

NASA/TM-2009-215767



# Cooper-Harper Experience Report for Spacecraft Handling Qualities Applications

*Randall E. Bailey and E. Bruce Jackson  
Langley Research Center, Hampton, Virginia*

*Karl D. Bilimoria, Eric R. Mueller, Chad R. Frost, and Thomas S. Alderete  
Ames Research Center, Moffett Field, California*

## NASA STI Program . . . in Profile

Since its founding, NASA has been dedicated to the advancement of aeronautics and space science. The NASA scientific and technical information (STI) program plays a key part in helping NASA maintain this important role.

The NASA STI program operates under the auspices of the Agency Chief Information Officer. It collects, organizes, provides for archiving, and disseminates NASA's STI. The NASA STI program provides access to the NASA Aeronautics and Space Database and its public interface, the NASA Technical Report Server, thus providing one of the largest collections of aeronautical and space science STI in the world. Results are published in both non-NASA channels and by NASA in the NASA STI Report Series, which includes the following report types:

- **TECHNICAL PUBLICATION.** Reports of completed research or a major significant phase of research that present the results of NASA programs and include extensive data or theoretical analysis. Includes compilations of significant scientific and technical data and information deemed to be of continuing reference value. NASA counterpart of peer-reviewed formal professional papers, but having less stringent limitations on manuscript length and extent of graphic presentations.
- **TECHNICAL MEMORANDUM.** Scientific and technical findings that are preliminary or of specialized interest, e.g., quick release reports, working papers, and bibliographies that contain minimal annotation. Does not contain extensive analysis.
- **CONTRACTOR REPORT.** Scientific and technical findings by NASA-sponsored contractors and grantees.

- **CONFERENCE PUBLICATION.** Collected papers from scientific and technical conferences, symposia, seminars, or other meetings sponsored or co-sponsored by NASA.
- **SPECIAL PUBLICATION.** Scientific, technical, or historical information from NASA programs, projects, and missions, often concerned with subjects having substantial public interest.
- **TECHNICAL TRANSLATION.** English-language translations of foreign scientific and technical material pertinent to NASA's mission.

Specialized services also include creating custom thesauri, building customized databases, and organizing and publishing research results.

For more information about the NASA STI program, see the following:

- Access the NASA STI program home page at <http://www.sti.nasa.gov>
- E-mail your question via the Internet to [help@sti.nasa.gov](mailto:help@sti.nasa.gov)
- Fax your question to the NASA STI Help Desk at 443-757-5803
- Phone the NASA STI Help Desk at 443-757-5802
- Write to:  
NASA STI Help Desk  
NASA Center for AeroSpace Information  
7115 Standard Drive  
Hanover, MD 21076-1320

NASA/TM-2009-215767



# Cooper-Harper Experience Report for Spacecraft Handling Qualities Applications

*Randall E. Bailey and E. Bruce Jackson  
Langley Research Center, Hampton, Virginia*

*Karl D. Bilimoria, Eric R. Mueller, Chad R. Frost, and Thomas S. Alderete  
Ames Research Center, Moffett Field, California*

National Aeronautics and  
Space Administration

Langley Research Center  
Hampton, Virginia 23681-2199

June 2009

## **Acknowledgments**

The work described in this report was funded by the Orion Crew Exploration Vehicle Program under the Exploration System Mission Directorate. The assistance and support of Jennifer Mitchell, co-lead of the Orion Guidance Navigation & Control, Operability and Pilotability mode team, was greatly appreciated. The authors also wish to acknowledge the contributions of the editorial review board for their help in creating a better report.

The use of trademarks or names of manufacturers in this report is for accurate reporting and does not constitute an official endorsement, either expressed or implied, of such products or manufacturers by the National Aeronautics and Space Administration.

Available from:

NASA Center for AeroSpace Information  
7115 Standard Drive  
Hanover, MD 21076-1320  
443-757-5802

## Table of Contents

EXECUTIVE SUMMARY .....	1
1. INTRODUCTION .....	6
2. BACKGROUND .....	6
2.1. PILOT-VEHICLE DYNAMIC SYSTEM .....	6
2.2. COOPER-HARPER RATING SCALE .....	7
2.3. ORDINAL VS. INTERVAL SCALE .....	8
2.4. PILOT COMMENTARY .....	11
2.5. SUBJECTIVE AND OBJECTIVE TASK DATA .....	11
3. HANDLING QUALITIES DEVELOPMENT .....	16
3.1. BEST PRACTICES PROCESS .....	16
3.2. COOPER-HARPER EXPERIENCE .....	16
3.2.1. SCALE DEFINITIONS .....	17
3.2.2. SCALE USAGE .....	17
3.2.3. DEFINITION OF THE “SELECTED TASK OR REQUIRED OPERATION” .....	18
3.2.4. EVALUATION SITUATION/EXTRAPOLATION .....	19
3.2.5. SPECIFICATION OF PERFORMANCE STANDARDS .....	20
3.2.6. TASK PERFORMANCE .....	23
3.2.7. USE OF PRE-TEST EVALUATION PILOTS .....	24
3.2.8. EVALUATION PILOT SELECTION .....	25
3.2.9. NUMBER OF EVALUATION PILOTS .....	25
3.2.10. EXPERIMENT DESIGN .....	25
3.2.11. BLIND EVALUATIONS .....	26
3.2.12. REPEAT EVALUATIONS .....	26
3.2.13. LENGTH OF EVALUATIONS .....	26
3.2.14. PRE-TEST FAMILIARIZATION .....	27
3.3. DEVELOPMENTAL TEST & EVALUATION .....	28
3.4. TEST AND VERIFICATION .....	31
3.4.1. US MILITARY FIXED-WING AIRCRAFT .....	32
3.4.2. JOINT STRIKE FIGHTER (JSF) .....	34
3.4.3. US ARMY ROTORCRAFT .....	34
3.4.4. FEDERAL AVIATION ADMINISTRATION .....	35
3.4.5. SUMMARY .....	39
4. SPACECRAFT HANDLING QUALITIES EXPERIENCE .....	39

4.1.	GEMINI .....	39
4.2.	APOLLO .....	40
4.2.1.	RENDEZVOUS, PROXIMITY OPERATIONS, AND DOCKING .....	41
4.2.2.	ATMOSPHERIC ENTRY .....	42
4.2.3.	HIGH-ALTITUDE ABORT .....	42
4.2.4.	LUNAR LANDING .....	43
4.3.	SPACE SHUTTLE.....	43
4.3.1.	RENDEZVOUS, PROXIMITY OPERATIONS, AND DOCKING .....	43
4.3.2.	APPROACH AND LANDING.....	44
4.4.	APPLICABILITY OF AIRCRAFT HQ TO SPACECRAFT .....	45
4.4.1.	AUTOMATION “HANDLING QUALITIES” .....	46
4.4.2.	SPACECRAFT TASK ANALOGIES .....	48
5.	RECOMMENDATIONS.....	55
6.	CONCLUDING REMARKS.....	56
	REFERENCES .....	57

## List of Figures

FIGURE 1: PILOT-VEHICLE DYNAMIC SYSTEM.....	7
FIGURE 2: FRONT AND BACK OF COOPER-HARPER PILOT RATING SCALE .....	9
FIGURE 3: EXAMPLE PILOT COMMENT CARD.....	12
FIGURE 4: TLX RATING SCALE .....	15
FIGURE 5: DESIRED AND ADEQUATE PERFORMANCE DURING BOOM AERIAL REFUELING (FROM LEGGETT AND CORD, 1994).....	21
FIGURE 6: LATERAL OFFSET LANDING TASK.....	22
FIGURE 7: HANDLING QUALITIES DEVELOPMENT PROCESS .....	28
FIGURE 8: FAA DEFINITIONS COMPARED TO COOPER-HARPER SCALE.....	36
FIGURE 9: DERIVATION OF PROBABILITY CONDITIONS.....	38
FIGURE 10: PROBE-AND-DROGUE REFUELING.....	49
FIGURE 11: TERRAIN FOLLOWING.....	52

## List of Tables

TABLE I: OFFSET LANDING TASK PERFORMANCE STANDARDS.....	22
TABLE II: MINIMUM ACCEPTABLE HANDLING QUALITIES LEVELS (FAA AC25-7).....	38
TABLE III: SUMMARY OF AIRCRAFT HANDLING QUALITIES REQUIREMENTS IN FAILURE STATE CONDITIONS (WITHIN OFE AND WITHOUT ATMOSPHERIC DISTURBANCES) .....	39
TABLE IV: PROBE AND DROGUE REFUELING TASK PERFORMANCE STANDARDS .....	50
TABLE V: TERRAIN FOLLOWING TASK PERFORMANCE STANDARDS .....	53
TABLE VI: TAKE-OFF AND CLIMB-OUT TASK PERFORMANCE STANDARDS.....	54



## Nomenclature

AC	Advisory Circular
ADS	Aeronautical Design Standard
ALT	Approach and Landing Test
BIUG	Background Information and Users Guide
CCTV	Closed-Circuit Television
CHPR	Cooper-Harper Pilot Rating
COAS	Crewman Optical Alignment Sight
CSM	Command and Service Module
DoD	Department of Defense
EP	Evaluation Pilot
FAA	Federal Aviation Administration
FAR	Federal Aviation Regulation
FCS	Flight Control System
HQR	Handling Qualities Rating
IEEE	Institute of Electrical and Electronics Engineers
JSF	Joint Strike Fighter
LFE	Limit Flight Envelope
LM	Lunar Module
Lidar	Light Detection and Ranging
NASA	National Aeronautics and Space Administration
NATO	North American Treaty Organization
NFE	Normal Flight Envelope
OFE	Operational Flight Envelope
PIO	Pilot-Induced Oscillations
RCAH	Rate Command Attitude Hold
RDT&E	Research, Development, Test and Evaluation
RPOD	Rendezvous Proximity Operations and Docking
RTO	Research Technology Organization
SFE	Service Flight Envelope
SME	Subject Matter Expert
STS	Space Transportation System
TLX	Task Load Index
US	United States

## Executive Summary

A synopsis of experience from the fixed-wing and rotary-wing aircraft communities in handling qualities development and the use of the Cooper-Harper pilot rating scale is presented as background for spacecraft handling qualities research, development, test, and evaluation (RDT&E). In addition, handling qualities experiences and lessons-learned from previous United States (US) spacecraft developments are reviewed. This report is intended to provide a central location for references, best practices, and lessons-learned to guide current and future spacecraft handling qualities RDT&E.

Handling qualities embody “those qualities or characteristics of an aircraft that govern the ease and precision with which a pilot is able to perform the tasks required in support of an aircraft role” (Cooper and Harper, 1969). These same qualities are as critical, if not more so, in the operation of spacecraft.

Handling qualities include more than just the flight control system and dynamic response characteristics of a spacecraft or aircraft. Handling qualities are characteristic of the coupled pilot-aircraft/spacecraft dynamic system which acts as a system in the accomplishment of a mission or task. Handling qualities therefore also depend upon the pilot-vehicle interface (controls and displays), the aural, visual, and motion cues involved in the required operation or task, and any stress due to the task or mission and potential external disturbances to the vehicle.

Handling qualities are assessed using the Cooper-Harper Rating scale. This scale has been the internationally accepted standard for the past 40 years. The Cooper-Harper pilot rating provides a shorthand notation summarizing the results of a piloted evaluation. The rating is subjective yet, due to the structure of the scale and by appropriate execution of the test, training, and its protocol, it quantifies the vehicle’s handling characteristics for a given task. This rating and the associated pilot comments and quantitative task performance data define the vehicles’ handling qualities.

Best practices and lessons-learned are provided in two areas:

- 1) the design and execution of handling qualities simulation and flight test evaluations (Section 3.2); and,
- 2) the overall design and development of an aircraft (or spacecraft’s) handling qualities (Section 3.3).

The best practices for the design and execution of handling qualities assessments largely follow the original basis established by Cooper and Harper in 1969, with amplification and instantiation from the works of the last 40 years. The design and definition of appropriate tasks to be flown and evaluated cannot be over-emphasized. Lessons-learned are given to improve the diagnostic results of, and reduce the variability in, the flying qualities evaluations by combinations of training, test structure, and test protocol.

For the overall design and development of a spacecraft's handling qualities, history has shown that when a vehicle designer considers handling qualities up-front and as an integral, critical part of the program, significantly less time and money are spent overall on handling qualities development than a comparable program where this up-front emphasis did not occur.

Best-practices for handling qualities suggest that a closed-loop design process should be used. The process is driven by design and mission requirements. Piloted simulation evaluations are the critical component used in the feedback loop, to assess and guide the design process. Increasing levels of sophistication, both in the model and in the simulation fidelity, should be used in this process and the resultant data must be weighed and assessed using off-line criteria and analyses.

Lessons-learned from aircraft developments suggest critical elements in ensuring excellent vehicle handling qualities and an efficient process by which to achieve them. These lessons-learned are simple in concept, but not always easy to put into effect by the vehicle management team. An overarching lesson-learned from the aircraft handling qualities history is that "more flight control system problems are caused by human behavior than for technical reasons." These lessons-learned are:

1. Understand the operational requirements and the piloting task in each phase of the mission. Ensure good communications with pilots is maintained in order to be fully aware of operational conditions.
2. Avoid over-complexity and aim to keep the flight control system design as simple and as visible as possible.
3. Beware of control systems which appear to achieve excellent performance, mainly by open-loop compensation of the nominal model. Such performance can deteriorate very rapidly when modeling tolerances are introduced or when external disturbances are applied. Such effects can be corrected by improving the closed-loop performance of the system, usually by increasing the feedback gains – although this is not always possible.
4. Plan for an integrated simulation program and ensure that all team- members (especially pilots and managers) are clear that the various simulators are for evaluation purposes, to feed data back into the analytical design process.
5. Identify the limitations of the simulation, including consideration of providing motion cues. Be aware that although simulators are of great value if used correctly, they can give misleading results if the assessments are not rigorously controlled. Simulation validation is highly desirable, if not essential.
6. Use the piloted simulator to complement the off-line design and development tools, and to intercept any design deficiencies at an early stage. The earlier problems are detected, the less it costs to fix them.

7. Use common code and data for off-line and piloted simulation to avoid unnecessary software maintenance or translation (time and cost) and the possible introduction of errors in control law functionality. Provide adequate off-line check cases to verify the control law implementation on the simulator.
8. Simulation displays and controls need to be representative, in order to avoid coloring pilot opinion of the control laws.
9. It is desirable that pilots are ‘calibrated’ in the use of development simulators, to aid their judgment of the simulated aircraft’s handling characteristics. One way of achieving such calibration is to allow them to familiarize themselves with the simulator, by flying an aircraft with which they have flying experience.
10. Deliberately search for handling problems, including the effects of design tolerances (parameter uncertainties) and failures. Identify the worst cases and any hidden weaknesses in the design, and fully explain any unexpected simulation results.
11. Evaluate the ability of the pilot to enter the control loop, to help out the automatic functions. Show that there is no tendency for divergence between the automatic and manual control functions.
12. An Integrated Product Team for flight controls/flying qualities should be formed, covering all the skill areas required to develop a flight control system. This team should be responsible for tracking the design, development, and test of each component, and the implementation and verification of each interface.
13. Deliberations should be encouraged, to bring into public view any problem or area of concern, so that all attendees can assess possible interactions with their area of responsibility, or where appropriate, potential solutions to “system” problems which may involve components other than those which encountered the anomaly. Including all components and interfaces in the discussions should be stressed, since a system problem can be generated by a component that is performing well within the performance boundaries specified for it as a unit.

Of these lessons-learned, the one that creates the most programmatic consternation is Number 10 – to deliberately search for handling problems. This process is typically cited as one of the greatest failings for a program. The reason is that it is counter-intuitive to management. Instead of the team working to build a good vehicle, it would seem this step is trying to find reasons for failure. In fact, it is just the opposite. By diligently searching for problems, including the effects of design tolerances (i.e., parametric uncertainties) and failures, testing worst case scenarios and searching for hidden weaknesses, the team is proactively heading off problems that might otherwise emerge late in the design. This period of “exploration and discovery” must be conducted. If properly done, this work makes the subsequent test and verification phase a “non-event.”

Future spacecraft are anticipated to be highly automated, if not, autonomous. Human-automation interface requirements are not often thought of as being part of “manual control” handling qualities requirements, but they should be. Although a control task may be automated, history has shown that the best automation designs are “human-centered.” A human-centered automation design takes advantage of the fact that the world’s best adaptive controller – the pilot – can intervene, adapt, and overcome as necessary in the event that the automation is not successful.

Aeronautics-domain research and development history with automation has shown that human-automation interface requirements should be developed and evaluated in parallel and in concert with the vehicle’s handling qualities. “Human-centered” automation designs principles are many and varied but should be adhered to in the design process.

The “handling qualities” of automated tasks (i.e., human-automation interface requirements) should be evaluated in three ways:

- 1) Pilots should conduct handling qualities of all tasks to be flown by the automation. By conducting these evaluations, the pilot gains an appreciation of what the automation must do to successfully control the vehicle. This knowledge is critical for understanding the behavior of the automation and what the pilot must do in the event (likely or unlikely) that they need or want to intervene or take-over. This task also defines what information (e.g., out-the-window visual cues or displays) are needed by the human to monitor, control, or interact with the automation.
- 2) Classic “handling qualities” evaluations must also be conducted in scenarios where, during the conduct of automated tasks, the pilot takes control of the vehicle and completes the task or temporarily takes-over and then re-engages the automation. This task evaluates the ability of the crew to take-over for, or to intervene with, the automation and the potential for upsets or discontinuities in the automation during this process.
- 3) Finally, “handling qualities” evaluations must also be conducted in scenarios where, during the conduct of automated tasks, the automation fails (passively or actively) and the pilot must take control of the vehicle and complete the task.

