Limited Evaluation of the Longitudinal Flying Qualities of a Centerstick Aircraft with Variations in Stick Feel Parameters

William M. Quinn Jr.*
HQ United States Air Force, Pentagon, Washington, D.C.
and
Malcolm A. Cutchins†
Auburn University, Auburn, Alabama

A two-phase study was performed using an Air Force NT-33A variable stability aircraft to simulate a Class IV centerstick-controlled aircraft during Category A flight phases. In Phase I, the effect of variations in the stick force per gram and stick force per inch on longitudinal flying qualities was evaluated using air-to-air tracking of a target aircraft. In Phase II, all configurations evaluated in Phase I were re-evaluated using a head-up display (HUD) command tracking task. Results from the evaluations with a target aircraft were compared to the results obtained using HUD command tracking. Cooper-Harper ratings consistently worse than 3.5 and pilot comments corresponding to Level 2 performance were obtained during all phases. Both test phases indicated that a stick force of 10 lb/*g* was too high and a stick force of 4 lb/in. was too low for favorable comments and ratings. In both Phase I and Phase II, the combination of 7 lb/*g* with 6 lb/in. and 7 lb/*g* with 8 lb/in. received the best ratings and comments. Phase II evaluations were consistent with those from Phase I, and the use of the HUD command tracking task was adequate for revealing potential flying qualities problems. Pilot comments were more consistent than Cooper-Harper ratings.

Nomenclature

δ_s = stick deflection (in.)
F_s = stick force (lb)
n = load factor (g)
α = angle of attack (rad)
β = angle of sideslip (deg)
$\omega_{n_{sp}}$ = short-period natural frequency (rad/s)
ϕ = angle of bank (deg)
↑ = too high
↓ = too low

I. Introduction

THE dynamic short-period response of an aircraft is characterized by pitch angle, pitch rate, and angle of attack changes while essentially at a constant airspeed and altitude.[1,2] The short-period mode is an especially important aircraft handling quality because its period can approach the limit of pilot reaction time, and it is the mode a pilot uses for longitudinal maneuvers in normal flight. When coupled with a pilot, the short-period mode is critical to high-gain tasks such as air-to-air and air-to-ground tracking and formation flying.[1]

The military has long recognized the importance of proper control feel system characteristics. Characteristics deemed "proper" by the pilot are, however, dependent on the aircraft's dynamic response. Based on empirical handling qualities data, an allowable range of short-period frequencies has been established by the military.[3] Even within this range, however, a pilot may find the aircraft's flying qualities unacceptable. A new military standard is under consideration.[4]

A pilot commands aircraft longitudinal response mainly through control column ("stick") inputs. The characteristics of the feel system (stick force per *g*, stick force/deflection) can, therefore, influence the handling qualities of the aircraft. Pilots may rate an aircraft unsatisfactory because it does not "feel" right to them.[2]

This paper presents the results of a limited in-flight investigation of the effects of variations in stick force per *g* and stick force per inch on the longitudinal handling qualities of a centerstick-controlled aircraft. The investigation was conducted at a short-period natural frequency at the upper limit of the Level 1 short-period frequency range (MIL-F-8785C)[3] using air-to-air gross acquisition and fine tracking tasks. A lower limit (3 rad/s) was also to be evaluated, but time did not permit this evaluation. It was felt that the results of the stick gradient changes would be more evident starting from a borderline value rather than one near the center of Level 1.

The objective of this study was to determine if an aircraft with borderline Level 1 characteristics could be moved into solid Level 1 ratings just by varying stick feel gradients—a quicker, less expensive means of improving an aircraft than attempting to change its natural frequency.

Air-to-air tracking tasks (Phase I) were accomplished using a T-38A target aircraft that followed a specified maneuver sequence. The maneuver sequences and Handling Qualities During Tracking (HQDT) techniques used were adapted from those recommended by Twisdale and Franklin[5] for air-to-air tracking tasks. T-38A jet trainer aircraft were used as target aircraft. Flight test engineers flew in the target aircraft to monitor the consistency of target performance.

Additional testing was conducted to obtain comparative pilot evaluation data for both air-to-air tracking and head-up display (HUD) command tracking using the same aircraft configurations and evaluation pilots. This testing was conducted to determine if HUD command tracking was a valid substitute for the air-to-air tracking task in handling qualities

Received Dec. 12, 1984; revision received Mar. 30, 1987.

*Major, USAF. Deputy for Aerospace Vehicle Technology. Member AIAA.

†Professor, Aerospace Engineering. Associate Fellow AIAA.

evaluations. This would permit the elimination of the requirement for a target aircraft, with resultant resource savings and enhanced flight scheduling flexibility. The use of HUD command tasks was suggested by the results of tests reported in Ref. 6. Reference 7 has similar suggestions.

The investigation was conducted in a NT-33A variable stability aircraft with a gunsight and programmable HUD.[8-10] This aircraft was a highly modified T-33A jet trainer capable of producing a wide range of dynamic and centerstick feel characteristics.

