stopped and started manually for each run. This caused loss of the control track between each run, which created severe difficulties with playback on the projector. It was eventually discovered that head switching was rarely stabilized before the start of each run, further compounding the problem. Since no playing deck can run properly without a continuous control track, even simply copying could not cure the problem because the playing deck could not supply stable video signals to the copying deck, and clean editing was impossible. The final set of tapes for the experiments were made on a 3/4 inch cassette machine and had very long gaps between runs to allow time for complete stabilization. The gaps had to be edited out later, but at least this could be done reliably.

Three different recording modes may be available, depending upon the recorder: Pure "record" puts down both video signal and a new control track independently of whatever is already on the tape; "insert" records only the video signal, synchronizing it with the existing control track; and "add" lays down new signal and control tracks, but synchronizes them both with the original control track immediately preceding the beginning of recording. Either of the last two are used for editing, and if possible should be used for the original recording for the best possible edits later. There are usually provisions for two audio tracks on video recorders, and it is recommended that both be used, one for making working notes during editing, and the other for terse run identification on the finished tape.

Since the Amphicon unit is black and white only, color recordings are unnecessary. Resist the temptation to make color tapes just because they look pretty on color monitors: Some resolution and contrast may be lost on both the original recordings and on playback on the Amphicon projector.

There are several different video tape standards. Half inch open reel is inexpensive and popular in portable units, but its performance is very heavily compromised to get low cost, and only relatively crude editors are made for it. Three quarter inch cassettes are becoming the industry standard for working tapes. These have better resolution than 1/2 inch tapes. Excellent editors, some of them micro-

processor-controlled for high precision, are made for 3/4 inch cassettes. Half inch cassettes are meant for low precision home use only. There were at one time no less than five different one inch open reel video formats, all incompatible. Industry has finally settled on a standard (the newest), but it is enormously expensive (an order of magnitude beyond 3/4 inch cassettes), being intended to replace the 2 inch format for commercial broadcasting. Performance, cost, and availability point to 3/4 inch cassettes as the recommended format for most laboratory work. If the slightly higher resolution of the one inch tapes is needed, be absolutely sure that all of the equipment to be used is actually compatible.

At present, Panasonic, Sony, and JVC make the best 3/4 inch cassette units in terms of performance and reliability. The first two are about equally preferred to JVC, but this depends upon the exact model. Video recorders are very temperamental and require continuous maintenance, so there is no guarantee that an individual Panasonic or Sony deck will work well on a particular tape, and perfectly adequate results may sometimes be had on a JVC.

Reasonably good viewing facilities are available from Video Services in the MIT Center for Advanced Engineering Study, but 3/4 inch cassette editing cannot be done there. The Media Center at the University of Massachusetts (Columbia Point Campus) has excellent editing facilities, and the Film

Section of the MIT Department of Architecture should eventually have a 3/4 inch cassette editor available to other users.

REFERENCES

- Amphicon Systems, Inc., Amphicon "260" Television Projector Operation and Service Manual, Norwood, New Jersey, 1 August 1967.
- Anderson, N.H. "Algebraic models in perception", Handbook of Perception, Vol. II, (Carterette and Friedman, eds.) New York: Academic Press, 1974.
- Armstrong, B.D. "Flight trials to discover whether peripheral vision is needed for landing", R.A.E. TR 70205, 1970.
- Barnes, A.G., "Simulating the visual approach and landing", AGARD Meeting on Piloted Aircraft Environment Simulation Techniques, Brussels, Belgium, 24-27 April 1978.
- Chase, W.D., "Evaluation of several TV display systems for visual simulation of the landing approach", NASA TN D-6274, 1971.
- Crow, E.L., Davis, F.A. and Maxfield, M.W., Statistics Manual, New York: Dover Publications, Inc., 1960.
- Dorfel, G., "Methods for the validation of synthesized images in visual flight simulation", 26th GCP Meeting Symposium on the Impact of Integrated Guidance and Control Technology on Weapon System Design, Sandefjord, Norway, 9-12 May, 1978.
- Driskell, C.R. "Wide angle system developments", AGARD Meeting on Piloted Aircraft Environmental Simulation Techniques,

- Brussels, Belgium, 24-27 April 1978.
- Fowler, V.J., Stone, S.M. and Schlafer, J., "Wide angle multicolor scanned laser display system for visual flight simulators", AIAA Paper No. 70-361, AIAA Visual and Motion Simulation Technology Conference, Cape Canaveral, Florida, 16-18 March 1970.
- Gold, T., "Visual perceptions of pilots in carrier landing", AIAA Paper No. 73-914, AIAA Visual and Motion Simulation Conference, Palo Alto, California, 10-12 September 1973.
- Gold, T. and Hyman, A., "Research in visual perception for carrier landing", Sperry Rand Corporation, Gyroscope Division, Report SGD-5265-0031, December 1968.
- Hasbrook, A.H., "The approach and landing: cues and clues to a safe touchdown", Business and Commercial Aviation, Vol. 37, No. 5, 39-43, November 1975.
- Huff, E.M. and Nagel, D.C., "Psychological aspects of aeronautical flight simulation", American Psychologist, Vol. 30, No. 3, 426-439, March 1975.
- Key, D.L., Odneal, B.L. and Sinacori, J.B., "Mission environment simulation for army rotorcraft development - requirements and capabilities", AGARD Meeting on Piloted Aircraft Environmental Simulation Techniques, Brussels, Belgium, 24-27 April 1978.
- Kraft, L.L. and Shaffer, L.W., "Visual criteria for out of the cockpit visual scenes", AGARD Meeting on Piloted Aircraft

- Environmental Simulation Techniques, Brussels, Belgium, 24-27 April 1978.
- McGregor, D.M., "Some factors influencing the choice of a simulator", AGARD Conference Proceedings No. 79 on Simulation, January 1971.
- Naish, J.M., "Control information in visual flight", Seventh Annual Conference on Manual Control, Los Angeles, California, 2-4 June 1971, NASA SP-281.
- Palmer, E.A., "Experimental determination of human ability to perceive aircraft aim point from expanding gradient cues", 1969 Annual Scientific Meeting of the Aerospace Medical Association, San Francisco, California, 5-8 May 1969, 176-177.
- Poulton, E.C. "The new psychophysics: six models for magnitude estimation", Psychological Bulletin, Vol. 69, No. 1, 1-19, January 1968.
- Robinson, J.F., Videotape Recording, London and New York: Focal Press, 1975.
- Snedecor, G.W. and Cochran, W.G., Statistical Methods, Ames, Iowa: Iowa State University Press, 1967.
- Spooner, A.M., "The development of visual systems for flight simulation", Second Flight Simulation Symposium, Royal Aeronautical Society, 16-17 May 1973.
- Stevens, S.S., "On the operation known as judgement", American Scientist, 54:385, 1966.

- Welch, B.C., "Recent advances in television visual systems",
AGARD Meeting on Piloted Aircraft Environmental Simula-
tion Techniques, Brussels, Belgium, 24-27 April 1978.
- Wempe, T. and Palmer, E., "Pilot performance with a simulated
pictorial landing display including different conditions
of resolution and update rate", Sixth Annual Conference
on Manual Control, Wright Patterons Air Force Base,
Ohio, 7-9 April 1970, 47-81.