Historical evidence has shown that a design which provides excellent handling qualities enables four key benefits:

1. Task performance which meets the mission requirements both in terms of precision and accuracy, with tolerable pilot workload.
2. A more robust vehicle system, elastic to changes in task, stressors, and external disturbances, including pilot distraction.

3. Less sensitivity to pilot technique and hence, lower training costs.
4. Less risk in the design and higher safety margins in the operation of the vehicle.

Although these historical lessons-learned are predominately aeronautics-domain based, US spacecraft developments (Gemini, Apollo, and the Space Shuttle) have shown similar experiences, thus demonstrating that these lessons-learned are directly applicable. This work provides a basis from which to avoid the construct that “those that cannot remember the past are condemned to re-live it.”

# 1. Introduction

In the following, a synopsis of experience from the fixed-wing and rotary-wing aircraft communities in handling qualities development and the use of the Cooper-Harper pilot rating scale is presented as background for spacecraft handling qualities research, development, test, and evaluation (RDT&E). In addition, an overview of handling qualities experiences and lessons-learned from previous United States (US) spacecraft developments are also reviewed. These data are not nearly as plentiful as the aircraft data (for obvious reasons) but are offered as insight for the future spacecraft programs.

This report is not intended to be a comprehensive, “one-stop” location for all data but rather, provides a central location for best practices and important lessons-learned to be used as “take-aways” for the spacecraft RDT&E. References are given for those that desire additional information behind these data.

## 2. Background

Handling qualities embody “those qualities or characteristics of an aircraft that govern the ease and precision with which a pilot is able to perform the tasks required in support of an aircraft role” (Cooper and Harper, 1969). These same qualities are as critical, if not more so, in the operation of spacecraft.

In the following, background of what “handling qualities” encompass and the use of the Cooper-Harper pilot rating scale for the assessment of handling qualities are reviewed. This review is extracted verbatim in many instances from the background information provided in the original Cooper-Harper report (NASA TN D-5153) and their Wright Brothers lecture paper written in 1984 to emphasize that many of the handling qualities issues being discussed within the current US spacecraft development programs are not new or novel. However, this work also emphasizes the criticality of developing good handling qualities and the methods by which to do so. This work must be based on understanding the historical lessons-learned, rather than re-living the unfortunate consequences of those that did not heed these lessons.

### 2.1. Pilot-Vehicle Dynamic System

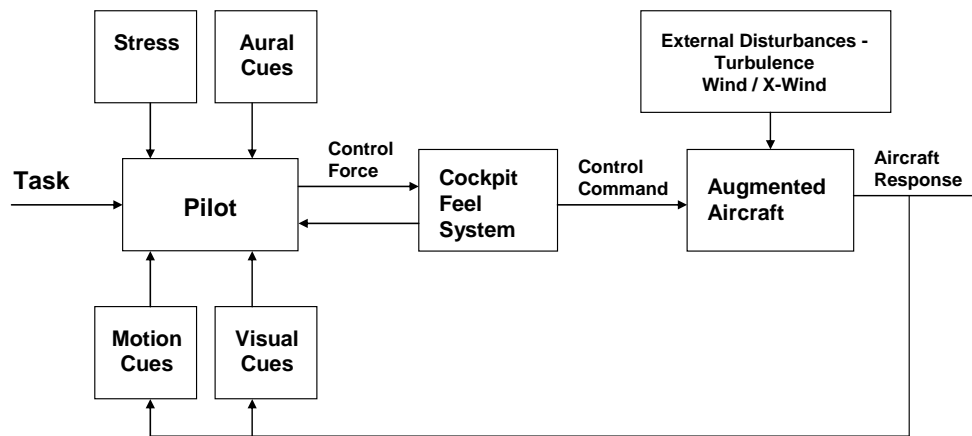
The term "handling qualities" includes more than just the flight control system and dynamic response characteristics of a spacecraft or aircraft. Handling qualities are a characteristic of the combined performance of the pilot and aircraft dynamic system acting together as a system in the accomplishment of a task.

The concept of the “pilot-vehicle dynamic system,” shown in Figure 1, illustrates this concept of handling qualities where the elements of this combined system form a closed-loop system, driven by a piloting task or objective. Handling qualities reflect the precision that the pilot can accomplish the task as the controller of the closed-loop system and the associated pilot workload and compensation to meet this level of performance. The diagram shows that the “augmented aircraft” – that is, the vehicle’s dynamic

response characteristics augmented by its flight control system – is the primary determinant of handling qualities, but other factors also influence the resultant handling qualities, including the cockpit interface (e.g., displays (visual cues), the presence or absence of motion cues, the controllers (cockpit feel system)), the environment (e.g., external visibility, control upsets, aural cues) and pilot stressors.

The pilot’s role as delineated in Figure 1 is to serve as “the decision-maker of what is to be done, the comparator of what’s happening vs. what he wants to happen, and the supplier of corrective inputs to the aircraft controls to achieve what he desires” (Harper and Cooper, 1984). Modifications or changes in the elements within this closed-loop system may be compensated for by the pilot, but possibly at a cost of pilot workload or changes in task performance. These effects cannot be segregated; thus, handling qualities must really be evaluated in the aggregate (Cooper and Harper, 1969). Historical data can provide perspective and estimates on the effects of changing elements within the system – for example, the effect of increasing the breakout force on a pitch controller – but the only truly accurate measure is to evaluate the aggregate closed-loop system.

### Pilot-Vehicle Dynamic System



**Figure 1: Pilot-Vehicle Dynamic System**

Since, by definition, simulation entails approximations for some or most of these elements within the pilot-vehicle dynamic system, the resultant handling qualities evaluations will be deficient in this regard. The magnitude and consequence of these deficiencies would ideally be assessed or quantified somehow (e.g., Brandon et al, 1995; Fields et al, 2003) so that the simulation evaluations can be related to the actual vehicle in flight.

### 2.2. Cooper-Harper Rating Scale

To the consternation of program managers and sometimes, engineers, what was true in 1969 is still true today: “pilot evaluation still remains the only method of assessing the interactions between pilot-vehicle performance and total workload in determining suitability of an airplane for the mission” (Cooper and Harper, 1969). In response to this



challenge, Cooper and Harper laid out a systematic method for the assessment and evaluation of handling qualities. The processes and definitions establish a common basis to achieve reliable and comparable data among pilots. In Section 3, the lessons-learned in conducting piloted handling qualities evaluations are reviewed to provide insight into how this process should be conducted to achieve reliable and comparable data.

Overwhelmingly, the focus of handling qualities has centered on the Cooper-Harper rating scale itself (Figure 2). As noted in Harper and Cooper, 1984: “The use of one scale since 1969 has been of considerable benefit to engineers, and it has generally found international acceptance. One problem has been that the background guidance contained in Cooper and Harper has not received the attention that has been given to the scale.” For a subjective rating scale to produce reliable and comparable data, two things are critical:

- First, the terminology used within the scale must be defined and understood by all persons (Cooper and Harper, 1969). The original scale that was distributed was a two-sided scale. The second side (shown in Figure 2) included the definitions to be used while using the scale. This “B-side,” as noted by Harper and Cooper, has been lost over the years.
- Second, a pilot rating will be meaningful only in proportion to the care taken in developing the program; that is, “in defining objectives, the role and mission, the evaluation task, what the rating applies to, the simulation situation and extent of pilot extrapolation involved.” “Large disagreement between pilot ratings is usually traced to incomplete program development.” Cooper and Harper, 1969

What also has been lost over the years is that the Cooper-Harper rating is not an end to itself, but a means to the end. The rating itself is an expedient – a shorthand notation to summarize the pilot’s evaluation. The pilot commentary data represent the more important result, capturing the “true” handling qualities data.

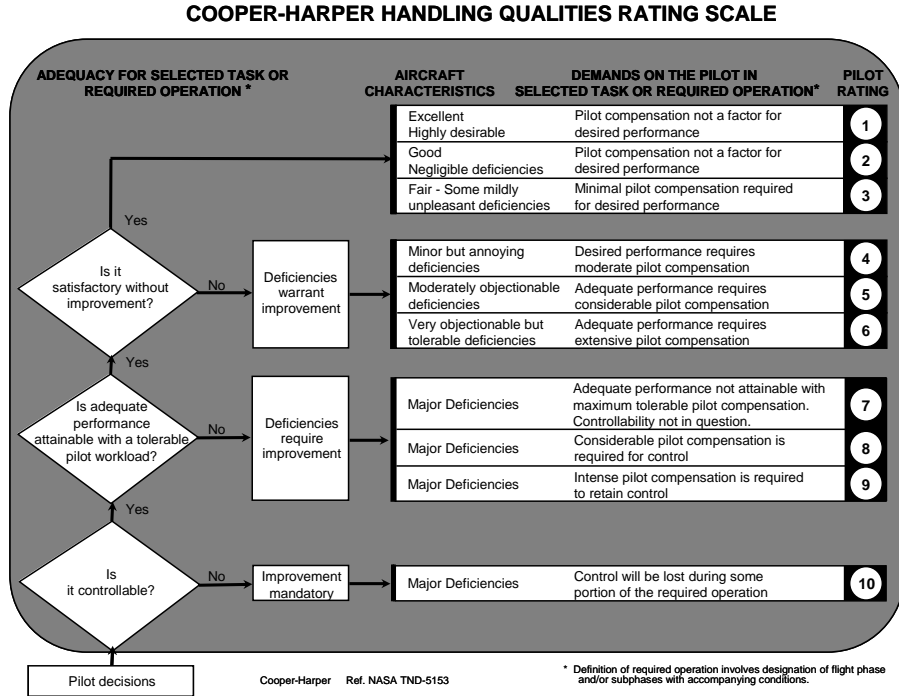
### **2.3. Ordinal vs. Interval Scale**

The Cooper-Harper Rating scale is an expedient – providing a shorthand notation – summarizing the results of the piloted evaluation. The critical data are the accompanying pilot comments and the engineering analysis and description of the mission, task, and simulation/flight environment under which the evaluation was conducted.

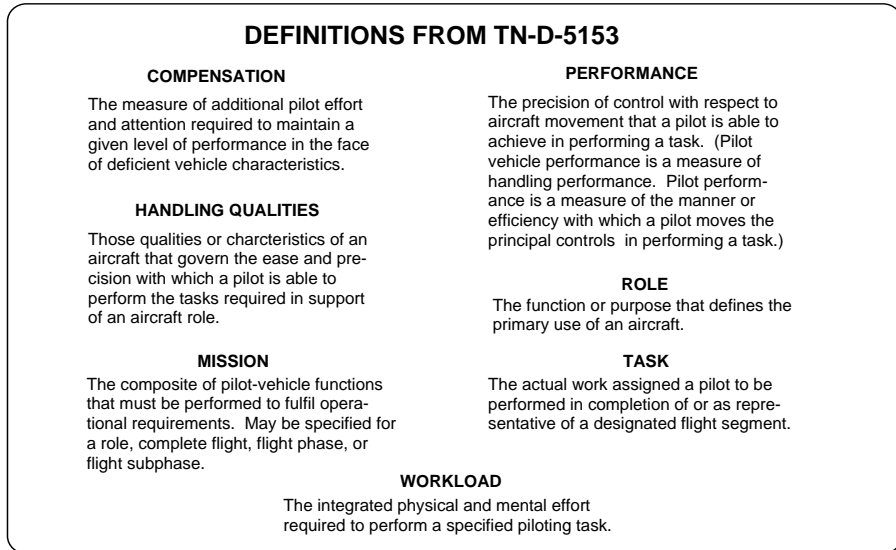
As part of the “expedient” representation, the term handling qualities “levels” are often used where:

- Level 1 denotes handling qualities that are satisfactory without improvement - Cooper-Harper ratings between 1 and 3;
- Level 2 denotes handling qualities which exhibit deficiencies that *warrant* improvement - Cooper-Harper ratings between 4 and 6; and,

- Level 3 denotes handling qualities which exhibit deficiencies that *require* improvement - Cooper-Harper ratings between 7 and 9.



(FRONT)



(BACK)

**Figure 2: Front and Back of Cooper-Harper Pilot Rating Scale**

As stated by Cooper and Harper, “an interval scale is desirable, but the proposed pilot rating scale cannot be shown to be an interval scale. The authors have accepted it as being ordinal.” Cooper and Harper were originally reluctant to associate numerals to handling qualities ratings with this ordinal scale since “engineers will attempt to treat the pilot rating data with mathematical operations that are rigorously applicable only to a linear interval scale” (Cooper and Harper, 1969). This, indeed, has happened on numerous occasions, whereby, for instance, McDonnell (1968) attempted to show an underlying interval scale (and an associated attempt at creating “conversions” between the two scales). This concept was reinvestigated in Mitchell and Aponso, 1990. Further instantiations have attempted to associate the ordinal Cooper-Harper scale ratings to the probability of loss-of-control (Hodgkinson, 1995). This latter work is intuitive and insightful but it mathematically takes liberties on the ordinal properties of the Cooper-Harper scale.

As a general rule, “although some insight is sometimes gained, analysis of specific pilot rating data should not be totally dependent on such mathematical operations” (Cooper and Harper, 1969). In the response to McDonnell’s IEEE journal article on his work (McDonnell, 1969), George Cooper responded that “The final Cooper-Harper scale is not claimed to be an interval scale as recommended by Mr. McDonnell. Furthermore, it is not clearly evident that this is an essential requirement. In fact, the main argument advanced for an interval scale is that it permits mathematical manipulation of pilot ratings, i.e., averaging etc. We are in fact concerned about indiscriminant attempts to perform such manipulations and would recommend careful examination of the supplementary pilot comments that accompany the pilot rating data and the determination of reasons for significant differences in rating by different pilots.”

As discussed in their recommendations (Cooper and Harper, 1969), each rating and the accompanying pilot comment and task data must be carefully considered in the assessment of handling qualities. The use of mathematical manipulations should be cautiously considered. If pursued, only analytical tools appropriate to ordinal data (and also typically, small sample sizes) should be used (e.g., see Dukes, 1985). Some analytical work, such as associating an increased probability of loss-of-control to poor flying qualities (Hodgkinson et al, 1992), is intuitively obvious and warrants merit as it highlights the importance of handling qualities to managers and engineers in the design process. But statistical rigor to mathematically prove this association and others is often difficult due to the ordinal nature of the scale and the small sample sizes from most experiments.

Another point that often needs to be considered in pilot evaluations is that the Cooper-Harper scale is an absolute scale rather than a relative one. “The pilot rating is given for a configuration in the context of its acceptability to the pilot for the specified flight phase (or task) and not in terms of its goodness with respect to a configuration already evaluated” (Cooper and Harper, 1969). Nonetheless, this issue continually creeps into the use of Cooper-Harper ratings, particularly as it affects the use of a Cooper-Harper rating of 1: “Pilots are reluctant to rate something as excellent or optimum for fear that a subsequent configuration will be better than anything they considered possible” (Cooper and Harper, 1969). This issue should be addressed in the pilot briefing and training

process (as detailed in Section 3) whereby pilots are encouraged to use the entire scale and that a rating of 1 does not mean optimal – it means that the aircraft characteristics are “excellent, highly desirable” and that “pilot compensation is not a factor in achieving desired performance.” More than one configuration in the world can exhibit characteristics meriting ratings of 1.

#### **2.4. Pilot Commentary**

Pilot ratings, without the comments, are only part of the flying qualities story. (Cooper and Harper, 1969). The comments reflect the “real data” which supplements the short-hand expedient rating.

Pilots should be encouraged to provide specific comments, using a comment card developed for the test, to serve as a “checklist” to tease out specific comments of interest for the engineers. An example pilot comment card from a recent Spacecraft Handling Qualities proximity operations and docking test is shown in Figure 3.

Experience has firmly cemented the fact that if the handling qualities are good, the pilot comments are short and sweet. On the other hand, pilot comments, when there are handling qualities deficiencies, are usually quite long and elaborate. In any event, the pilots should be briefed to follow the comment card, be concise yet comprehensive in describing the observed handling qualities characteristics while not trying to engineer a fix or hypothesize solutions in real-time. (Fixes and solutions are evaluated and created ideally using a slightly different process as discussed in Section 3.)

Pilot comments are critical for two reasons:

- 1) “to the airplane designer who is responsible for improving the handling qualities and to the engineer who is attempting to understand and use the pilot rating data.” (Cooper and Harper, 1969)
- 2) as “a means of assessing whether his objections (which lead to his summary rating) were related to the mission, to some extraneous factor in the execution of the experiment, or to his inaccurate interpretation of various aspects of the mission.” (Cooper and Harper, 1969)

In essence, the commentary provides the basis by which to fix a design (if necessary), use the data, or identify if the test is measuring something unintended or unexpected by the testers. Without these data, the Cooper-Harper ratings can be misleading or irrelevant to these intended purposes.