Pilot qualitative comments supported by pilot Cooper-Harper ratings were the primary data source. Project pilots attempted to limit this evaluation to the longitudinal response of the aircraft and to ignore or minimize the effects of lateral response. To limit the effects of lateral directional characteristics on the evaluation, these characteristics (dutch roll natural frequency and damping ratio, roll and spiral mode time constants, lateral stick force and deflection gradients, etc.) were held constant, and the tracking reference was set at the fuselage reference line (0 mils depression angle) for air-to-air tracking.

It should be noted that the results of these tests apply in the strictest sense only to the limited tasks performed during this evaluation. While these maneuvers were adequate for the intended evaluation, care should be exercised in extending these results to the entire spectrum of air-to-air and air-to-ground maneuvers. Only a limited range of load factors with a concurrent limited range of pilot stress could be explored due to aircraft performance limitations (only 3.5 g). No air-to-ground maneuvers were accomplished. These results *may* well be valid for any Class IV aircraft, but that cannot be proved from *this* test. In future evaluations, verification testing should be accomplished using higher performance aircraft in an environment that includes more difficult maneuvers as well as more variety.

II. Test and Evaluation

A. Objectives

The overall objective was to investigate the effects of variations in stick force per g and stick force per inch on the longitudinal handling qualities of a centerstick-controlled NT-33A aircraft. The aircraft was evaluated as a Class IV aircraft in flight phase Category A.[3] The specific objectives during each of the two phases were (1) to determine the Level 1 boundary for variations in the stick force per g and stick force per inch at a high short-period frequency during air-to-air gross acquisition and fine tracking tasks and (2) to determine whether pilot evaluations using HUD command tracking tasks were consistent with those obtained using air-to-air acquisition and tracking tasks.

B. Test Item Description

The NT-33A is a modified T-33A jet trainer capable of variable dynamic response and control system characteristics. The response feedback variable system modified the static and dynamic responses of the basic NT-33A by commanding control surface positions through full-authority electrohydraulic servos. This arrangement, through a response-feedback system, allowed the normal T-33A stability derivatives to be augmented to the extent that the handling qualities of an existing aircraft (or a hypothetical research configuration) could be simulated. A programmable analog computer, associated aircraft response sensors, control surface servos, and an electrohydraulic force-feel system provided the total simulation capability.

The safety pilot could vary the computer gains through controls located in the rear cockpit, allowing changes in airplane dynamics and control system characteristics in flight. Functionally, the variable stability system was divided into two independent parts. The first part, the variable feel system, provided the evaluation pilot with the control feel forces, gradients, and displacements. The second part, the response-feedback flight control systems, augmented the normal T-33A dynamic responses to represent those of the vehicle being simulated. An angle-of-attack vane, a sideslip probe, accelerometers, rate and attitude gyros, and dynamic pressure pickups were used as the sensor elements.

Each aircraft configuration was determined by a set of stability derivatives. The HUD flight director configuration was a set of preprogrammed HUD steering bar commands.

C. Test Methods and Conditions

The testing consisted of two phases. Phase I examined the effect of variations in the stick force per g and stick force per inch with a high short-period natural frequency (10 rad/s) on air-to-air tasks. Phase II evaluated the HUD to present a command signal ("target") for air-to-air tracking evaluation tasks using aircraft configurations tested in Phase I. The HUD was used as a target for the tracking task. A computer-driven pitch command bar was presented on the HUD, and the pilot attempted to maintain the apex of the HUD waterline symbol superimposed on the command bar. The computer generated both step and ramp inputs to the pitch command bar. A nominal or baseline dynamics and control configuration was interspersed during all phases of testing to provide a baseline for comparison. This configuration was located in the center of the MIL-F-8785C short-period dynamics and stick force gradient requirements.[3] Test configurations were as shown in Table 1.

The evaluation pilots rated each configuration flown using the Cooper-Harper Rating Scale.[11] The test variables associated with air-to-air configurations were short-period natural frequency, stick force per g, and stick force per inch. The remaining characteristics were selected by the Calspan safety pilot to minimize deleterious effects on test results. These characteristics are shown in Table 2.

D. Phase I Testing

The assignment of test points (aircraft configurations) was accomplished with a random number generator using the following specific test constraints: insure each of the four pilots accomplished the baseline configuration once; insure each pilot accomplished one of the test configurations twice; evaluate each test configuration by at least two different pilots; evaluate each test configuration at least three times; insure each test configuration was flown twice by one pilot. Randomly selecting the configuration test order within these constraints improves confidence in the data obtained by allowing pilot rating consistency checks.

Air-to-air tracking tasks were accomplished using the sequence of maneuvers described in the following subsections. All maneuvers were initiated with the NT-33A trimmed for level flight 1,500 ft behind the T-38A target aircraft, with both aircraft in a 30-deg bank turn at 300 KIAS and 13,000 ft

Table 1 Phase I air-to-air points (13,000 ft msl, 300 KIAS)

Configuration	Stick force per g (lb/g)	Stick force per inch (lb/in.)
A	5	4
B	5	6
C	5	8
D	7	4
E	7	6
F	7	8
G	7	8
H	10	6
I	10	8
J (baseline)	6	6

Note: Short-period natural frequency was 10 rad/s for all configurations except J, which was 6 rad/s. Planned tests at 3 rad/s were not done due to time limitations.

pressure altitude. The NT-33A was not retrimmed during the tasks. Tasks were performed without using the rudder (i.e., feet on the floor).

1. *Gross Acquisition Task*

From a position approximately 30 deg angle off, the NT-33A project pilot aggressively maneuvered the gunsight pipper to the T-38A tailpipe junction. When the pipper was stable within 10 mils of the tailpipe junction, the gross acquisition task was considered complete. This task was evaluated using Maneuver #1. Adequate gross acquisition performance was defined as acquiring the target within 10 mils of the pipper with no more than two overshoots and no overshoots greater than 20 mils. Desired performance was defined as acquiring and maintaining the target within 10 mils of the pipper with one overshoot not greater than 20 mils.

2. *Maneuver #1* (Fig. 1)

The gross acquisition task began when the target initiated a 2-*g* turn upon command by the NT-33A pilot. The project pilot waited until the T-38A crossed the NT-33A canopy bow (approximately 30 deg angle off), called "hack," and aggressively maneuvered to acquire the T-38A. After the project pilot called "clear to reverse," the T-38A pilot performed a 1-*g* level reversal to a level 30-deg bank turn in the opposite direction and reestablished trim at 300 KIAS, 13,000 ft pressure altitude. The NT-33A pilot waited until the T-38A pilot called "ready" before initiating the maneuver again with a "cleared to maneuver" call. The project pilot repeated the gross acquisition task and called "knock it off" when he considered sufficient reversals had been made. Both aircraft then terminated maneuvering and set up for Maneuver #2.

3. *Fine Tracking Task*

The project pilot precisely and aggressively attempted to keep the pipper on the T-38A tailpipe junction. This task was evaluated using Maneuvers #2 and #3. Fine tracking criteria for adequate performance was a maximum of 10 mils deviation of the pipper, 90% of the time, from the target's tailpipe junction. Desired performance was 5 mils maximum deviation, 90% of the time.

4. *Maneuver #2* (Fig. 2)

Upon command by the NT-33A pilot, the T-38A pilot initiated a 2-*g* level turn at 300 KIAS, 13,000 ft pressure altitude. The project pilot called "tracking" (pipper within 5 mils of the T-38A tailpipe junction) when fine tracking began. After approximately 20 s of tracking or until the configuration was satisfactorily evaluated, the project pilot called "clear to reverse," and the T-38A pilot performed a 1-*g* reversal to a level 30-deg bank turn and re-established trim at 300 KIAS, 13,000 ft pressure altitude. The NT-33A pilot repeated the maneuver and called "knock it off" when sufficient reversals had been made to evaluate the configuration. After the "knock it off" call, both aircraft terminated maneuvering and set up for Maneuver #3.

Table 2 Air-to-air dynamic configuration (13,000 ft pressure altitude, 300 KIAS)

Parameter	Value
Short-period natural frequency (rad/s)	10, 6
Short-period damping ratio (g/rad)	0.70
n/α	29.00
Dutch roll natural frequency (rad/s)	3.20
Dutch roll damping ratio	0.35
ϕ/β	2.00
Roll mode time constant (s)	0.35
Phugoid natural frequency (rad/s)	0.09
Phugoid damping ratio	0.05
Spiral mode time constant	∞

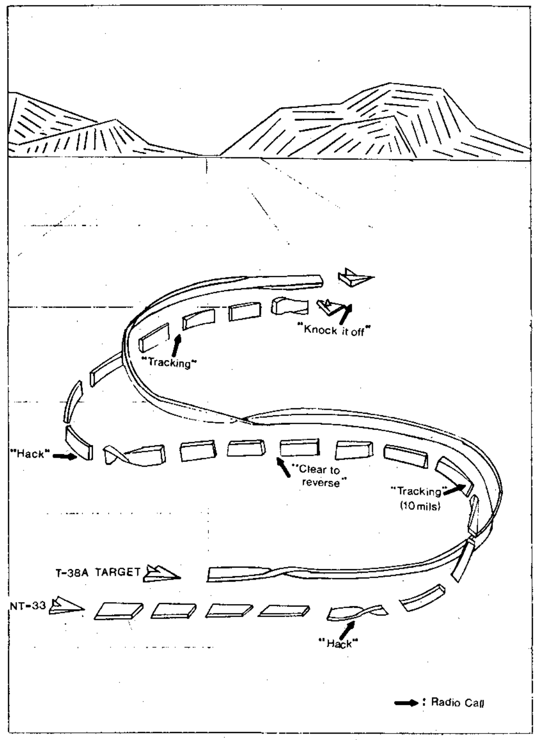


Fig. 1 Gross acquisition task (Maneuver #1).