#### **2.5. Subjective and Objective Task Data**

The three critical handling qualities elements for a given task are the resultant performance, pilot workload, and pilot compensation. Again, to the consternation of engineers, in particular, “pilot evaluation still remains the only method of assessing the interactions between pilot-vehicle performance and total workload in determining suitability of an airplane for the mission.” “Pilot evaluation is like most forms of

experimental data ... since it is a subjective output of the human, it can be affected by factors not normally monitored by engineers.” (Harper & Cooper, 1984). Best practices show that the design of the experiment and protocol is critical – as discussed in Section 3 – and that subjective and objective data should be obtained to provide insight and diagnosticity into the handling qualities evaluations and the rating processes.

**SHaQ Pilot Comment Card**

1) Assign Cooper-Harper Pilot Ratings

2) Translational Control

*Rate Acceptability of Translational Control for Mission/Task:*

Very Unacceptable                      Neutral                      Very Acceptable

1      2      3      4      5      6      7

*Please Comment On:*

- a. Translational Control Power / Sensitivity?
- b. Ability to Precisely Control / Maintain Position & Closure

3) Rotational Control (If Applicable)

*Rate Acceptability of Rotational Control for Mission/Task:*

Very Unacceptable                      Neutral                      Very Acceptable

1      2      3      4      5      6      7

*Please Comment On:*

- a. Rotational Control Power / Sensitivity?
- b. Ability to Precisely Control / Maintain Attitude

4) Control Coupling:

*Rate Acceptability of Coupling for Mission/Task:*

Very Unacceptable                      Neutral                      Very Acceptable

1      2      3      4      5      6      7

5) Summary / Overall Comments

- a. Any Change in Pilot Rating?
- b. TLX Rating

**Figure 3: Example Pilot Comment Card**

The inherent subjective nature of handling qualities hasn’t stopped engineers from trying to quantify handling qualities or remove the “uncontrollable” pilot-subjective element from the process of determining handling qualities. The techniques have ranged from cockpit-based expert systems (Gungras et al, 1996) to automated rating techniques using physiological data (Suchomel, 1996) and fuzzy-logic (Tseng, Guptat, and Schumann, 2006).

As discussed in Section 3, the definition of the mission, the task, and the desired and adequate performance standards are critical to the handling qualities evaluation process. Obviously, this task performance should be measured. The measured task performance,

not only provides quantitative measure of the pilot-vehicle system performance, but provides for diagnosis as to whether the cause or effect for any handling qualities deficiencies were induced by performance. One limitation however, is that these measures do not typically quantify “control” per se. Although concepts attempting to identify pilot-induced oscillations – a form of loss-of-control – from time history data have been tried (Elliott, 2007; Mitchell et al, 2005; Fabre and Raimbault, 2001), the reliability of these methods and their correspondence to the handling qualities processes have not been verified and validated.

Additional diagnostic insight is provided by workload measures. The NASA Task Load Index (TLX) is ideal for this task since its multi-dimensional aspects provide insight into the sources of workload and hence, specific influences on handling qualities (Figure 4). The only drawback to TLX is that to develop a highly reliable workload measure for each pilot, a paired-comparison between the TLX workload dimensions should be conducted to develop weighting factors for subsequent computation of an overall workload score (Hart and Staveland 1988). Fortunately, others have shown that the paired-comparison weighting determinations are not necessary, as a simple averaging results in no loss of accuracy (e.g., see Moroney, Biers, Eggemeier, and Mitchell, 1992).

The TLX and other scales like it reflect subjective pilot workload. Objective measures of workload, such as physiological indices have also been attempted and, with some success, provide quantitative measures (Martin, G.F., and Eggemeier, 1991). One intuitively obvious measure, pilot control activity, is continually attempted but has been found to be an inherently flawed measure of pilot workload. The control activity measures are confounded by pilot strategy, training, and experience as well as the fact that these measures are task-specific (Gawron, 2000).

For example, one would intuitively expect that as the handling qualities degrade that pilot workload would increase and increasing control activity would be an indicator of this trend. First, at a minimum, control activity may only be related to physical work. It is certainly not associated with mental or other dimensional aspects of pilot workload. In the Cooper-Harper rating process, pilot workload is the “integrated physical and mental effort” required to complete the task. Second, as handling qualities degrade, different control strategies may be employed by the pilot. For instance, as control system delay increases, pilots typically reduce their “gain” to avoid large control inputs which may induce adverse coupling with the vehicle (Bailey et al, 1987). Because the control activity would be decreasing, pilot control activity metrics would suggest improved flying qualities; whereas, the subjective handling qualities ratings would reflect degraded handling qualities due to increased pilot compensation, increased mental workload, and degraded task performance and control. Lastly, control activity is an individualistic quantity, influenced by experience and training, which changes from task to task. The same task, in different environmental conditions (e.g., winds and turbulence) will necessitate different control activity to achieve the same performance. Numerous other disassociations like these have been found (e.g., Schultz et al, 1970). As such, subjective measures, while not ideal, are indicative of the flying qualities rating process.

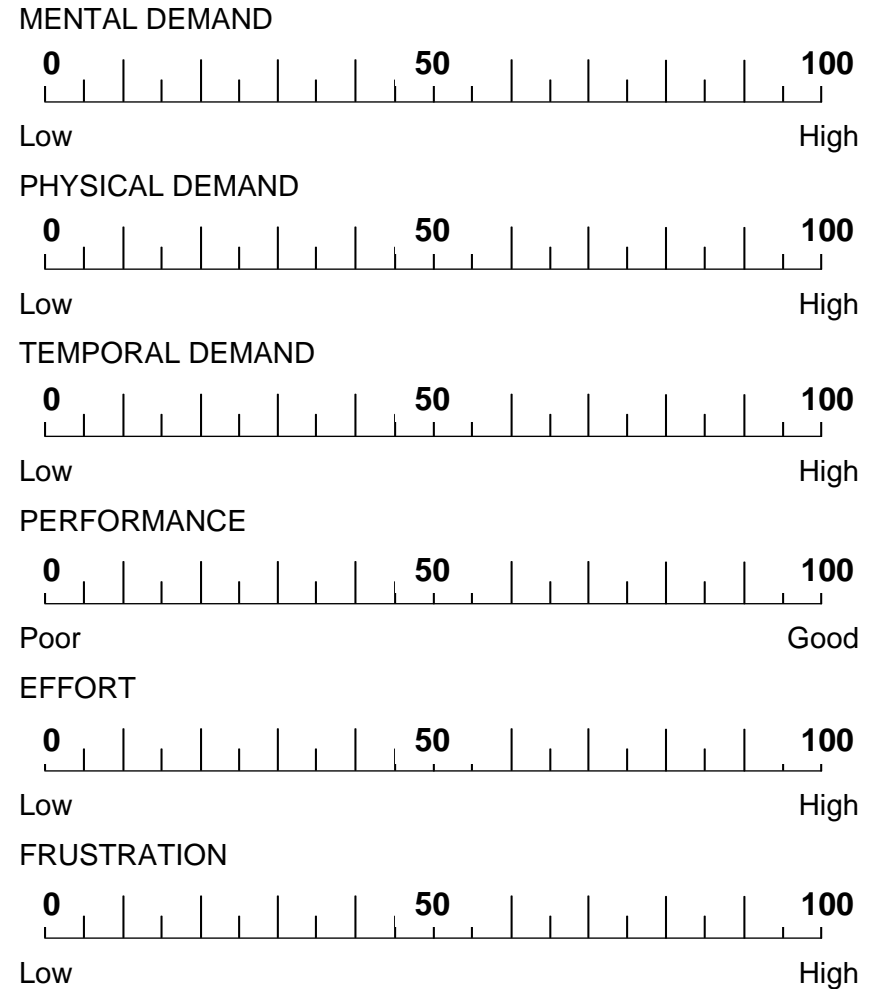
The last “ingredient” – compensation – is even less amenable to objective data analysis than the others. Compensation is defined as “the measure of additional pilot effort and attention required to maintain a given level of performance in the face of deficient vehicle characteristics.” The variables of training and experience again play a significant role in the assessment of control compensation.

Some closed-loop handling qualities criteria (closed-loop in the sense that they include a pilot model in their formulation) offer a sense of how control compensation plays a role in handling qualities (e.g., Neal and Smith, 1974 and Bailey and Bidlack, 1995), but objective and subjective measures for control compensation are neither readily available nor either validated or verified for handling qualities evaluations.

The subjective nature of handling qualities evaluations has fostered numerous endeavors to eliminate or reduce the subjectivity of pilot evaluations in defining handling qualities. While these intentions are good, the key isn’t so much eliminating the pilot as an evaluator, but precisely designing the test, conducting adequate training, and performing structured evaluations, as detailed in the following, to minimize the variability in the resultant evaluations. The test results which result from this procedure will accurately reflect the vehicle handling qualities – for better or worse.

## Rating Scale Definitions

Title	Descriptions
MENTAL DEMAND	How much mental and perceptual activity was required (e.g., thinking, deciding, calculating, remembering, looking, searching, etc.)? Was the task easy or demanding, simple or complex, exacting or forgiving?
PHYSICAL DEMAND	How much physical activity was required (e.g., pushing, pulling, turning, controlling, activating, etc.)? Was the task easy or demanding, slow or brisk, slack or strenuous, restful or laborious?
TEMPORAL DEMAND	How much time pressure did you feel due to the rate or pace at which the tasks or task elements occurred? Was the pace slow and leisurely or rapid and frantic?
PERFORMANCE	How successful do you think you were in accomplishing the goals of the task set by the experimenter (or yourself)? How satisfied were you with your performance in accomplishing these goals?
EFFORT	How hard did you have to work (mentally and physically) to accomplish your level of performance?
FRUSTRATION LEVEL	How insecure, discouraged, irritated, stressed and annoyed versus secure, gratified, content, relaxed and complacent did you feel during the task?



**Figure 4: TLX Rating Scale**



### **3. Handling Qualities Development**

The advent of digital computing and fly-by-wire flight control system technologies provided the capability whereby the handling qualities of aircraft could be tailored to the desires of the designer and essentially without regard of the vehicle's aerodynamics (stability and control). Despite this promise, handling qualities problems were rampant in these modern aircraft (National Research Council, 1997). An overarching lesson-learned from this history was that "more flight control system problems are caused by human behavior than for technical reasons" (Hodgkinson, 1990). These lessons-learned are reviewed to avoid re-living this history within the current and future US spacecraft development programs.

#### **3.1. Best Practices Process**

Fixed-wing and rotary-wing experiences in handling qualities developments are briefly reviewed to provide "best practices" guidance.

These best practices are reviewed in two areas:

- 1) for the design and execution of handling qualities simulation and flight test evaluations (Section 3.2); and,
- 2) for the overall design and development of an aircraft (or spacecraft's) handling qualities (Section 3.3).

Finally, in Section 3.4, an overview of industry standards for test and verification of handling qualities for fixed-and rotary-wing vehicles is provided. This overview is intended for perspective.

It should be noted that the test and verification process (Section 3.4) should be integrally linked to the best practices for the design and development phase. In fact, if a "best practices" approach to handling qualities is used in the aircraft (and spacecraft) design and development, the test and verification phase should be a "non-event." The flying qualities in test and verification should be thoroughly understood by the entire team. Any flying qualities deficiencies that do result in the test and evaluation phase would be due to factors, uncertainties, or inaccuracies in the actual flight hardware that could not be accurately simulated or tested before hand. However, these issues would have been identified and tracked as "risks" during the design and development phase and exploratory studies would have been conducted to assess the risk impacts. Thus, the risks and potential mitigations would be already identified and in place.

#### **3.2. Cooper-Harper Experience**

In the following, the best-practices for the conduct of handling qualities test and evaluations are summarized. The intent of these best practices is to: a) achieve reliable and comparable data among pilots; and b) provide accurate and sufficient handling qualities testing. These practices emphasize evaluations conducted during the design and

development phase, but they are generalizable to evaluations conducted throughout the design, test, and evaluation cycle.

The processes and definitions established by Cooper and Harper form the primary basis to achieve reliable and comparable data among pilots. The lessons-learned in conducting piloted handling qualities evaluations in accordance with these processes are reviewed to provide insight into these processes. New data and experience to support the work of Cooper and Harper are also provided.

### **3.2.1. Scale Definitions**

First and foremost, the Cooper-Harper Scale depends upon precise definitions of the words used. (Harper and Cooper, 1984) The definitions – essentially the B-side of the Cooper-Harper scale – should be used. These definitions should be reviewed during the pre-evaluation training and referenced during the course of the evaluations.

Best-practices would suggest that a review of the definitions is conducted and discussed during the pre-test briefing. Also, the *same* engineers/experimenters should conduct the pre-test briefing for each evaluation pilot to ensure that this discussion and the resulting interpretation of the definitions is consistent for all evaluation pilots.

### **3.2.2. Scale Usage**

How to use the scale must be understood by all evaluation pilots (EPs). While familiarity within the test pilot community may be assumed, the usage of the scale and the evaluation process should be repeated in pre-test briefings. The amount of training should be based on the evaluation pilot's familiarity and currency with the scale and the process to the point that all pilots are equally familiar.

In the use of the scale, it should be emphasized that the Cooper-Harper pilot rating is a shorthand notation which best represents the pilot's overall assessment of the evaluation. It reflects the pilot's summary opinion as to whether the handling qualities are satisfactory without improvement or if there are deficiencies which warrant or require improvement.

As noted by Cooper-Harper, "There tends to be some disagreement among pilots as to how they actually arrive at a specific numerical rating. Some pilots lean heavily on the specific rating description and look for the description that best fits their overall assessment. Other pilots prefer to make the dichotomous decisions sequentially, thereby arriving at a choice between two or three ratings." Experience has shown that using the dichotomous decision-tree reduces rating variability and provides consistency in the application of the scale.

In one attempt to reduce pilot rating variability, the decision tree process was enforced by using an interactive computer program which only displayed the parts of the Cooper-Harper scale – the decision tree - based upon the pilot's answers to the decision tree questions (Wilson and Riley, 1989). This feature enforced the dichotomous decision tree, but experienced EPs didn't necessarily appreciate this feature since they couldn't read the

entire scale to determine their rating. Since their rating is a shorthand notation to the words on the scale that best described their evaluation, they could not reflect across the scale ratings when only part of the scale was displayed.

For example, on one program (Monagan, Smith, and Bailey, 1981), evaluation pilots could achieve desired performance flying some aircraft configurations but they had extremely abrupt roll response dynamics to pilot control inputs. These configurations presented a quandary, in that, while desired performance could be achieved, their handling qualities characteristics were not necessarily desirable. The evaluation pilots had to adopt smooth and deliberate control actions to avoid unacceptable accelerations in the cockpit. These configurations exhibited pilot ratings that were rated either 4 or 7. This rating difference at first seems illogical, but the evaluation comments and ratings were consistent and truly indicative of the aircraft's handling qualities. To some EPs, the pilot could achieve desired performance and the control compensation was only moderate - the deficiencies warranted improvement. On the other hand, other pilots could achieve desired performance but the deficiencies were so objectionable that they required improvement; hence, ratings of 7 (control was not in question). For experienced evaluation pilots, the whole scale should be shown and used to give ratings.

This example obviously does not promote confidence in the repeatability and reliability of Cooper-Harper handling qualities data; but it does accurately reflect the closed-loop pilot-vehicle characteristics. "There will always be cases where different regions of aircraft characteristics will maximize an individual's performance and minimize their workload, due to a pilot's experience, training, and personal "tastes." (e.g., see Wilson and Riley, 1989). There are regions, nonetheless, that maximize performance and minimize workload for the vast majority of pilots. These rating differences typically occur at the boundaries between good and bad handling qualities (so-called "Level 2" configurations) or for aircraft that exhibit nonlinear or "cliff-like" characteristics.

Some have suggested that the scale itself has fundamental flaws, which promote pilot rating variability and uncertainty (Hoh, 1990; Riley and Wilson, 1990; Moorhouse, 1990). However, these so-called "flaws" do not out-weight the universal acceptance of the scale as written. Further, these issues are inconsequential compared to the bigger sources of pilot rating variability. As stated by Cooper and Harper, "precise definitions for the aircraft (or spacecraft) role and mission, the evaluation task, what the rating applies to, the simulation situation and the extent of pilot extrapolation are required." (Cooper and Harper, 1969). Without these factors being precisely defined, uncontrollable variability in handling qualities data will result.

### **3.2.3. Definition of the "Selected Task or Required Operation"**

The Cooper-Harper scale and the pilot evaluation reflect the adequacy of the vehicles' handling qualities for the "selected task or required operation." The selected task or required operations must be precisely defined. "The explicit description of the mission by delineation of the "required operations" is probably the most important contributor to the objectivity of the pilot evaluation data." (Cooper and Harper, 1969).

The complete task is composed of (1) the control task, and (2) auxiliary tasks (Cooper and Harper, 1969). “A task in the sense that it is used in handling qualities evaluations is defined as “the actual work assigned a pilot to be performed in completion of, or as representative of, a designated flight segment.” (Cooper and Harper, 1969).

The tasks should be defined by examining the actual mission context and then, with the existing evaluation tools, one should examine what can be done to assess that situation. “The tasks which the piloted airplane must perform, the weather (instrument, visual) and environmental conditions (day, night) which are expected to be encountered, the situation stressors (emergencies, upsets, combat), the disturbances (turbulence), distractions (secondary tasks), the sources of information available (displays, director guidance)—all these and more need to be considered. Secondary piloting tasks (voice communication, airplane and weapon system management) as well as primary tasks should be considered as they affect the attention available and total pilot workload.” (Harper & Cooper, 1984).