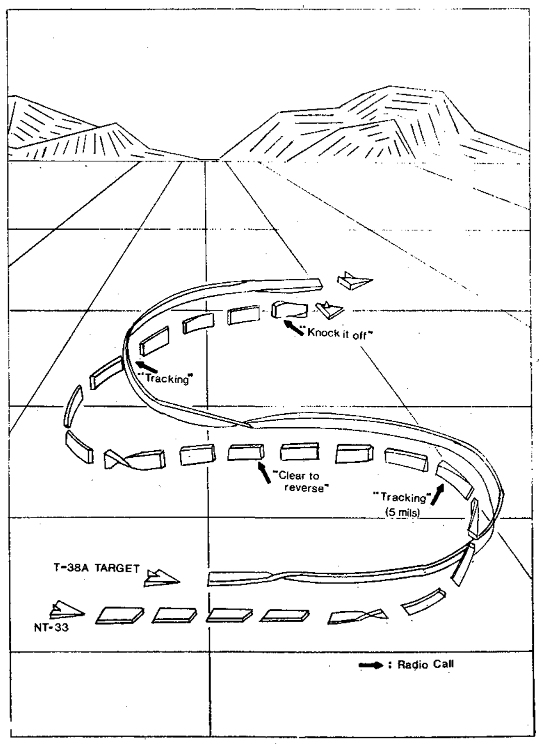


Fig. 2 Fine tracking task (Maneuver #2).

5. *Maneuver #3* (Fig. 3)

A wind-up turn was commenced on the call from the NT-33A pilot. At the call, the T-38A pilot started a wind-up turn increasing *g* at a rate of 0.2 *g*/s until reaching 3.5 *g*. This load factor was held until the evaluation pilot was satisfied with the evaluation and called "knock it off." Load factors were called out by the T-38A target pilot every 0.5 *g*. The maneuver was set up and reaccomplished in the opposite direction until sufficient reversals had been made to evaluate the configuration.

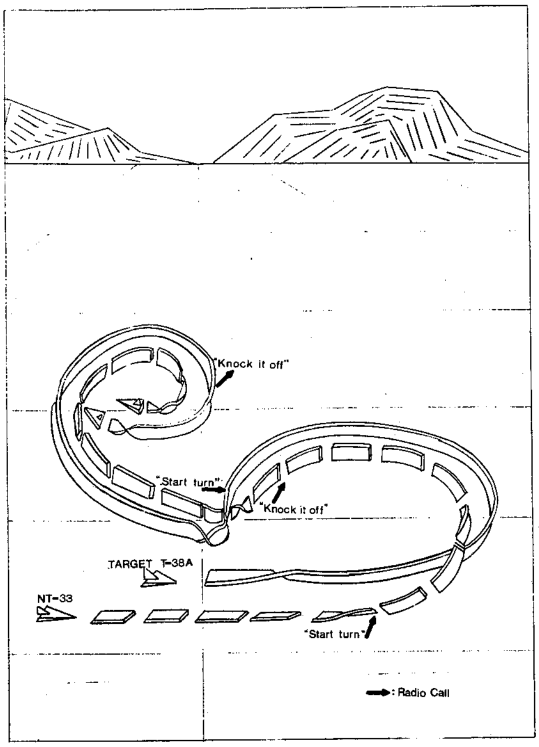


Fig. 3 Fine tracking task (Maneuver #3).

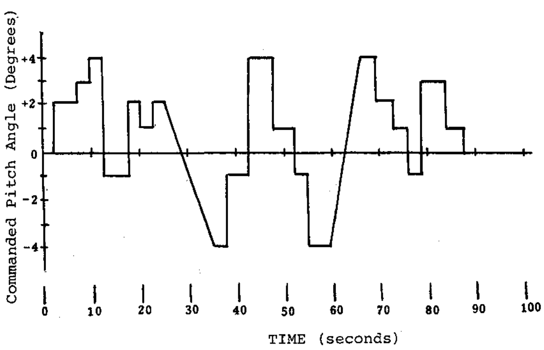
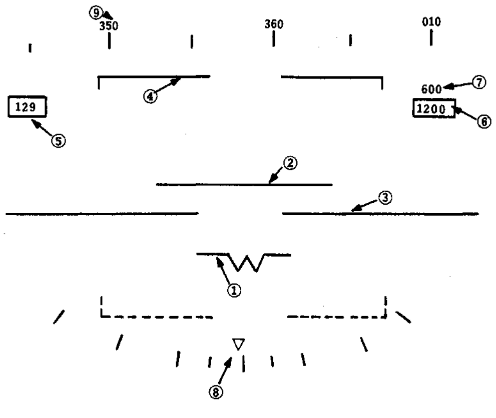


Fig. 4 HUD pitch command sequence.



1. Aircraft Waterline Symbol
2. HUD Command Bar
3. Horizon Line
4. Pitch Ladder
5. Indicated Airspeed
6. Barometric Altitude
7. Vertical Velocity
8. Bank Angle Scale (30° Maximum)
9. Heading Scale

Fig. 5 HUD symbology.

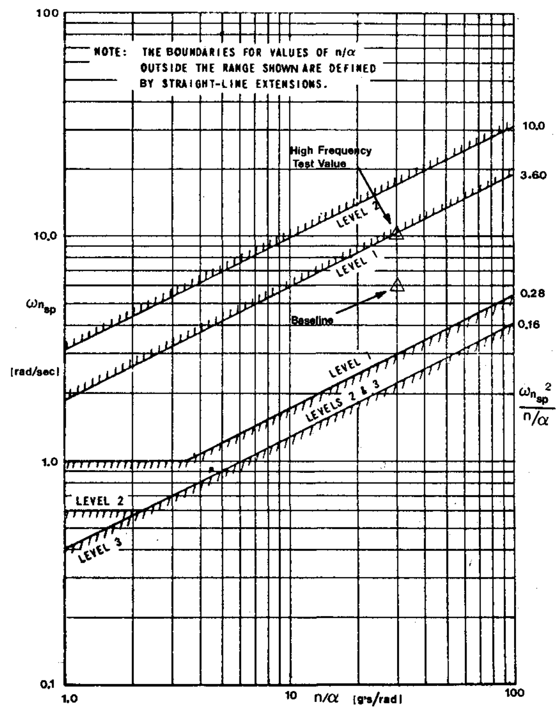


Fig. 6 Military specification (MIL-F-8785C) short-period frequency requirements, Category A flight phases (Ref. 3).

For Phase I evaluations, the pilot accomplished the gross acquisition and fine tracking task before completing an in-flight comment card and assigning a Cooper-Harper rating to each task. The pilot then accomplished the wind-up turn, completed the in-flight comment card, and assigned a rating to this task. In addition, a Cooper-Harper rating was assigned to the overall performance configuration being tested.

E. Phase II Testing

The Phase II HUD tracking task was a preprogrammed series of step and ramp movements of the pitch command bar generated by the HUD computer. The specific HUD pitch command sequence used during testing is shown in Fig. 4. During this phase, the evaluation pilot attempted to maintain the apex of the HUD waterline symbol superimposed on the pitch command bar symbol (Fig. 5).

Each pilot flew one HUD flight, during which each of the different aircraft configurations were evaluated. After each configuration was flown, the pilot completed an in-flight comment card indicating a Cooper-Harper rating. Phase II used all configurations from Phase I. All 10 configurations were flown by each pilot, and at least one configuration was repeated for each pilot to provide an indication of pilot rating consistency.

A double-blind method of testing and debriefing was used. Neither the test pilot nor the flight test engineer debriefing the mission were aware of the particular configuration parameters. This procedure eliminated any preconceptions that might have contaminated the opinions that described the test results.

Immediately after each sortie, the test pilot was debriefed. A Cooper-Harper rating was assigned to each task and an overall rating was assigned to each configuration. The pilot was also asked to comment on pipper movement, stick force and deflections, and pilot workload and compensation required to accomplish the task.

III. Results and Analysis

In general, test results for both phases reflected the fact that the stick force per gram and stick force per inch were evaluated at a Level 1 boundary value short-period natural frequency.

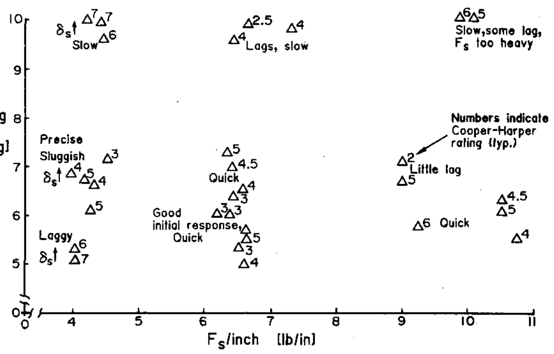


Fig. 7 Phase I, Task 1 pilot ratings and comments vs configuration.

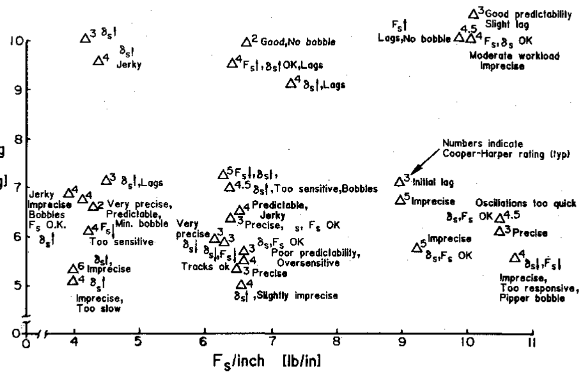


Fig. 8 Phase I, Task 2 pilot ratings and comments vs configuration.