“The pilot must be given a clear description and understanding reached between the engineer and the evaluation pilot as to their interpretation of the required operations. This description must include:

- a. What the pilot is required to accomplish with the aircraft, and
- b. The conditions or circumstances under which the mission is to be conducted.”  
Cooper and Harper, 1969

#### **3.2.4. Evaluation Situation/Extrapolation**

In virtually every context, simulation or flight testing is used in an attempt to understand or forecast the handling qualities of the actual vehicle during actual operation. Since the engineer (and pilot) wants to get this assessment, there is a temptation to “extrapolate” the results to the “real-world.”

As discussed by Harper and Cooper (1984), “Some would have the pilot assess only the simulated operation; others would have him use the simulation results to predict/extrapolate to the real world operation.” The issue must be discussed and addressed a priori; otherwise, different pilots may produce different results.” As a minimum, the idea of extrapolation needs to be agreed upon.

For pilot rating repeatability, experience shows that “the best extrapolation is no extrapolation.” If you ask the pilot to extrapolate their ratings to other situations and circumstances, a significant degree of variability and uncontrollability is introduced.

However, as also discussed by Harper and Cooper (1984), “An important aspect to this question of extrapolation is: if the pilot doesn’t do it, who will? And what are his credentials for doing so? Some differences (especially simulator deficiencies) would seem to be primarily left to the engineer to unravel (perhaps with the aid of a test pilot), for it is difficult for the evaluation pilot to fully assess the effects, for example, for missing motion cues or time delays in the visual scene. But when the simulation tasks do not include all of the real situation, one would perhaps rather depend upon the pilot to

assess what he sees in the simulator in the light of his experience in the real-world tasks.” (Harper & Cooper, 1984).

Extrapolation is usually an issue when evaluating particular combinations of failure and operating conditions. “Such questions of probability of occurrence and levels of disturbances must be resolved as part of the mission description in the design of the experiment, with special attention often being required with respect to pilot orientation and the reporting of results.” (Cooper and Harper, 1969). Again, in the pre-brief, the evaluation pilot should be made aware that some combinations of failures and operating conditions may be more or less likely to occur.

For pilot rating repeatability, the degree of extrapolation would be discussed and agreed upon – with consistency across all EPs. The best extrapolation is no extrapolation – the pilot should rate the configurations and scenarios strictly in the context for which they were flown. The “extrapolation” based on the probability of occurrence can be handled in the post-test de-brief and by engineering analysis. In some cases, relaxed task performance standards have been used for contingency or low-probability of occurrence scenarios to reflect that the configurations might be a “last-ditch” effort to save the aircraft or mission. “... “operating problems” consisting of combination of failure or weather conditions are apt to raise the question of probability of occurrence. Again, this question should be separated from the pilot assessment wherever possible.” (Cooper and Harper, 1969).

### **3.2.5. Specification of Performance Standards**

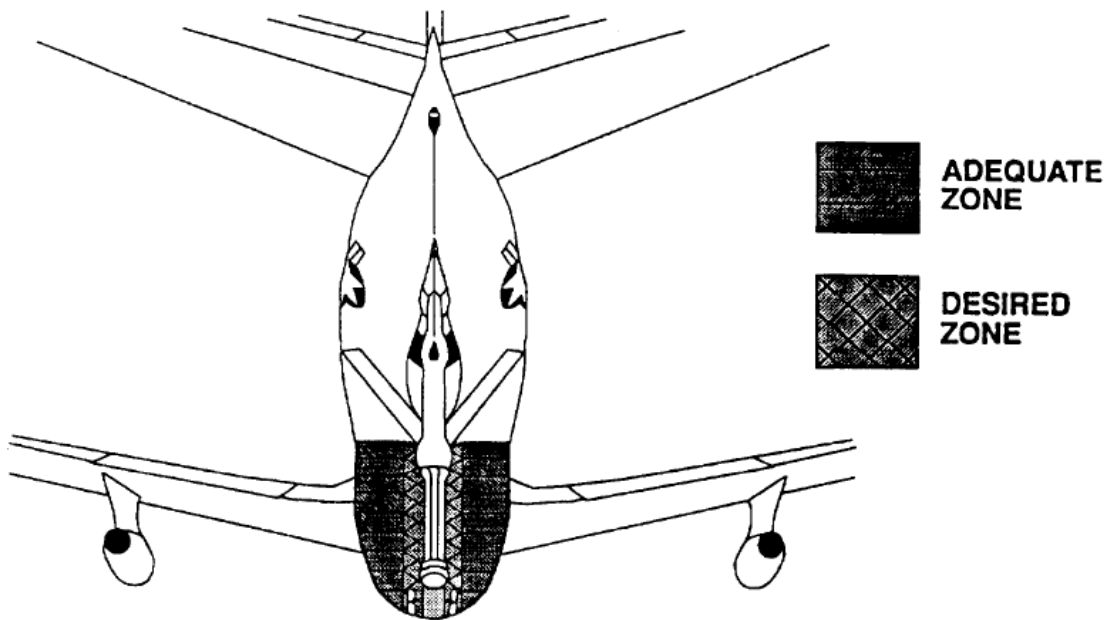
In addition to explicitly defining the task or required operation, precise definitions of task performance standards must be established.

These task performance standards must be: a) germane to the required operation or task as they apply to the actual mission; b) include variables or outcomes controllable by the pilot; c) observable to the pilot; and, d) sufficiently demanding that high closed-loop pilot-vehicle “task bandwidth” is required to aptly stress and test the handling qualities characteristics.

In many cases, task performance standards naturally flow from the performance demanded for the actual mission. For instance, handling qualities testing for aircraft landing task naturally use the touchdown point as a vital part of the performance standard. The results are clearly observable to the pilot and under their control.

On the other hand, sometimes indirect standards, germane to the mission, are necessary. For instance, in boom tanker refueling, the mission is to off-load fuel from the tanker to the receiver aircraft. The handling qualities performance standards for the receiver in this task do not directly reflect the mission, but assess the ability of the receiver to move into and then maintain the contact position. The rest of the mission – the ability of the boom operator to make contact and off-load fuel – is not part of the handling qualities test. The performance standards have involved using visually observable and controllable parameters for the pilot of the receiver aircraft (see Figure 5 from Leggett and Cord,

1994). The angular position of tanker lights and markings make it possible for the pilot of the receiver aircraft to control and observe performance in real-time.



**Figure 5: Desired and Adequate Performance During Boom Aerial Refueling (from Leggett and Cord, 1994)**

The term “task bandwidth” qualifies the extent that the task demands stress the pilot and test the closed-loop handling qualities. Task bandwidth is determined by: 1) the precision that is demanded of the pilot-vehicle performance; and, 2) the time to complete the task. A time constraint may be naturally occurring for certain tasks. If it is not naturally part of a task, it should be imposed. For instance, in a landing task, the time constraint is established by the approach speed and winds. For a refueling task, a closure rate should be established which dictates how quickly the vehicle must move from the pre-contact position into a refueling position. Without stipulating the closure rate, pilots when flying aircraft with handling qualities deficiencies may stop or slow the closure so they have more time to keep the aircraft under control.

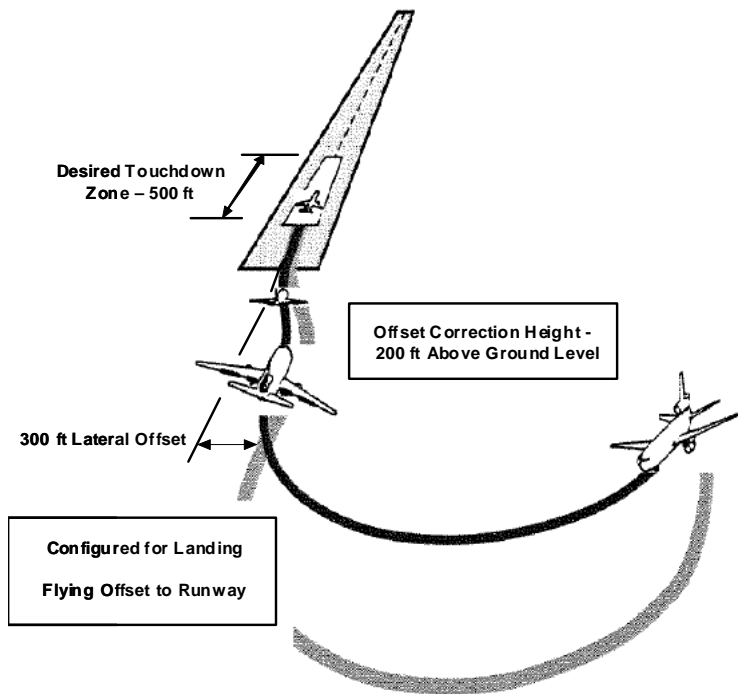
Task bandwidth is also modulated by the task definition. Introducing positional offsets in a task requires that the pilot null these position errors by actively entering the control loop. How much time is afforded the pilot to do the task, establishes the bandwidth. For instance, an offset landing task (Figure 6) is often used. The offset landing is the nominal mission task, but the offset ensures that the pilot actively enters the control loop, requiring pitch and roll inputs, to correctly align and land the aircraft. The idea of standardized task performance standards has emerged. The concept is that, for aircraft to meet a particular mission, the handling qualities task performance standards should be invariant. This work has taken many forms, from the Mission Task Elements which are a vital part of the Aeronautical Design Standard (ADS)-33, which is the tri-service rotary wing handling qualities standard – to the “Standard Evaluation Maneuver Set” (Leggett and Cord, 1994) which serves as a verification basis within MIL-HDBK-1797 - the

military flying qualities handbook for fixed-wing aircraft. The Joint Strike Fighter program developed task performance standards for their prototype fly-off and are using them for the verification and validation in the production vehicle.

Finally, the task performance standards – for completeness sake – should include a stipulation against undesirable characteristics so that pilot-induced oscillations (PIO), for example, cannot be considered “desired performance” (see Table I as an example of the offset landing task performance standards). Pilot-induced oscillations are instabilities in the closed-loop pilot-vehicle dynamic system, sometimes referred to as airplane-pilot coupling. Explicitly stating “No PIO” within the desired performance box in Table I, achieves two goals: 1) it continually reminds the pilots that PIO is undesirable and they should be searching for its presence and its negative connotations; and, 2) it prevents there from being any doubt that IF a pilot were to be so lucky as to be in a PIO which just so happens be within the desired performance touchdown zone, that this behavior is still not desirable. (In fact, it could be argued that this type of performance is not adequate performance either.)

**Table I: Offset Landing Task Performance Standards**

Desired Performance	Desired Performance	Adequate Performance
Longitudinal Touchdown	1000 to 1500 ft past threshold	750 to 2250 ft past threshold
Lateral Touchdown	+/- 10 ft of centerline	+/- 27 ft of centerline
Sink Rate	< 4 ft per sec	< 7 ft per sec
Other	No PIO	



**Figure 6: Lateral Offset Landing Task**

### **3.2.6. Task Performance**

At the completion of the task, the task performance should be reported to the pilot. This procedure can reduce variability in the Cooper-Harper Pilot Rating (CHPR) process; however, it is not necessarily eliminated. As per the scale definitions, desired performance is a necessary but not necessarily sufficient condition for Level 1 ratings (CHPRs of 1, 2 and 3). Adequate performance is a necessary but not necessarily sufficient condition for Level 2 ratings (CHPRs of 5 and 6).

Particularly during a verification and validation phase, pilot rating variability can be problematic. The definition and structure of the evaluation task can be used to reduce structural causes for evaluation differences.

First, the task initial conditions establish the start of the evaluation and the task performance standards typically dictate the outcome or “end condition.” But task performance variability (and consequently, pilot rating variability) is induced by differences in how the evaluation pilots flies the task between these established initial and end-conditions.

For instance, in evaluating landing handling qualities for fixed-wing aircraft, an offset landing task is often used. The task requires the pilot to fly parallel to the nominal approach but laterally offset (nominally 250 ft) from the runway centerline. Upon reaching a predetermined altitude above the runway, the pilot corrects the offset by performing a sidestep maneuver and completes the landing. The lateral offset task forces the pilot to actively control the vehicle and does not allow use of “open-loop”-type maneuvers which wouldn’t adequately stress the aircraft’s handling characteristics. The sidestep acts as a “disturbance” function. How quickly and aggressively each pilot performs the side-step maneuver can significantly influence the evaluation. Gradual sidestep corrections can keep flying qualities problems from appearing, whereas the same or a different pilot who might more quickly re-acquire the runway centerline and excite a “lurking” deficiency.

For handling qualities “exploration and discovery” (see Section 3.3), structure in how to fly a task is not necessarily desired. To search for lurking handling qualities problems or “cliffs,” the engineering team wants the pilot to explore different control techniques and different ways to fly a particular task. The results of this exercise may generate pilot rating variability, but this type of data is invaluable in understanding a configuration.

On the other hand, this non-uniformity may frustrate a statistically-driven design validation and verification process. If ratings “conformity” is desired, all aspects of flying the task should be briefed and the test structured – somehow – to enforce uniformity in how the task is flown. Pilot briefings should be used but a briefing alone is not typically sufficient. Test methods to control how the task is to be flown are the best method if rating variability is to be minimized. Pilots should be briefed what rates and accelerations are to be used in the task execution.



Second, evaluation pilots should be briefed to always strive to achieve desired performance and use compensation and control techniques to do so. Using this philosophy promotes rating consistency. On the other hand, if “exploration and discovery of a vehicle’s handling qualities” is desired, pilots may and should be encouraged to use different control compensation and techniques to try to desired performance or adequate performance, when desired performance does not appear achievable. (Adequate performance is not a disaster. Adequate performance means that the mission can still be completed, albeit with a slight pilot workload and/or compensation penalty and reduced performance margins.) By “relaxing” the standard of performance to adequate, the pilot is typically relaxing how tightly they are controlling the vehicle. As such, they may not excite deficient handling qualities and end up, possibly, with desired task performance. Experienced evaluation pilots will recognize this behavior – i.e., the need to stay “out of the loop” to avoid exciting deficiencies - as control compensation and workload. Their ratings and comments for this configuration will track this behavior. But because the training and experience of the pilot with these types of configurations is critical (i.e., to recognize the compensation and their subconscious change in task “objective”), this behavior may contribute to inter-pilot rating variability if the training and experience of the evaluation pilots differ.

Lastly, a test technique called Handling Qualities During Tracking (HQDT) was championed whereby the pilot attempts to eliminate any error in the performance of the task (Twisdale and Franklin, 1975). The objective was to create a "stress test" of handling qualities. This type of test is not truly a handling qualities test – in fact, adequate and desired performance standards are not supposed to be defined for HQDT and Cooper-Harper ratings are not recommended (Leggett & Cord, 1994). The HQDT technique demands that the pilot strive for “perfection,” using near limit-cycle inputs, without the pilot trying to compensate for aircraft deficiencies. This task and striving for “perfection” should not be used as the driving task requirement for a true handling qualities evaluation. But this type of technique may be part of the handling qualities “exploration” process. If used, the fact that the HQDT is not a handling qualities test should be emphasized and “abusive” and unrealistic pilot inputs should be avoided.

### **3.2.7. Use of Pre-Test Evaluation Pilots**

It is critical that the engineering team use a subject matter expert (SME) pilot in a pre-test phase to assess and develop the briefings, the tasks, and the task performance standards.

This pilot should review and critique all aspects of the test. The SME should explore the handling qualities behaviors of the planned configurations, tasks, and rating techniques and scales.

This pilot should focus on the test set-up as well as providing preliminary data for the engineering team who are developing the test. “It is generally recommended to use only a few pilots (sometimes only one) until the experiment has matured through the engineer’s understanding of the comment and rating data. His task of sorting out, organizing, and digesting the comments and ratings to understand the pilot-airplane system is complex and often frustrating. By working closely with one or a few pilots

initially, the engineer can often acquire this understanding sooner.” (Harper & Cooper, 1984).

The pre-test developmental pilots, once their work is complete, should not be used in the “formal” data collection process.

### **3.2.8. Evaluation Pilot Selection**

The first rule of thumb is that “the evaluation pilots represent the user population and should be experienced in the required operation” (Harper & Cooper, 1984). Since the evaluations are subjective in nature, the background, training, experience, and point of view of the evaluation pilots must be considered in their selection based on the program objective as well as by the intended use of the resultant data (Cooper and Harper, 1969).

The effects of differing skills should be determined from the results obtained from evaluation pilots of different, but representative, levels of experience and training. Exceptions to this general rule have occurred, however, when the research or development test pilot is asked to evaluate handling qualities with respect to his understanding of the lowest degree of skill and training existent in a group of operational pilots” (Cooper and Harper, 1969).

### **3.2.9. Number of Evaluation Pilots**

Once the pre-test is completed, the number of pilots to use for the test matrix is always a tradeoff between pilot schedule/availability and the cost to run the simulation or flight test. Ideally, you’d like the sample size to be as large as possible to obtain statistical significance, but this is typically impractical. Riley and Wilson (1990) suggested a flowchart to determine the number of evaluation pilots required as a balance between quality and cost. Practically speaking, the number of highly qualified test pilots required to attain statistical significance using sophisticated simulations, extensive pre-test training, and an ordinal scale (Bailey, 1990) is much greater than what most programs can afford. On the other hand, the best approach might be that “if you want statistical significance, measure big differences.”