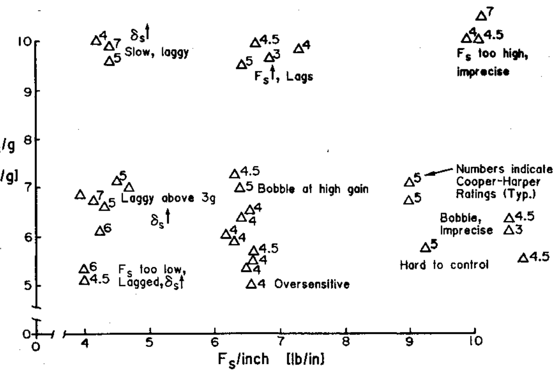


Fig. 9 Phase I, Task 3 overall pilot ratings and comments vs configuration.

Figure 6 illustrates that a short-period frequency of 10 rad/s at an n/α equal to 29 was a high-borderline Level 1 case.

Cooper-Harper ratings consistently worse than 3.5 (improvements desired) and pilot comments corresponding to Level 2 performance or worse were consistently obtained during all phases. Again, with reference to Fig. 6, either the test n/α was lower than desired or the Level 1 boundaries were too far apart. Since airspeeds and altitudes were tightly controlled during most maneuvers and significant differences from the target n/α were not observed in strip chart data, the latter seems more likely. Almost all configurations at short-period frequency equal to 10 rad/s produced comments of "oversensitive... abrupt... jerky... bobble... too quick."

Although the effects of stick force per g and stick force per inch variations were observable, the trends were heavily influenced by short-period frequency. In no case did any one value or range of values for either the stick force per g or stick force per inch improve the performance to a consistent Level 1. This called into question the validity of the Level 1 short-period frequency upper boundary of MIL-F-8785C[3] (Fig. 6).

In the remainder of this section, the effects of the stick force per g and stick force per inch on ratings and comments will be examined (by phase). A comparison of Phase I and Phase II results will be made. Finally, a discussion of pilot rating consistency will be presented. Note that the results of this limited test program may not be valid for high g loadings, since testing was not conducted above 3.5 g.

A. Phase I (Target Tracking)

Figures 7, 8, and 9 present summaries of pilot comments and Cooper-Harper ratings vs the stick force per g and stick force per inch for Tasks 1, 2, and 3, respectively. For the gross acquisition task (Fig. 7), Level 1 performance was most frequently obtained at 5.5–6.5 lb/g and 6–7 lb/in. For the fine tracking tasks (Figs. 8 and 9), similar results were obtained. However, during the wind-up turn tracking tasks (Maneuver 3), Level 1 performance could not be obtained, and this was reflected in the overall ratings for each configuration.

For all tasks, a stick force per g of 10 lb/g was too high for satisfactory performance. In spite of the short-period frequency of 10 rad/s, comments at these high stick force per g values were, "slow... laggy... stick force was too high... lags... some lag... imprecise." These comments were irrespective of variation in the stick force per inch. Pilot debriefings indicated that at a high stick force per g, initial pilot inputs were not strong enough. The pilot then observed a lack of response and increased his input. However, the lag in the process was significant and was reflected in the pilot ratings and comments. This effect was possibly partly due to pilot fatigue from higher forces and previous pilot experience with aircraft having lighter stick forces.

For all tasks, a stick force per inch of 4–5 lb/in. produced comments of, "sluggish... laggy... imprecise... stick deflection too high," irrespective of the stick force per g. This result appears analogous to the high stick force per g result discussed previously. The pilot, whether exerting light or heavy stick forces, expected a certain range of motion in the stick for the desired response. At low stick force per inch values, stick deflection was too large and required excessive arm motion to precisely track a target and make small corrections.

Figure 10 shows overall configuration ratings and comments vs the stick force per g and stick force per inch, along with the corresponding values of inch per g computed from actual stick force per g and stick force per inch settings. The high stick force per g (9–10 lb/g) and low stick force per inch (4–5 lb/in.) areas received poor pilot ratings and comments. Cooper-Harper ratings averaged above five in most cases, with individual pilots assigning ratings up to seven ("adequate performance not obtainable with maximum pilot compensation"). This leaves the lower right corner to consider. The comments here seem to be predominantly, "too sensitive... jerky... too responsive," and the pilot ratings were not much better than those for the other configurations evaluated. However, some preferences were indicated by the debriefing question, "Was this the baseline configuration?" Responses to this question are summarized in the following paragraph.

Since the pilots equated the baseline configuration to the "best" configuration, incorrect guesses that a particular configuration was indeed the baseline served to rank order the remaining test configurations. Configurations incorrectly guessed to be the baseline (or best) were: A (once), B (once), E (twice), and F (twice). Most comment data suggested that configuration A was in fact, "poor," but B, E, and F were "not all that bad." The corresponding inch per g range (0.8–1.2 in./g) for these three configurations indicated an optimum value of 1 in./g.