Instead of focusing on statistical significance, fewer evaluation pilots often produce better results. As told by Harper and Cooper (1984), “A classic handling qualities experiment (Kidd and Bull, 1963) showed that a few pilots evaluating for a longer period of time produced the same central tendency of the rating excursions as a larger group conducting shorter evaluations. What was lost with the larger group, however, was the quality, consistency, and meaningfulness of the pilot comment data.”

### **3.2.10. Experiment Design**

While statistical significance in pilot ratings evaluations may not be realizable - given limited time and money - good experimental design practices are still important (e.g., see Kirk, 1982; Dukes, 1985). The configuration run matrix should be properly balanced to avoid learning effects but it should also be tailored as possible and practical to create periodic "re-calibration" of the evaluation pilots as to "good" and "bad" flying qualities

(Bailey, 1990). This offers the opportunity to maintain an "absolute" rather than relative sense of the rating scale and the possible range of flying qualities characteristics for the task.

### **3.2.11. Blind Evaluations**

Experience has shown, and it is recommended (Harper and Cooper, 1984) that handling qualities evaluations should be conducted “blind” in that the pilots do not know the engineering details of the configuration. This approach is not without detractors and dissenters, but it typically renders the evaluations free from unintended influences.

Blind evaluations do not imply that the pilots do not know information that is critical to their evaluation or that the cockpit displays are changed or masked to create this anonymity. The intent of withholding the details of a configuration is to provide a “clean-slate” for each evaluation. If each configuration is evaluated in this way, the handling qualities influence will truly emerge, rather than pre-disposing an evaluation. For the purposes of evaluation, the evaluation pilot should observe and report the handling qualities of the configuration. It is not necessarily constructive to bias them or have them try to “dig into” a configuration just to reveal the influence of an announced engineering change.

During this process, encouragement should be given to the evaluation pilot to reassure them that their evaluations are not “in left field.” Otherwise, it can somewhat disconcerting to perform continual blind evaluations.

Upon completion of the evaluations, the details of the configuration and the evaluation results should be revealed and discussed. Additional “exploration” if warranted should be done at this time to tease out additional comments and engineering ideas and concepts. This is also useful for the evaluation pilot’s knowledge and training.

### **3.2.12. Repeat Evaluations**

Repeat evaluations should be conducted to test for evaluation consistency and training effects. The repeated evaluations should be conducted just as the others – in the blind.

Roughly 10 to 20% of the matrix is typically repeated. Often these evaluations are selected as the most “interesting” configurations. But experimental design and “re-calibration” considerations (see Section 3.2.10) also are entered into the selection process for repeats.

### **3.2.13. Length of Evaluations**

Just as in choosing the number of evaluation pilots, the same trade-off exists between pilot schedule/availability and the cost to run the simulation or flight test in determining the length of the evaluations.

The typical protocol is to allow a practice run with the evaluation configuration, followed by two runs with the configuration for the record. A third run for data can be flown by

the evaluation pilot, at the evaluation pilot's request, if there was a discrepancy in the evaluation in the two evaluations. The third run is the "tie-breaker."

#### **3.2.14. Pre-Test Familiarization**

To achieve consistent handling qualities data, each evaluation pilot should be given training in the handling qualities and Cooper-Harper rating process sufficient to bring each participant to a common level. This is particularly critical for new or untrained evaluation pilots.

Part of the training for new evaluation pilots is indoctrination into the fact that handling qualities evaluations are an "observational process." Too often, inexperienced evaluators think that performance problems with the pilot-vehicle dynamics are their fault – not the result of deficient handling qualities. They must be trained to understand that: a) they are serving as a pilot and the resultant pilot-vehicle dynamics are the result of, not the fault of, the pilot. They should observe and report on the resultant dynamics and characteristics.

For new evaluation pilots, the concepts of control compensation and workload, to a lesser extent, are new and somewhat difficult to fully grasp and accurately report. Hands-on training must be used to ingrain these concepts.

As noted by Harper and Cooper, 1984, "it is helpful to allow the pilot to evaluate handling qualities that span the range of the rating scale; that is, let him see good, bad, and in-between characteristics. This is perhaps less important with experienced evaluation pilots, but it can be an important factor with operational pilots whose experience is confined to one or two airplanes."

In addition, the pre-evaluation phase should be used to re-emphasize the briefing and training for the handling qualities evaluation process. In particular, this phase should provide hands-on training for:

1. The selected task or required operation that the EP will evaluate.
2. The task performance standards.
3. The Cooper-Harper rating scale definitions and decision tree usage.
4. The pilot comment card and associated rating scales.
5. The manner in which the task should be performed and the degree of latitude that the EPs have in the exploration of handling qualities.
6. The fact that "the pilot rating is given for a configuration in the context of its acceptability to the pilot for the specified flight phase (or task) and not in terms of its goodness with respect to a configuration already evaluated." (Cooper and Harper, 1969). Therefore, they should not be "reluctant to rate something as excellent or optimum for fear that a subsequent configuration

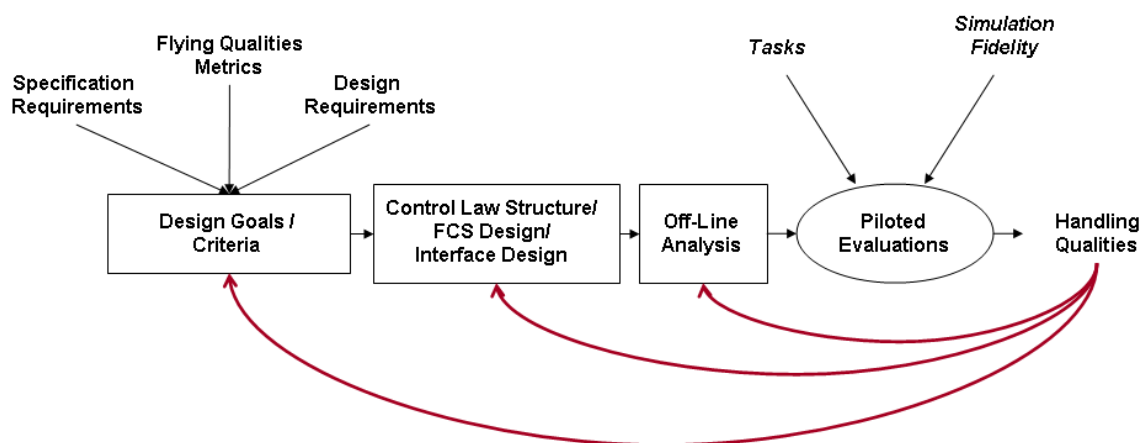
will be better than anything they considered possible.” (Cooper and Harper, 1969). Pilots should be encouraged to use the entire rating scale including ratings of 1.

7. The testing procedure including: 1) the amount of information that the EP will receive regarding the configurations – even though, the configurations will be evaluated “in the blind;” 2) the number of runs to be used to formulate an evaluation; and, 3) their option to do additional runs should variability in the closed-loop characteristics or performance occur, thereby clouding their evaluation.

### 3.3. Developmental Test & Evaluation

Handling qualities evaluations are a critical component in the successful development of a new vehicle. In the following, the lessons-learned in how to best develop fixed-wing and rotary-wing handling qualities and the use of handling qualities evaluations in this process are discussed.

The ideal handling qualities development process is notionally shown in Figure 7. This process captures the historical best practices for handling qualities development (Harper and Cooper, 1984; Bailey, 1990; National Research Council, 1997; NATO, 2000).



**Figure 7: Handling Qualities Development Process**

First and foremost, history has shown that when the design considers handling qualities up-front and as an integral, critical part of the program, significantly less time and money are spent overall on handling qualities development than a comparable program where this up-front emphasis was not taken (e.g., Hodgkinson, 1990; NATO, 2000). Fixing problems late in the design cycle are significantly more expensive (and typically less effective) than those fixed early in the design process. The process involves the following constructs:

- Handling qualities design goals and criteria are established as an integral part of the program. These goals are based on design requirements, applicable specifications, and existing flying qualities metrics and past experience.
- From these goals, the control law structure, flight control system (FCS) design and interfaces are established.
- Off-line analysis are integral to understanding the handling qualities design and the technical issues that drive this design, and providing a basis from which to interpret the handling qualities evaluations.
- Handling qualities evaluations form the principle means by which the design is assessed and improved. To be effective, the evaluations must be fed-back into the design process to adjust the design goals, the FCS and interface design, and complement the off-line analyses. In particular, the results of the handling qualities evaluations must be weighted and analyzed according to the selected tasks and the simulation fidelity used in the evaluations.
- In this closed-loop process, the level of simulation fidelity and maturity of the control law implementation and vehicle dynamics and the associated analytical toolboxes should grow together as the design process matures. Testing results from actual hardware should be included as soon as possible and iron-bird, hardware-in-the-loop, with flight-fidelity interfaces (controllers and displays) should be planned and “flown” as early as possible.
- The goal should be that, just prior to flight testing, very few differences, if any, will exist between the simulations and the actual flight test. The only differences that should exist may be motion-cues (i.e., it is extremely difficult to simulate the complete motion environment of a new aircraft, even with sophisticated ground-based or in-flight simulators) and aero-dynamics parameter variations that were outside of the simulation math models.

For the design of rotary-wing and fixed-wing aircraft, an enormous body of time-tested FCS design criteria, experience, handling qualities criteria, and metrics are available. These works significantly simplify and improve the aircraft handling qualities design process. This same wealth of available data is, unfortunately, not available for spacecraft handling qualities due to the four or five orders of magnitude difference in the number of manned spacecraft developed versus the number of manned aircraft.

Without this strong legacy data, the design process and the use of handling qualities will be especially critical for future spacecraft developments. An overarching lesson-learned from the aircraft handling qualities history is that “more flight control system problems are caused by human behavior than for technical reasons” (Hodgkinson, 1990).

The following best-practices (NATO, 2000) are presented as they are directly relevant to the spacecraft handling qualities developments. These are not meant to be comprehensive, but are “take-aways” appropriate for spacecraft:

1. Understand the operational requirements and the piloting task in each phase of the mission. Ensure good communications with pilots is maintained in order to be fully aware of operational conditions.
2. Avoid over-complexity and aim to keep the FCS design as simple and as visible as possible.
3. Beware of control systems which appear to achieve excellent performance, mainly by open-loop compensation of the nominal model. Such performance can deteriorate very rapidly when modeling tolerances are introduced or when external disturbances are applied. Such effects can be corrected by improving the closed-loop performance of the system, usually by increasing the feedback gains – although this is not always possible.
4. Plan for an integrated simulation program and ensure that all team- members (especially pilots and managers) are clear that the various simulators are for evaluation purposes, to feed data back into the analytical design process.
5. Identify the limitations of the simulation, including consideration of providing motion cues. Be aware that although simulators are of great value if used correctly, they can give misleading results if the assessments are not rigorously controlled. Simulation validation is highly desirable, if not essential.
6. Use the piloted simulator to complement the off-line design and development tools, and to intercept any design deficiencies at an early stage. The earlier problems are detected, the less it costs to fix them.
7. Use common code and data for off-line and piloted simulation to avoid unnecessary software maintenance or translation (time and cost) and the possible introduction of errors in control law functionality. Provide adequate off-line check cases to verify the control law implementation on the simulator.
8. Simulation displays and controls need to be representative, in order to avoid coloring pilot opinion of the control laws.
9. It is desirable that pilots are ‘calibrated’ in the use of development simulators, to aid their judgment of the simulated aircraft’s handling characteristics. One way of achieving such calibration is to allow them to familiarize themselves with the simulator, by flying an aircraft with which they have flying experience.
10. Deliberately search for handling problems, including the effects of design tolerances (parameter uncertainties) and failures. Identify the worst cases and any hidden weaknesses in the design, and fully explain any unexpected simulation results.

11. Evaluate the ability of the pilot to enter the control loop, to help out the automatic functions. Show that there is no tendency for divergence between the automatic and manual control functions.
12. An Integrated Product Team for flight controls/flying qualities should be formed, covering all the skill areas required to develop a flight control system. This team should be responsible for tracking the design, development, and test of each component, and the implementation and verification of each interface.
13. Deliberations should be encouraged, to bring into public view any problem or area of concern, so that all attendees can assess possible interactions with their area of responsibility, or where appropriate, potential solutions to “system” problems which may involve components other than those which encountered the anomaly. Stress should be placed on including all components and interfaces in the discussions, since a system problem can be generated by a component that is performing well within the performance boundaries specified for it as a unit.

Of these lessons-learned, the one that creates the most programmatic consternation is Number 10 – to deliberately search for handling problems. This process is typically cited as one of the greatest failings for a program. The reason is that it seems counter-intuitive to management. Instead of the team working to build a good vehicle, it would seem this step is trying to find reasons for failure. In fact, it is just the opposite. By diligently searching for problems, including the effects of design tolerances (i.e., parametric uncertainties) and failures, testing worst case scenarios and searching for hidden weaknesses, the team is proactively heading off problems that might otherwise emerge late in the design. This period of “exploration and discovery” must be conducted. If properly done, this work makes subsequent test and verification phase a “non-event.” The team will have tested the system to its limits and fully understands the physics of the problem and its associated uncertainty levels to the extent that the test and verification will almost be an afterthought.

### **3.4. Test and Verification**

Test and verification for flying qualities has been an issue in aircraft procurement from the very beginnings of aviation. Handling qualities “acceptance test” was, in fact, a clause in the very first aircraft procurement – the Advertisement And Specification For A Heavier-Than-Air Flying Machine (Army Signal Corps, Specification No. 486, dated December 1907) - which specified that:

“Before acceptance, a trial endurance flight will be required of at least one hour during which time the flying machine must remain continuously in the air without landing. It shall return to the starting point and land without any damage that would prevent it immediately starting upon another flight. During this trial flight of one hour, it must be steered in all directions without difficulty and at all time under perfect control and equilibrium.”



From this simple specification, test and verification requirements for handling qualities have significantly expanded. Nonetheless, the essence of this first requirement is retained – that the vehicle must be “steered in all directions without difficulty and at all time under perfect control and equilibrium.”

A survey of handling qualities practices is given in the following for commercial and military aircraft requirements for test and verification of handling qualities.

### **3.4.1. US Military Fixed-Wing Aircraft**

The first comprehensive military handling qualities specifications were issued by the Navy Bureau of Aeronautics in 1942 and by the U.S. Army Air Force (AAF-C-1815) in 1943 (Harper and Cooper, 1984). AAF-C-1815 gave way in 1954 to a new version, MIL-F-8785. More importantly, a subsequent version, MIL-F-8785B, began the precedence within the handling qualities community that the true value in a specification document was not the detailed requirements, but an elaborate background information and users guide (BIUG) wherein the data which form the specification are contained. The BIUG forms the historical lessons-learned for handling qualities which provide a continual improvement process for vehicle handling qualities.

The use of military specifications fell out of favor in the 1980s. The last in this series was MIL-F-8785C issued in 1980 (see DoD, 1980).

MIL-F-8785C was re-worked and updated into a military standard - MIL-STD-1797A in 1995 - and this document was re-designated in 1997 as a handbook MIL-HDBK-1797A to be used for procurement guidance purposes (DoD, 1997). Like MIL-F-8785B before it, the MIL-HDBK incorporates exhaustive BIUG material as its basis.

Under MIL-HDBK-1797A, test and verification of handling qualities are stipulated using the following levels of flying qualities definitions:

- Level 1 (Satisfactory): Flying qualities clearly adequate for the mission flight phase. Desired performance is achievable with no more than minimal pilot compensation.
- Level 2 (Acceptable): Flying qualities adequate to accomplish the mission flight phase, but some increase in pilot workload or degradation in mission effectiveness, or both, exists.
- Level 3 (Controllable): Flying qualities such that the aircraft can be controlled in the context of the mission flight phase, even though pilot workload is excessive or mission effectiveness is inadequate, or both. The pilot can transition from Category A flight phase tasks (i.e., those tasks that rapid maneuvering, precision tracking, precise flight-path control) to Category B or C flight phases and these tasks (i.e., non-terminal and terminal flight phases such as cruise, approach and landing) can be completed. It is to be noted that Level 3 is not necessarily defined as safe.

The handling qualities requirements depend upon whether the vehicle is operating in a normal or failure state:

- For aircraft normal states, the minimum handling qualities requirement is Level 1 within the operational flight envelope, and Level 2 within the service flight envelope.
- No single failure of any component or system shall result in dangerous or intolerable flying qualities.
- Otherwise, for airplane failure states, the probability of encountering Level 2 handling qualities must be less than 10<sup>-2</sup> per flight within the operational flight envelope and the probability of encountering Level 3 handling qualities must be less than 10<sup>-4</sup> per flight within the operational flight envelope, and less than 10<sup>-2</sup> per flight within the service flight envelope. (The Service Flight Envelope encompasses the Operational Flight Envelope with sufficient margins established between.)

Verification is conducted in demonstration tasks using the resultant pilot comments and Cooper-Harper ratings as follows to define handling qualities levels:

- For Level 1, the pilot comments must indicate satisfaction with aircraft flying qualities, with no worse than "mildly unpleasant" deficiencies, and median Cooper-Harper ratings must be no worse than 3.5 in calm air or in light atmospheric disturbances.
- For Level 2, the pilot comments must indicate that whatever deficiencies may exist, aircraft flying qualities are still acceptable, and median Cooper-Harper ratings must be no worse than 6.5 in calm air or light atmospheric disturbances.
- For Level 3, the pilot comments must indicate that the aircraft is at least controllable despite any flying qualities deficiencies, and median Cooper-Harper ratings must be no worse than 9.5 in calm air or light atmospheric disturbances.