B. Phase II (HUD Command)

Figure 11 illustrates comments and ratings vs the stick force per g and stick force per inch. A high stick force per g of

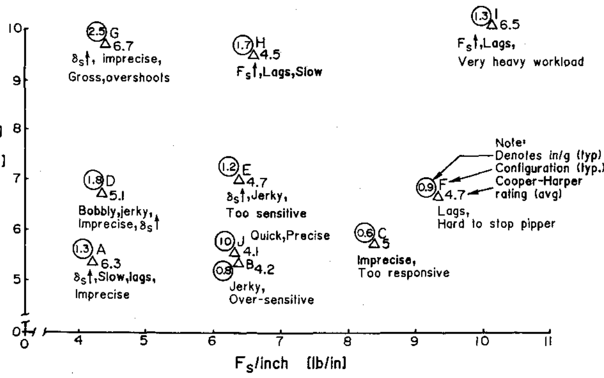


Fig. 10 Phase I overall pilot ratings and comments vs configuration.

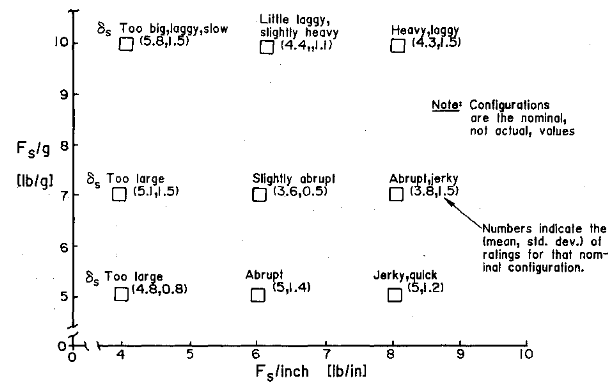


Fig. 11 Phase II pilot ratings and comments vs configuration.

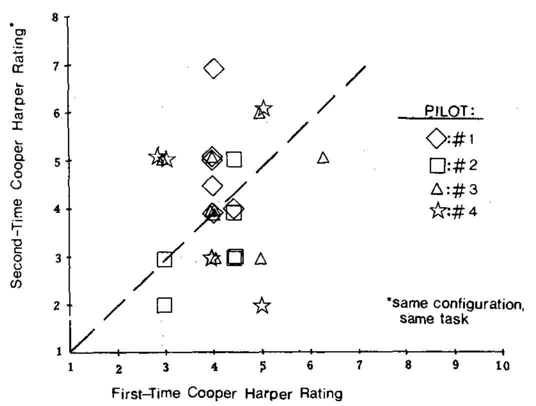


Fig. 12 Consistency of pilot ratings.

10 lb/g was excessive, and a stick force per inch of 4 lb/in. was too low for favorable comments and ratings. Configurations E and F had the best ratings. The "abruptness" comments were likely due to the high short-period frequency (10 rad/s).

C. Phase I to Phase II Comparison

Pilot comments were weighed more heavily than numerical Cooper-Harper ratings in this comparison since comments were more consistent. In this light, Phase II results were quite consistent with Phase I, and the use of HUD tracking tasks was adequate for revealing potential flying qualities problems. However, the pilots also indicated that the HUD tasks were easier than the T-38A target tracking maneuvers and somewhat artificial. Further development work should be done to improve the fidelity of the HUD director programs and symbols.

D. Consistency of Pilot Ratings and Comments

Pilot comments proved to be more consistent than the Cooper-Harper ratings. The variation of the Phase I Cooper-Harper ratings for pilots repeating specific configurations is shown in Fig. 12. Second-time ratings are plotted against first-time ratings. Ideally, a 45-deg straight-line relationship would result between the two. Not only is there a large amount of scatter about the 45-deg line, but variation in the second-time rating of the configuration is clearly greater than variation in the first-time ratings (2–7 vs 3–6). This variation indicates a learning curve in using the Cooper-Harper rating scale.[11,12] Also, the phraseology of the project pilots' comments during the initial flights of this evaluation (Phase I) showed poor standardization. Terms such as "jerky," "bobbly," or "oversensitive" were used, each with seemingly different meanings, depending on the particular pilot. This seemed to indicate different aircraft behavior. However, post-evaluation comparisons, discussions between project pilots and flight test engineers, and review and comparison of gun camera film indicated that closely similar aircraft behavior was being reported. Reference 13 addresses the relationship between adjectives used by pilots and aircraft behavior.

IV. Conclusions and Recommendations

The results of the three specific test phases are summarized herein. Phase I tests were conducted to determine the Level 1 boundary for variations in the stick force per g and stick force per inch at a short-period frequency of 10 rad/s during air-to-air gross acquisition and fine tracking tasks. Phase II tests were conducted to determine whether pilot evaluations using head-up display (HUD) command tracking tasks were consistent with those obtained using air-to-air gross acquisition and fine tracking tasks.

Cooper-Harper ratings worse than 3.5 and pilot comments corresponding to Level 2 performance were consistently obtained during all phases. The $\omega_{n_{sp}}$ chosen, while high, *was at borderline Level 1.* Pilot ratings should, therefore, have averaged around that *borderline*. Instead, the ratings were almost all Level 2, with many borderline *Level 3*. Although the variations of the effects of the stick force per g and stick force per inch were observable, the trends were heavily influenced by short-period frequency. Additional testing should be conducted to determine the validity of the Level 1 short-period frequency upper boundary of MIL-F-8785C (Fig. 6).

For the gross acquisition tasks in Phase I, Level 1 performance was most frequently obtained at 5.5–6.5 lb/g and 6–7 lb/in. For all Phase I tasks, a stick force per g of 10 lb/g was too high for satisfactory performance, irrespective of the stick force per inch. For all Phase I tasks, a stick force per inch of 4–5 lb/in. was too low for satisfactory performance, irrespective of the stick force per g. Most Phase I comment data suggested that configuration combinations of 5 lb/g with 6 lb/in., 7 lb/g with 6 lb/in., and 7 lb/g with 8 lb/in. were better than the other configurations tested. The corresponding inch per g range (0.8–1.2 in./g) for these three configurations indicated an optimum value of 1 in./g.

Phase II data indicated that a stick force per g of 10 lb/g was excessively high and a stick force per inch of 4 lb/in. was too low for favorable comments and ratings. Configuration combinations of 7 lb/g with 6 lb/in. and 7 lb/g with 8 lb/in. had the best ratings in Phase II testing. Phase II results were quite consistent with Phase I, and the use of HUD tracking tasks, although somewhat artificial, was adequate for revealing potential flying qualities problems. A higher degree of unpredictability should be incorporated into the HUD director symbol motion according to pilot comments.

Pilot comments (although not identical) proved to be more consistent than the Cooper-Harper ratings. Differences in the pilot comments/descriptions of pipper movement could be reconciled using conferences and a review of gun camera film. A learning curve in the pilots' use of the Cooper-Harper rating scale was evident. Perhaps more teaching was required, or perhaps the performance criteria were overspecified; however, the *relative* ranking of configurations would probably have stayed the same.

The effect of the learning curve on test data could have been diminished if the project pilots had flown the same configurations on a pre-evaluation training/standardization flight. They then, after making their comments and ratings individually, could have reconciled their comments/ratings (using their descriptions and a review of gun camera film) and agreed on standardized definitions for aircraft behavior.

The consistency of the Cooper-Harper ratings would also have benefited from the pilots having an opportunity to "calibrate" their perceptions of the rating scale. The possibility of the first sortie by each evaluation pilot being a training flight with each flying the same configuration should be investigated. This would enable the pilots to develop a baseline for their Cooper-Harper ratings and comments and move them further down the learning curve.

References

[1] *Flying Qualities Theory and Flight Test Techniques: Dynamics*, USAF Test Pilot School, Edwards AFB CA, June 1980.

[2] Etkin, B., *Dynamics of Atmospheric Flight*, Wiley, New York, 1972.

[3] *MIL-F-8785C, Flying Qualities of Piloted Airplanes*, Nov. 5, 1980.

[4] Hoh, R.H. and Mitchell, D.G., "Status of the Flying Qualities Mil-Standard," *Proceedings of the IEEE 1983 National Aerospace and Electronics Conference*, IEEE, NY, 1983, pp. 1316–1322.

[5] Twisdale, T.R. and Franklin, D.L., *Tracking Techniques for Handling Qualities Evaluation*, AFFTC-TD-75-1, April 1975.

[6] Monagan, S.J., and Smith, R.E., "Head-Up-Display Flight Tests," *Symposium Proceedings of the 24th SETP Aerospace Symposium*, Society of Experimental Test Pilots, Beverly Hills, CA, 1980, pp. 75–87.

[7] Boothe, E.M., Chen, R.T.N., and Chalk, C.R., *A Two-Phase Investigation of Longitudinal Flying Qualities for Fighters*, AFFDL-TR-74-9, April 1974.

[8] Hall, G.W. and Huber, R.W., *System Description and Performance Data for the USAF/Calspan Variable Stability T-33 Airplane*, AFFDL-TR-70-71, Aug. 1970.

[9] *Briefing Guide for the Head-Up Display Demonstration Program at the USAF Test Pilot School*, Calspan Corp., Buffalo, NY, Oct. 1980.

[10] *Briefing Material for Flight Demonstrations with the USAF/Calspan NT-33A Variable Stability Aircraft for the USAF Test Pilot School*, Calspan Corp., Buffalo, NY, April 1979.

[11] Cooper, G.E. and Harper, R.P. Jr., *The Use of Pilot Rating in the Evaluation of Aircraft Handling Qualities*, NASA TN D-5153, April 1969.

[12] Cooper, G.E., "Understanding and Interpreting Pilot Opinion," *Aeronautical Engineering Review*, Vol. 16, March 1957, pp. 47–51, 56.

[13] Smith, R.H. and Torgerson, W.S., "A Multivariate Approach to Handling Qualities Rating Scale Development," AFOSR-TR-80-0876, DDC Accession ADA089825, Jan. 1980.