This standard shows that the required flying qualities level depends upon: 1) the task; 2) the presence or absence of failures; 3) if the aircraft is flying within its nominal flight envelope; and, 4) the atmospheric conditions.

The test and verification process under MIL-HDBK-1797 recommends three to six pilots per test condition. Careful selection of the evaluation pilots is also recommended to reduce the variability in results. All of the evaluation pilots must be test pilots trained in the use of the Cooper-Harper scale and they all must be experienced in the class of aircraft under evaluation.

### **3.4.2. Joint Strike Fighter (JSF)**

To illustrate the application of MIL-HDBK-1797 concepts, the Joint Strike Fighter (F-35) aircraft requirements are reviewed. (The following material is based on recent personal communications with Mr. James “Buddy” Denham, Aeromechanics Senior Engineer at Naval Air Systems Command in Patuxent River, MD.)

For normal operations of the F-35 aircraft, Level 1 handling qualities are required. For any single failure or combination of failures with a probability of occurrence greater than  $10^{-7}$  per flight hour, the vehicle shall be capable of:

- Aerial refueling and landing at the original/alternate destination with Level 2 (or better) handling qualities.
- Cruise and descent with Level 3 (or better) handling qualities.
- Terminating precision tracking or maneuvering tasks with Level 3 (or better) handling qualities.

Handling qualities verification testing is conducted by a joint contractor/military pilot team who fly each task and provide Cooper-Harper ratings and comments. Based on these combined Cooper-Harper ratings and comments, the military determines if the required handling qualities Level has been achieved; if outlier(s) exist, the comments should indicate if the problem was with the vehicle, the task workload, or unique to the particular evaluation.

### **3.4.3. US Army Rotorcraft**

Helicopter flying qualities requirements were first established in MIL-H-8501 which was issued in 1952. The U.S. Army began development of a handling qualities specification in 1982, and its initial Aeronautical Design Standard-33 (ADS-33A) was published in 1987. The current version is ADS-33E-PRF; see U.S. Army, 2000.

Rotorcraft flying tasks are called Mission Task Elements (MTEs). There are 23 MTEs such as: slalom, sidestep, hover, and landing.

Handling qualities requirements for rotorcraft normal states mirror the fixed-wing standards:

- The minimum Levels of flying qualities shall be Level 1 in the Operational Flight Envelopes and Level 2 in the Service Flight Envelopes.
- Under failure conditions, the probability of encountering Level 2 handling qualities must be less than  $2.5 \times 10^{-3}$  per flight hour within the operational flight envelope. The probability of encountering Level 3 handling qualities must be less than  $2.5 \times 10^{-5}$  per flight hour within the operational flight envelope, and less than  $2.5 \times 10^{-3}$  per flight hour within the service flight envelope

Under ADS-33, handling qualities verification is conducted by piloted evaluation where the Cooper-Harper rating scale assesses the workload and task performance required to perform the designated MTEs.

Each MTE is assessed by at least three pilots. These pilots shall each assign a subjective rating using the Cooper-Harper rating scale. The arithmetic mean across all pilots of the Cooper-Harper Handling Qualities Ratings (HQRs) forms the overall rating for the MTE. Level 1 is defined as the average HQR  $\leq 3.5$ , Level 2 is defined as the averaged HQR  $\leq 6.5$ , and Level 3 is defined as an average HQR between 6.6 and 8.5.

To meet Level 1 handling qualities requirements, the rotorcraft shall be rated Level 1 for all of the MTEs designated as appropriate to the rotorcraft's operational requirements.

#### **3.4.4. Federal Aviation Administration**

Unlike military specifications which serve as procurement documents, the Federal Aviation Administration (FAA) conducts certifications of air vehicles against Federal Aviation Regulations (FARs) which establish whether they are safe to operate.

The introduction of fly-by-wire flight control system technology in commercial transport aircraft necessitated that the FAA establish a systematic methodology to conduct certification flight testing for handling qualities (McElroy, 1988). The unique handling characteristics, systems complexity, and failure modes effects rendered many of the previous and existing regulations and test techniques useless or inappropriate. The FAA developed flight test certification procedures which follow, in many respects, the major elements of existing handling qualities evaluation methodologies, tailored to the civil application (Advisory Circular 25-7A, FAA, 1998). Particular attention is focused within these techniques for the exploration of deficient handling qualities in the form of pilot-induced oscillations (PIO) or Aircraft-Pilot Coupling (APC).

Roughly following the Cooper-Harper decision tree, three levels of handling qualities are defined as shown in Figure 8. The FAA levels of SAT, ADQ, and CON are given equivalents to Cooper-Harper pilot ratings and military standards (MIL-STD-1797 and MIL-F-8785).

FAA HQ RATING	FAA DEFINITION	COMPARISON		
		CH*	MIL STANDARD	
			LEVEL	QUAL
Satisfactory	Full performance criteria met with routine pilot effort and attention	1-3.5	1	SAT
Adequate	Adequate for continued safe flight and landing: full or specified reduced performance met, but with heightened pilot effort and attention.	3.5-6.5	2	ACCEPT
Controllable	Inadequate for continued safe flight and landing, but controllable for return to safe flight condition, a safe flight envelope, and/or reconfiguration so that HQ are at least ADEQUATE.	6.5-8.0	3	CON

Note: \*Cooper-Harper Rating

**Figure 8: FAA Definitions Compared to Cooper-Harper Scale**

Three categories of general handling qualities tasks are flown:

- Trim and unattended operation (e.g., dynamic response to pulse input)
- Large amplitude maneuvering (e.g., pitch/roll upset recover)
- Closed-loop precision regulation of flight path (e.g., ILS and precision touchdown)

The minimum acceptable level of handling qualities (FAA, 1998) depends upon combinations of three factors:

- Atmospheric Disturbance Level (Light, Moderate, or Severe)
- Flight Envelope: (Normal Flight Envelope (NFE), Operational Flight Envelope (OFE), or Limit Flight Envelope (LFE) )
- Flight Control System Failure State

Note that the FAA uses three flight envelope definitions (NFE, OFE, and LFE) as opposed to the military's usage two (i.e., the OFE and SFE) where the NFE is associated with routine operations, the OFE is outside of the NFE and associated with warning

onset, and the LFE is the most outside flight envelope associated with aircraft design limits.

The fundamental concept is that handling qualities should be evaluated and found acceptable for all conditions and tasks for which their probability of occurrence is not extremely remote (probability of occurrence less than  $10^{-9}$  per flight hour).

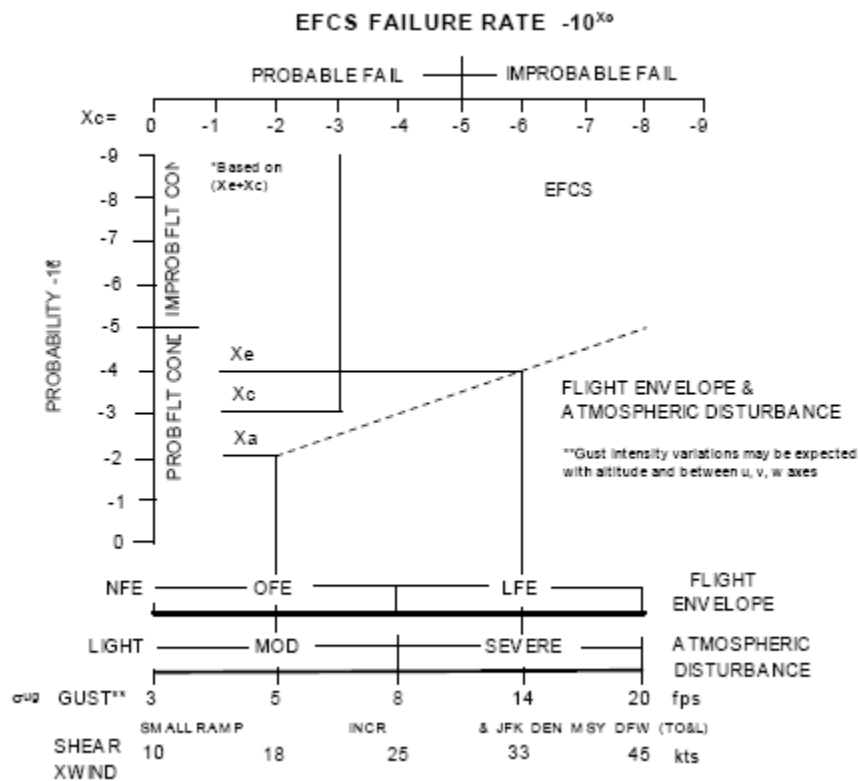
To identify which conditions must be evaluated and what the required handling qualities levels are, the following procedure is used.

1. First, the probability of atmospheric condition ( $10^{X_a}$ ) and flight envelope ( $10^{X_e}$ ) are determined using Figure 9. For instance, the probability of light turbulence is 1.0 thus,  $X_a$  equals 0. For nominal moderate turbulence conditions, the probability of occurrence is  $10^{-2}$  per flight hour, so  $X_a$  equals -2. The probability of being in the NFE is unity, thus,  $X_e = 0$ .; whereas, the nominal probability of being in the LFE is  $10^{-4}$  per flight hour; thus,  $X_e$  is -4.
2. Second, the probability of occurrence ( $10^{X_c}$ ) for failure conditions of the FCS are computed.
3. All combinations which are extremely remote are ignored; thus, those conditions where  $X_c + X_e + X_a \leq -9$  are not considered for evaluation.
4. For those not extremely remote, the probability of the condition is determined as the combination of the FCS and flight envelope condition where:
  - a. Probable Condition,  $(X_c + X_e) \leq -5$ ;
  - b. Improbable Condition:  $-5 < (X_c + X_e) < -9$ ;
5. The minimum acceptable handling qualities levels are given in Table II.

**Table II: Minimum Acceptable Handling Qualities Levels (FAA AC25-7)**

Flight Condition ( $X_a+X_c$ )	Atmospheric Disturbance ( $X_a$ )								
	Light			Moderate			Severe		
	Flight Envelope ( $X_e$ )								
	NFE	OFE	LFE	NFE	OFE	LFE	NFE	OFE	LFE
Probable Condition	S	S	A	A	C	C	C	C	C
Improbable Condition	A	A	C	C	C				

Similar to the military standards for fixed- and rotary-wing handling qualities, this evaluation procedure provides a probabilistic approach to handling qualities requirements that includes flight control system failure mode. However, unlike the military, the influence of flight envelope and atmospheric conditions are included in the probabilistic analysis.



**Figure 9: Derivation of Probability Conditions**

The verification procedure to ensure that the design meets these requirements is less clear. While equivalence of the SAT, ADQ, and CON levels to the Cooper-Harper scale are given, the AC does not explicitly state whether the pilots provide specific Cooper-Harper ratings or simply provide overall FAA ratings of SAT, ADQ, or CON. The use of

comments and ratings are also not detailed. The number of evaluation pilots is also not stipulated, but common practice is that a consensus approach from FAA test pilots is used to ultimately decide whether the aircraft’s handling qualities meet the requirements.

**3.4.5. Summary**

The military and commercial handling qualities standards are consistent but differ based on whether they are a certification standard (FAA) or a procurement vehicle (MIL-STDs). A summary is provided in Table III.

Each requires Level 1 handling qualities for normal operating conditions within the operational flight envelope. In all cases, a slight degradation in handling qualities is allowable:

- When operating outside of the operational flight envelope.
- In atmospheric conditions greater the “light” turbulence.
- When failure conditions are present.

The allowable degradation depends upon the probability of occurrence. Otherwise, Level 1 handling qualities for even low probability conditions could create unreasonable expense in designing extensive flight control system redundancy and elaborate control laws.

**Table III: Summary of Aircraft Handling Qualities Requirements In Failure State Conditions (Within OFE and Without Atmospheric Disturbances)**

Handling Qualities Requirements	Failure State Probability of Occurrence	
	<i>Level 1</i>	<i>Level 2</i>
MIL-HDBK-1797	<10 <sup>-2</sup> per flight	<10 <sup>-4</sup> per flight
F-35	<10 <sup>-7</sup> per flight hour	
ADS-33	<2.5 x 10 <sup>-3</sup> per flight hour	<2.5 x 10 <sup>-5</sup> per flight hour
FAA	<10 <sup>-5</sup> per flight hour	<10 <sup>-9</sup> per flight hour

**4. Spacecraft Handling Qualities Experience**

While the number of experiences within the space-/exploration-domain pales in comparison to the wealth of aeronautics knowledge and experience, there are space-domain lessons-learned and experiences in manual control and handling qualities. An overview of these works is provided in the following.

**4.1. Gemini**

Under the Gemini program, manual control and handling qualities investigations primarily focused on rendezvous, proximity operations, and docking (RPOD) since this



task was one of the primary goals of Gemini as a stepping-stone for the lunar landing objective of the Apollo program.

During Gemini, this work investigated the handling qualities influence of such parameters as:

- Spacecraft attitude control mode, control power, target lighting and target oscillatory motion (Riley, Jaquet, Bardusch, and Deal, 1965; Riley, Jaquet, and Cobb, 1966);
- “Remote” docking using Closed-Circuit Television (CCTV) (Long, Pennington, and Deal, 1965);
- Visual aides, in day and night conditions, to align to a docking target (Pennington, Hatch, Long, and Cobb, 1965);
- Hand controllers, instruments, and control modes (Jaquet and Riley, 1965);
- Visual simulation compared to “full-size” docking (Riley, Jaquet, Pennington, and Brissenden, 1966);
- Visual aides and attitude control modes in lunar orbit (Hatch, Riley, and Cobb, 1964).

These studies provided a wealth of data especially for evaluating “first-principles” since on-orbit RPOD had never been conducted prior to much of this work.

Unfortunately, application of these works to modern systems suffers in two critical aspects:

- 1) the handling qualities data used the Cooper rating scale (Cooper, 1957), not the Cooper-Harper pilot rating scale (Cooper and Harper, 1969) which is today’s accepted standard. A translation from the Cooper scale to the Cooper-Harper scale is not possible; they are two distinctly different rating methodologies.
- 2) for these “foundational” works, analog simulations were almost exclusively conducted; hence, critical digital flight control system effects and control law techniques were not evaluated.

It should also be noted that the docking tolerances for Gemini were significantly larger than those identified for the current and planned spacecraft (e.g., see AIAA, 1993).

## **4.2. Apollo**

In preparation for Apollo, spacecraft manual control and handling qualities evaluations focused on specific Apollo issues.

#### **4.2.1. Rendezvous, Proximity Operations, and Docking**

Since the basic principles of RPOD were proven under Gemini, RPOD evaluations investigated specific Apollo issues.

One significant difference between Apollo and Gemini, specific to RPOD, was the evaluation of Lunar Module (LM) docking. Two studies, in particular, were conducted. Full-size, pilot-in-the-loop simulations were conducted at LaRC's rendezvous docking simulator.

The first study (Pennington, Hatch, and Driscoll, 1966) was designed to investigate the pilot's ability to complete a successful docking by using only visual information for command and service module (CSM) transposition docking with the LM during the trans-lunar trajectory (between earth and moon). Evaluations consisted of:

- Control modes (rate command, rate command with attitude hold, and direct) with no thruster failures and single and double thruster quadrant failures.
- Dockings with tumbling targets

A second study (Hatch, Pennington, and Cobb, 1967) evaluated the Lunar-Orbit-Rendezvous of the LM with the CSM.

The CSM was the target vehicle - docking the ascent stage of the LM with its top hatch to the CSM using only visual observation of the target for guidance information. The objectives of the simulation program included:

1. The feasibility of docking the LM with its top hatch to the CSM
2. The efficacy of different visual aides
3. The effect of lighting conditions
4. The impact of different control modes (rate-command-attitude-hold mode, rate command, and direct attitude control).
5. The influence of the pilot wearing a pressurized suit.

On orbit test and evaluation of lunar orbit docking showed that using the LM to complete the docking was not desirable. The "torque to inertia ratio" for the LM was too high (Stafford, Armstrong, Collins, and Cooper, 1969) to allow precise, repeatable docking. Instead, the LM was the target and the CSM served as the chaser for the last 3 to 4 feet of docking in lunar orbit.

The parametric evaluations of Gemini and the specific developments for Apollo established the RPOD state-of-the-art. So much so, in fact, that as the Space Shuttle was being developed, RPOD was assumed to be a lower priority in the early days of the program (i.e., it had less technical risk) compared to other system development tasks

(Goodman, 2006). It wasn't until later in the program that the unique attributes of the Shuttle design for RPOD were recognized, resulting in "complex operational work-arounds over the life of the program" (Goodman, 2006) that might have otherwise been mitigated earlier in the Shuttle design by continued parametric evaluation of handling qualities requirements.

#### **4.2.2. Atmospheric Entry**

Several studies were conducted evaluating the pilot's ability to monitor and control an Apollo-type vehicle during atmospheric entry. Several of these studies were performed in centrifuges to approximate the motion cues (g-levels) for entry.

For instance, in one study (Wingrove, Stinnett, and Innis, 1964), pilot-in-the-loop evaluations were conducted to evaluate the ability of a pilot to: 1) control a reentry vehicle to various constant acceleration levels, 2) control entry range based on various displays and control augmentation levels; and, 3) perform monitoring and recovery procedures.

The results illustrated that basic velocity and range-to-go information displays are unsatisfactory for normal operation for pilot control of range-to-go; however, satisfactory performance is attainable if lead-information is displayed and the displays include roll rate or roll-angle command information. Simple methods for pilot monitoring of automated entries were developed and proven feasible. In fact, successful pilot-controlled transitions from skip-out to safe entry trajectories were documented.

#### **4.2.3. High-Altitude Abort**

A fixed-base simulation study (Meintel, Garren, and Driscoll, 1966) was conducted to determine whether a pilot could manually orient the Apollo vehicle to the proper reentry attitude following a high altitude (120,000 ft and above) abort by only using the "out-the-window" visual scene as an attitude reference. This work served as a follow-on study to a contractor study which showed that the flight crew could perform stabilization and orientation of the vehicle during a high-altitude abort, if attitude information was provided on head-down displays.

This study showed that:

1. If a visual yaw reference is available, such as a landmark or a vapor trail, manual orientation is possible for all aborts above a 120,000-foot altitude except the single-control-system Saturn V 120,000-foot abort.
2. Heading determination by using ground tracking, which requires a broken cloud cover along the orbital track, appears feasible for aborts above a 150,000-foot altitude. This technique is not usable for 120,000-foot aborts because of insufficient control time for accurate ground tracking.

4. Control system failures which occur during the visually-controlled abort maneuver greatly affect the pilot's ability to orient the vehicle to the heat-shield-forward attitude.
5. An unpressurized pressure suit does not affect the pilot's ability to perform the orientation maneuvers.

#### **4.2.4. Lunar Landing**

Fixed-base simulation and flight test evaluations were conducted using flight test and in-flight simulation vehicles (Jarvis, 1967; Matranga and Walker, 1965; Matranga et al, 1963) to support the development of acceptable control law and control characteristics for the lunar landing task. These works were used to develop the controllers and pilot-vehicle interface requirements for Apollo.

The initial Lunar landing handling qualities requirements (Cheatham and Hackler, 1966) revolved around the acceptability/utility of control law types for this task:

1. Proportional attitude thruster design with direct command through the rotational hand controller (RHC).
2. Proportional attitude thruster design with a rate command system through a RHC.
3. On-off Thruster design with direct command through the RHC.
4. On-off Thruster design with a rate command system through the RHC.

In Stengel (1970), the influences of digital phase plan controllers and quadratic input shaping from this baseline were addressed which created the “second generation” capability for Apollo. Additionally, handling qualities and pilot-vehicle interface issues, specific to the Apollo program, are well-summarized in Hatcher et al (1968), including the ability to monitor the automatic systems, re-designate the planned landing site, establish appropriate thrust response sensitivity, and provide the critical out-the-window visibility for Lunar Landing in Apollo. In fact, “the constraints placed on crew visibility by the design of the LM window and by trajectory parameters make the viewing of the programmed landing site a major problem.”

### **4.3. Space Shuttle**

#### **4.3.1. Rendezvous, Proximity Operations, and Docking**

As mentioned before, the success of Apollo led the Space Shuttle program to assume that RPOD had less technical risk compared to other system development tasks (Goodman, 2006). It wasn't until later in the program that the unique attributes of the Shuttle design for RPOD were recognized. The issues included rendezvous operations changes, control law modifications for reduced plume impingement (i.e., the “low-Z” control law mode), and the unplanned development of relative navigation sensors (Crewman Optical

Alignment Sight (COAS) and CCTV camera ranging rulers) and, by the 1990s, trajectory control sensor and hand-held Lidar. (Goodman, 2006).

#### **4.3.2. Approach and Landing**

The Space Shuttle program took on the challenge of providing a manual landing capability for an operational vehicle returning from orbit. The initial plans required the development of an operational capability of landing day or night in all types of weather on a 15,000-ft runway. The control system design was complicated by the requirement for a center-of-gravity position that ranged from statically stable to statically unstable (Powers, 1986). Flight path control is complicated because the center of rotation of the pitch axis is ahead of the pilot's position in the cockpit which means the pilot does not perceive any change in flight path for almost a full second after control input (NATO, 2000)

Three levels of control system performance were required for Shuttle (Powers, 1986), depending on the type and number of system failures.

- Level 1 was generally required for nominal and one failure state operating conditions, Level 2 was required in the event of two failures, and Level 3 was required for two failed auxiliary power units.
- Level 1 requirements consisted of system stability margins (high-frequency crossover gain margin of 6-dB and a 30-deg phase margin); time response criteria that were derived from a composite of then-available handling qualities criteria; and pilot ratings had to be better than 3 from real-time simulation.
- Level 2 requirements included lower stability margins and pilot ratings better than 6.
- The Level 3 requirements specified that the vehicle would be controllable.

The design process was heavily reliant on real-time simulation.

In 1977, low-speed characteristics of the Orbiter were evaluated in-flight during the approach and landing test (ALT) program. The first four landings were on the Edwards dry lakebed and were uneventful. In general, the flying qualities were quite good. The fifth landing, however, was on the 15,000-ft concrete runway.

ALT-5 resulted in a spectacular PIO. The pilot's touchdown aim point was about 5,000 ft beyond the runway threshold. The pilot perceived an overshoot of the intended touchdown point and attempted to correct, triggering an ensuing PIO. Records indicate almost continuous elevon rate limiting with a pitch PIO started seven seconds prior to first touchdown and a lateral PIO five seconds before touchdown (NATO, 2000). After the first touchdown and bounce, a more pronounced lateral PIO occurred, followed by a series of over-controlled skip and hop motions.

This event resulted in numerous activities, none of which included a fundamental change in the flight control system or control laws. A control law re-design to fundamentally fix the handling qualities deficiencies was not pursued since this late change would have been extremely expensive (cost and schedule). Instead, the modifications included:

- An adaptive stick gain algorithm that would reduce the pilot and system command gain whenever PIO conditions were approached.
- Increasing the stick force gradient by a factor of two.
- Changing priority rate-limiting logic of the elevens to reduce the interactions between the roll and pitch axes.
- A dedicated, elaborate training regiment using Shuttle Training Aircraft, since the required pilot skill is very exacting, as the pilot must learn new control techniques to avoid exciting the inherently poor flight path response characteristics.

The redesign and training regiment significantly improved the Shuttle handling qualities performance but did not completely eliminate the underlying deficiencies. For instance, on STS-3 (March, 1982) Columbia was in an incipient PIO at touchdown. A high crosswind caused an overshoot of the final approach course coming off of the Heading Alignment Circle. The result was a high-gain situation on short final, with the vehicle being in the beginning phases of a PIO just as it touched down. There was also about a 1 ½-cycle oscillation as the nose was lowered, resulting in a pronounced “slam-down” (NATO, 2000).

Lessons-learned (NATO, 2000) from the Shuttle recommended the following:

- Improved flight control system designs in the form of better flight path command logic were needed.
- Improved flight path awareness for the pilot, in terms of display logic was needed.
- The flight control system (time domain) design criteria were inappropriate; other available criteria were not used but should have been; new criteria were also needed to cover the unique characteristics of the Shuttle.
- Data from other lifting body flight tests (e.g., M2F2) were discounted but should have been referenced.

#### **4.4. Applicability of Aircraft HQ to Spacecraft**

The handling qualities issues cited above have almost exclusively been for situations involving direct manual control of a vehicle. Planned and future spacecraft, on the other hand, will be nominally operated, for a vast majority of scenarios, under automatic control. In the following, lessons-learned in defining “handling qualities” in the case of

highly automated vehicles are discussed in terms of how this work applies to these emerging vehicles.

Also, an attempt is made to draw analogies between spacecraft manual control tasks and aircraft evaluation task. The concept is to draw from the lessons-learned in the development of these aircraft-based evaluation tasks for the development of spacecraft handling qualities tasks.

#### **4.4.1. Automation “Handling Qualities”**

Human-automation interface requirements are not often thought of as being part of “manual control” handling qualities requirements, but they should be. Handling qualities is an evaluation of the performance and workload associated with the pilot-vehicle dynamic system (Figure 1). The only substantive difference between “manual control” and “automated control” is the physical process of activating the controller (i.e., whether it is a human or automation system at the controls).

Although a control task may be automated, history has shown that the best automation designs are “human-centered.” They are designed understanding the human needs for monitoring, intervention, and adaptation in concert with the automation. The rationale is that unless the automation is fool-proof and perfect, a human-centered automation design takes advantage of the fact that the world’s best adaptive controller – the pilot – can intervene, adapt, and overcome as necessary in the event that the automation is not successful.

The aeronautics-domain experience with automation has shown that human-automation interface requirements should be developed and evaluated in parallel and in concert with the vehicle’s handling qualities. Automation technology was originally developed with the hope of “increasing the precision and economy of operation while, at the same time, reducing the operator workload and training requirements. It was considered possible to create a system that required very little or no operator intervention and therefore, reduced or eliminate the chance of human error within the system” (Sarter, Woods, Billings, 1997). These precepts have gone largely unrealized. Research has shown that automation has, in fact, caused a rash of aviation accidents and incidents due to the introduction of automation aids, because of the adverse interaction between the automation and the human operators (Billings, 1997).

“Human-centered” automation designs principles are many and varied (e.g., see Billings, 1997). “Human-centered designs” should not merely be a process where a human factors team has been involved or tests have shown that humans can operate them. While such designs and design practices may produce acceptable results, three key principles for human-automation interaction must be addressed:

1. The automation must be observable by its human operator, (i.e., the human operator is appropriately informed).
2. The current and near-term future behaviors of the automation must be comprehensible, understood and predictable by the human operator.

3. The automation must be contextually appropriate for its application, designed to complement the human operator, and not automated just because it is possible..

These human-automation interaction issues are mirrored in the lessons-learned from Apollo and their use of automated entry guidance and control (Graves and Harpold, 1972):

- “Guidance logic must be simple... the more complicated the guidance logic, the more difficult the guidance is to monitor during the mission. The monitoring difficulty complicates the development of the monitoring procedures and increased the time required for flight-crew training.”
- “The guidance logic should be compatible with a backup or an alternate trajectory control procedures. That is, once an anomaly is detected in the trajectory control of the primary guidance system, an alternative technique must be available that will allow satisfactory trajectory control to be implemented so that the spacecraft will land near the originally selected target.”
- “The interaction between guidance system performance and attitude control system performance must be recognized. Realistic attitude control system response requirements must be established, and guidance-logic design must minimize the need for rapid response.”

The “handling qualities” of automated tasks (i.e., human-automation interface requirements) should be evaluated in three ways:

- 1) Pilots should conduct handling qualities of all tasks to be flown by the automation. This is not to imply that these tasks could or should be manually flown. By conducting these evaluations, the pilot gains an appreciation of what the automation must do to successfully control the vehicle. This knowledge is critical for understanding the behavior of the automation and what the pilot must do in the event (likely or unlikely) that they need or want to intervene or take-over. This task also defines what information (e.g., out-the-window visual cues or displays) are needed by the human to monitor, control, or interact with the automation.
- 2) Classic “handling qualities” evaluations must also be conducted in scenarios where, during the conduct of automated tasks, the pilot takes control of the vehicle and completes the task or temporarily takes-over and then re-engages the automation. This task evaluates the ability of the crew to take-over for or to intervene with the automation and the potential for upsets or discontinuities in the automation during this process.
- 3) Finally, “handling qualities” evaluations must also be conducted in scenarios where, during the conduct of automated tasks, the automation fails (passively



or actively) and the pilot must take control of the vehicle and complete the task.

The aeronautics-domain has numerous handling qualities requirements associated with automation, principally focusing on the failure mode effects (e.g., “No single failure of any component or system (of the automatic FCS) shall result in dangerous or intolerable flying qualities” MIL-STD-1797A) or the transition from automatic to manual control (e.g., “The aircraft motions following sudden aircraft system or component failures shall be such that dangerous conditions can be avoided by the pilot, without requiring unusual or abnormal corrective action,” MIL-STD-1797A). Requirements are available which attempt to quantify the handling qualities due to transient motions (aircraft acceleration) between automated and manual flight.

Further, the requirements stipulate that a reasonable time delay between the automation failure and initiation of pilot corrective action should be used in determining the acceptability of the failure transient. A minimum time delay value of 1 second is often used, but greater delays are suggested depending upon whether the task is being conducted with the pilot’s undivided or divided attention and their associated workload, if the failure occurrence is cued by a readily apparent acceleration, rate, or sound that will definitely indicate to the pilot that a failure has occurred, whether they are flying with their hands on or off the controls, plus an additional delay may be necessary to represent the time required for the pilot to diagnose the situation before they initiate corrective action (e.g., see FAA AC25-07, MIL-STD-1797A)

#### **4.4.2. Spacecraft Task Analogies**

An attempt is made to draw analogies between spacecraft manual control tasks and aircraft evaluation tasks. The concept is to draw from the lessons-learned in the development of these aircraft-based evaluation tasks for the development of spacecraft handling qualities tasks. Obvious differences in control response characteristics and the lack of aerodynamic forces may obviate these comparisons, but this effort to extract some lessons-learned is, nonetheless, tried.

##### **RPOD: Probe-and-Drogue Refueling:**

The aeronautics-domain task most like RPOD would be probe-and-drogue refueling (Figure 10). This task requires the aircraft (i.e., the refueling aircraft, or “maneuvering vehicle” in spacecraft terminology) to plug its probe into the drogue from the tanker (i.e., the “target vehicle” in spacecraft terminology).



**Figure 10: Probe-and-Drogue Refueling**

The probe-and-drogue refueling task involves the following:

- An initial condition behind the tanker is used with possible vertical and horizontal displaced from the drogue as well. Vertical and horizontal displacement increases the task severity but may not be operationally acceptable from a safety-of-flight perspective (e.g., unacceptable engine wake or wing-tip vortices).
- A closure-rate is established to approach the drogue and make contact. Constant closure-rates are usually encouraged since, to do otherwise, would introduce task variability.
- Sometimes, the closure-rate is stopped at a pre-contact position, and the evaluation pilot is asked to “touch” the basket at the 12, 3, 6, and 9 o’clock basket positions with the probe. The ability to do this task with precision and low workload demonstrates the ability of the pilot to control the probe (as opposed to a “lucky stab” at the basket.). This type of procedure can also be used as a form of “tracking task” – see Table IV.
- Once a connection is made, the pilot must maintain an acceptable refueling station-keeping position for a certain amount of time.

- Numerous attempts at contact are used. Because of aerodynamics, the basket will typically move as the refueling vehicle approaches; not always in a repeatable manner.
- Typical task performance standards are shown in Table IV.

The primary difference between this aviation task and the spacecraft RPOD task is that the basket (drogue) is subject to continual, time-varying aerodynamic influences. These influences manifest themselves as basket movements which depend upon the characteristics of tanker and its refueling system (e.g., the type of basket and drogue, the length of hose deployed, its location on the aircraft – e.g., body mounted or wing mounted), the atmospheric conditions (winds and turbulence), and the location of the refueling aircraft probe and its forebody aerodynamics.

Unlike the RPOD task, the accuracy of hitting the target is not used in the probe-and-drogue refueling task. The primary reason is that it would be difficult to instrument and record these parameters. Instead, successful hook-up is the manifestation of achieving the desired task performance. If the probe hits too hard, is off-angled, or doesn't hit hard enough, a successful hook-up will not be achieved. Observation can also be used to directly see if the probe hit the webbing or not.

**Table IV: Probe and Drogue Refueling Task Performance Standards**

<b>Probe-and-Drogue References</b>	<b>Task Performance Standards</b>	
	<i>Desired Performance</i>	<i>Adequate Performance</i>
MIL-STD-1797A	No PIO. Hook-up without touching basket webbing in at least 50% of attempts.	Hook-up in at least 50% of attempts.
F-35 Probe and Drogue Tracking (1 Min Duration) (Unpublished)	Maintain the probe between the edge of the basket and 10 ft aft of the basket. Maintain the probe vertically and horizontally within one-half basket diameter. No contact with basket. Note: Momentary excursion outside the desired limits that are considered to be a result of basket motion and beyond the control of the EP should be ignored. .	Maintain the probe between the edge of the basket and 10 ft aft of the basket. Maintain the probe vertically and horizontally the edges of the basket (unless sudden basket motion is caused by the tanker or external influences). No contact with basket. Note: Momentary excursion outside the desired limits that are considered to be a result of basket motion and beyond the control of the EP should be ignored. .

The probe-and-drogue refueling task can be broken down into components as a way to get better diagnostics of task performance and its handling qualities since there are many, many uncontrollable factors which influence the resultant task performance and handling qualities. The probe tracking task is one such way of generating diagnostic data.

Other than the notable dynamics differences, another key difference between the Space- and Aero-domains is that the probe-and-drogue refueling task is done strictly with out-the-window visual cues. No cockpit displays are measurably involved.

### Atmospheric Entry: Terrain-Following

The closest analogy to the atmospheric entry task may be a terrain-following task (Figure 11).

In a terrain-following task, the pilot is tasked to fly at a constant altitude above the terrain, following a planned route over the ground. The pilot's task is use attitude changes to maintain this trajectory.

The terrain-following task involves:

- A route of flight is chosen. Obviously, the more undulations in the terrain, the higher degree of maneuvering and hence, the greater difficulty of the task.
- In addition, the speed along the route and the altitude above ground level, also drive the task difficulty. Higher speeds and lower altitudes significantly increase the task demands. Pitch and roll control are essential to the task.
- Typically, guidance is provided on a Head-Up Display (HUD) whereby the task involves following a guidance cue which provides a reference for the pilot's manual control to achieve a terrain following (TF) performance objective.
- This task can also be flown by an auto-pilot; i.e., automatic TF system where the pilot monitors the performance and intervenes as necessary using this same display data with supplemental information from other aural and visual cues in the cockpit.
- If the TF guidance cue is commanding an *attitude* profile and thus, indirectly commanding the vertical and lateral TF flight path profile, the pilot will use an attitude reference symbol as the control reference. The advantage of commanding an attitude profile is that attitude reference is easier for a pilot to fly because of the direct interconnection between the control reference and pilot controller input. The difficulty in using attitude command guidance is that the guidance command will be continually moving and adjusting to maintain the TF flight profile. The pilot may lose an intuitive concept of the flight path and will only be reacting to attitude commands. This increases the task workload. Desired and adequate performance standards for an attitude tracking task, analogous to a TF profile, is shown in Table V.
- If the TF guidance cue is directly commanding a vertical and lateral *flight path* profile, the pilot will use the HUD Flight Path Marker (i.e., the velocity vector) as the control reference. The advantage of commanding a flight path

profile is that it is directly related to the performance objective and it should intuitively correspond to the outside visual references, if available to the pilot. (The HUD could also be augmented to show the TF profile (e.g., a pathway display concept), allowing the pilot's to see the upcoming path profile and better anticipate the required control activity and reduce the task workload.) The disadvantage is that the flight path control task is more demanding for the pilot because of the aerodynamic lag between attitude and flight path.

- Desired and adequate performance standards for a TF task are shown in Table V. The performance standards differ if a flight path reference is used (e.g., see Christensen et al, 1998) or if an attitude reference is used.



**Figure 11: Terrain Following**

The task performance standards shown in Table V would need to be adapted to the display interface for the particular task, to ensure that the performance standards are controllable and observable for the pilot. For instance, the flight path reference performance standard assumed that the guidance command was in the form of a box. This may not necessarily be so in all cases, so the task standard would have to be changed accordingly. For the attitude tracking reference, time-on-target statistics can also be used as a performance measure.

**Table V: Terrain Following Task Performance Standards**

<b>Terrain Following Task</b>	<b>Task Performance Standards</b>	
	<i>Desired Performance</i>	<i>Adequate Performance</i>
FPM Reference (Christenson et al, 1998)	Keeping at least half of the flight path marker (FPM) in the Terrain Following Guidance box for the entire route with no more than five excursions per minute. (This represented the pilot following the DBTC cue to within +/- 0.2 g.) No PIO (pilots were told that overshoots while correcting back to the box were not to be counted as excursions.)	Following the Terrain Following Guidance with the flight path marker and keeping the FPM at least touching the guidance box for the entire route with no more than five excursions per minute. (This represented the pilot following the cue to within +/- 0.4 g.) (pilots were told that overshoots while correcting back to the box were not to be counted as excursions.)
Attitude Reference (MIL-STD-1797 et al)	Time to acquire: TBD Pitch and roll attitude maintained within 5 mils in pitch, 5 degrees in bank; Overshoots: no more than one greater than 5 mils, none to exceed 10 mils in pitch; or 5 degrees, none to exceed 10 degrees in roll; No PIO	Time to acquire: TBD Pitch and roll attitude maintained within 10 mils in pitch, 10 degrees in bank; Overshoots: no more than two greater than 5 mils, none to exceed 20 mils in pitch; or 5 degrees; none to exceed 20 degrees in roll.

The analogy between the aircraft terrain-following task and the spacecraft atmospheric re-entry task is loose.

- The similarity is that the task may be flown using attitude or flight path references. Ultimately, the flight path trajectory performance is the desired result. The task may be automatically flown. Also, the commanded pitch and roll commands typically lead the flight path by a significant amount to avoid overshooting; thus, it is not intuitively or immediately obvious how the attitude tracking performance relates to the flight path tracking objective.
- The analogy between aeronautics and spacecraft quickly breaks down from the standpoint that the aero-TF task is directly visible for the pilot – that is, they can see the terrain in front of the aircraft if flown in visual flight conditions. Also, the flight dynamics in the terrain-following task are typically constant, unlike the atmospheric re-entry task where the aerodynamic influences vary tremendously. Finally, the atmospheric re-entry task is primarily a roll-only task, unlike the aeronautics-domain TF task.

Ascent Control: Take-off and Climb-Out:

The closest analogy to the ascent control task may be a takeoff and climb-out task.

In a takeoff and climb-out task, the pilot is tasked to rotate the aircraft to a target attitude reference. Upon lift-off, the pilot regulates the aircraft attitude to maintain a flight path (vertical and lateral).

The takeoff and climb-out task involves:

- A take-off rotation speed and target attitude is determined.
- A pitch rotation rate is also chosen. (The rotation rate will influence the task demands.)
- Upon rotation, the pilot is tasked to regulate the aircraft attitude and maintain a target climb-rate and hold a constant track (i.e., lateral flight path). The specific task demands are dependent upon the aircraft cockpit instrumentation. If HUD equipped, vertical flight path angle and track angle requirements can be used. Otherwise, rate-of-climb and heading angles targets can be used.
- Through the climb-out, the aircraft undergoes configuration changes (flap and gear position) and accelerates to a target climb-out speed. These changes induce significant aerodynamic changes which add additional complications to the pilot’s ability to maintain the required flight path.
- Attitude and / or flight path command guidance may be given to the pilot to assist in this task. Typically, however, the primary flight information already on the displays is sufficient for the pilot to complete this task.
- Task performance standards for this task are shown in Table VI.

**Table VI: Take-off and Climb-Out Task Performance Standards**

Take-off and Climb-Out Task	Task Performance Standards	
	<i>Desired Performance</i>	<i>Adequate Performance</i>
MIL-STD-1797A	Attitude control on rotation: Keep within $\approx 1$ degree of takeoff attitude with no more than one overshoot, not to exceed TBD degrees. Flightpath control: Keep within $\approx 1$ degree of specified climb-out angle. Groundtrack: Keep aircraft within $\approx 10$ feet of runway centerline or within $\approx 2$ degrees of runway heading. No PIO	Attitude control on rotation: Keep within $\approx 2$ degrees of takeoff attitude with no more than one overshoot, not to exceed TBD degrees. Flightpath control: Keep within $\approx 2$ degrees of specified climb-out angle, but not less than 0 deg. Groundtrack: Keep aircraft within $\approx 25$ feet of runway centerline or within $\approx 5$ degrees of runway heading

The analogy between the aircraft take-off and climb-out task and the spacecraft ascent control task is weak. The similarity is that the task uses attitude to track a flight path

reference trajectory. The flight dynamics and g-loading differences between the two tasks are very significant.

## 5. Recommendations

Historical evidence has shown that a design which provides excellent handling qualities enables four key benefits:

1. Task performance which meets the mission requirements both in terms of precision and accuracy, with tolerable pilot workload.
2. A more robust vehicle system, elastic to changes in task, stressors, and external disturbances, including pilot distraction.
3. Less sensitivity to pilot technique and hence, lower training costs.
4. Less risk in the design and higher safety margins in the operation of the vehicle.

Handling qualities evaluations are a critical component in the successful development of a new vehicle. The lessons-learned from aircraft developments (and spacecraft) should be applied. In particular, a “best practices” approach to spacecraft handling qualities development has been identified and should be adopted. Two key elements of this best practices approach include the following, somewhat non-intuitive concepts:

- 1) Deliberately search for handling problems, including the effects of design tolerances (parameter uncertainties) and failures. Identify the worst cases and any hidden weaknesses in the design, and fully explain any unexpected simulation results.
- 2) Evaluate the ability of the pilot to enter the control loop, to help out the automatic functions. Show that there is no tendency for divergence between the automatic and manual control functions.

In the unlikely event that handling qualities deficiencies are uncovered, the following prioritized recommendations for improving handling qualities are as follows:

- Fix the problem. A direct resolution to a handling qualities problem is always the most straight-forward and least costly (in terms of life cycle costs). Since changes in a design are dramatically lower the earlier in the design process they are uncovered, a vigorous handling qualities “exploration” process is critical in the very early stages of a program. One cannot wait for a “mature” design to begin this exploration. It must be done concurrently with the design process to help the design process and explore the design space, including parametric uncertainty and failure conditions.
- Develop work-around solutions. Indirect resolutions to handling qualities problems – for example, patch-work flight control solutions – can be



reasonably effective but they often bring along operational or training “baggage” which increase the overall life cycle costs and introduce elements of operational risk or operational constraints.

- Mitigate the problem by operational procedures/changes. Indirect resolutions to handling qualities problems by precluding particular operations to avoid a problem area can work but, again, this restricts the operation and additional training costs are incurred which increases the overall life cycle costs. Also, it is difficult, if not impossible, to prevent inadvertent entry into these problem areas and thus, avoid an accident or incident.
- Improve the Pilot-Vehicle Interface. Many times a handling qualities problem can be improved by providing the pilot with “observability” - the ability to see the “problem” - and “controllability” – the ability to acceptably control and manage the “problem” and keep it within acceptable performance and tolerance. The best adaptive controller ever made is the human pilot. Their talents can overcome many problems as long as they can see what’s wrong and can control it. Of course, the limitations of human pilots (e.g., fatigue, attention, physical strength) must also be considered in this trade-off.
- Training. Lastly, the best adaptive controller ever made – the human pilot – performs best when properly trained to the problem or trained to adapt to unique or changing circumstances near the problem area. Training should not be considered as a “sufficient” solution. It can also be expensive – since the training must have positive transfer and simulation of spacecraft situations can be difficult without elaborate facilities on Earth – and simulation is never perfect. Training by simulation is always an approximation.

## 6. Concluding Remarks

A synopsis of experience from the fixed-wing and rotary-wing aircraft communities in handling qualities development and the use of the Cooper-Harper pilot rating scale is presented as background for the US spacecraft handling qualities RDT&E. In addition, an overview of handling qualities experiences and lessons-learned from previous US spacecraft developments are also reviewed. These data are not nearly as plentiful as the aircraft data (for obvious reasons) but are offered as insight for the future spacecraft developments.

This report is not intended to be a comprehensive, “one-stop” location for all data but rather, provides a central location for best practices and important lessons-learned to be used as “take-aways” for the future spacecraft developments. References are given for those that desire additional information behind these data.

## References

- Alexander, J.D., and Becker, R.W. (1973): *Apollo Experience Report – Evolution of the Rendezvous-Maneuver Plan for Lunar-Landing Missions*. NASA TN D-7388. Aug 1973.
- American Institute of Aeronautics and Astronautics (1993): *Guide for the Serviceable Spacecraft Grasping/Berthing/Docking Interfaces*. AIAA G-056-1992. 1993.
- Bailey, R.E., Knotts, L.H., and Levison, W.H. (1987): *An Investigation Of Time Delay During In-Flight And Ground-Based Simulation*. Calspan Corporation Technical Report No 7205-16, August 1987.
- Bailey, R.E. and Bidlack, T.J. (1996): *A Quantitative Criterion For Pilot-Induced Oscillations - Time Domain Neal-Smith Criterion*. Paper presented at the AIAA Atmospheric Flight Mechanics Conference, Paper No. AIAA-1996-3434. San Diego, CA, July 29-31, 1996.
- Bellman, D.R. and Matranga, G.J. (1965): *Design And Operational Characteristics Of A Lunar-Landing Research Vehicle*. NASA-TN-D-3023. 1965. p. 1-50.
- Bennett, F.V. (1972): *Apollo Experience Report – Mission Planning for Lunar Module Descent and Ascent*. NASA TN D-6846. Jun 1972.
- Bennett, F.V. and Price, T.G. (1964): *Study of Powered-Descent Trajectories for Manned Lunar Landings*. NASA TN D-2426. August 1964.
- Bennett, F.V. (1970): *Apollo Lunar Descent and Ascent Trajectories*. NASA TM X-58040. presented at the AIAA 8th Aerospace Sciences Meeting, New York, New York. Jan 19-21, 1970.
- Brandon, J.M., Glaab, L.J., Brown, P.W., and Phillips, M.R. (1995): *Ground-to-Flight Handling Qualities Comparisons for a High Performance Airplane*. Paper presented at the Atmospheric Flight Mechanics Conference, AIAA paper No. AIAA-95-3457. Baltimore, Maryland. August 7-9, 1995.
- Cheatham, D.C., and Hackler, C.T. (1966): *Handling Qualities for Pilot Control of Apollo Lunar-Landing Spacecraft*. Journal of Spacecraft, Vol. 3, No. 5, May 1966.
- Christensen, K., Weber, G., Seelos, M., Moua, C., and Gillen, S. (1998). *F-16 Database Terrain Cueing. An Investigation Of Display Handling Qualities*. Proceedings of the IEEE Aerospace Conference, Volume 3, 21-28 Mar 1998, pp. 51-58
- Cooper, G.E. (1957): *Understanding and Interpreting Pilot Opinion*. Aeronautical Engineering Review, Vol. 16, No. 3, March 1957, pp. 47-51.
- Cooper, G.E., and Harper, Jr., R.P.: *The Use of Pilot Rating in the Evaluation of Aircraft Handling Qualities*. NASA TN D-5153, April 1969.
- Department of Defense (1980), *Military Specification: Flying Qualities of Piloted Airplanes*, MIL-F-8785C, Wright-Patterson Air Force Base, OH.
- Department of Defense (1997), *Handbook: Flying Qualities of Piloted Airplanes*, MIL-HDBK-1797, U.S. Air Force, OH.

- Dukes, T.A. (1985): *Guidelines for Designing Flying Qualities Experiments*. Naval Air Development Center Report No. NADC-85130-60, June 1985.
- Elliott, C. (2007): *Evaluating the Ability of Real Time Detection with Two Recent Pilot-Induced Oscillations*. Paper presented at the AIAA Atmospheric Flight Mechanics Conference and Exhibit, Paper No. AIAA-2007-6388. Hilton Head, South Carolina, Aug. 20-23, 2007
- Fabre, P. and Raimbault, N. (2001): *Probabilistic Neural Detector Of Pilot Induced Oscillations (PIO)*. Paper presented at the AIAA Guidance, Navigation, and Control Conference and Exhibit, Paper No. AIAA-2001-4353. Montreal, Canada, Aug. 6-9, 2001.
- Field, E.J., Armor, J.B., Rossitto, K.F., Schroeder, J.A., and Weingarten, N.C. (2002): *Comparison Of In-Flight And Ground Based Simulations Of Large Aircraft Flying Qualities*. Paper presented at the AIAA Atmospheric Flight Mechanics Conference and Exhibit, AIAA paper No. AIAA 2002-4800. Monterey, California. 5-8 August 2002.
- Federal Aviation Administration (1998), *Flight Test Guide for Certification of Transport Category Airplanes*, Advisory Circular AC 25-7A, Washington, D.C.
- Gawron, V.J. (2000): *Human Performance Measures Handbook*. Lawrence Erlbaum Associates, 2000.
- Goodman, J.L. (2006): *History of Space Shuttle Rendezvous and Proximity Operations*. Journal Of Spacecraft And Rockets, Vol. 43, No. 5, September-October 2006.
- Graves, C.A. and Harpold, J.C. (1972): *Apollo Experience Report – Mission Planning For Apollo Entry*. NASA TN D-6725, March 1972.
- Gungras, D. R., Linse, D.J., and Wood, C. (1996): *Expert System Cockpit Aid For The Evaluation Of Handling Qualities*. Paper presented at the AIAA Atmospheric Flight Mechanics Conference, Paper No. AIAA-1996-3378. San Diego, CA, July 29-31, 1996.
- Hackler, C.T., Brickel, J.R., Smith, H.E., and Cheatham, D.C. (1968): *Lunar Module Pilot Control Considerations*. NASA TN D-4131, February 1968.
- Harper, R. P., JR., and Cooper, G. E. (1986): *Handling Qualities and Pilot Evaluation*. Journal of Guidance, Control, and Dynamics, Vol. 9, No. 5, 1986, pp. 515-529.
- Harris, J.J. and Black, G.T. (1996): *F-22 Control Law Development And Flying Qualities*. Paper presented at the AIAA Atmospheric Flight Mechanics Conference, AIAA Paper No. 96-3379, San Diego, CA, July 29-31, 1996.
- Hart, S.G., and Staveland. L.E. (1988): *Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research*. In P.A. Hancock and N. Meshkati (Eds.), *Human Mental Workload*, (pp. 139-183). Amsterdam: North Holland Publishers.
- Hatch, H.G., Jr.; Riley, D.R.; and Cobb, J.B. (1964): *Simulating Gemini-Agena Docking*. *Astronautics and Aeronautics*, Vol. 2, No. 11, Nov. 1964, pp. 74-81.
- Hatch, H.G., Jr., Pennington, J.E., and Cobb, J.B. (1967): *Dynamic Simulation Of Lunar Module Docking With Apollo Command Module In Lunar Orbit*. NASA TN D-3972, June 1967.

- Hodgkinson, J., Potsdam, E.H., and Smith, R.E. (1990): *Interpreting The Handling Qualities Of Aircraft With Stability And Control Augmentation*. Paper presented at the 1990 AIAA Atmospheric Flight Mechanics Conference, Paper No. AIAA-1990-2825, Portland, OR, Aug 20-22, 1990.
- Hodgkinson, J., Page, M., Preston, J., and Gillette, D. (1992): *Continuous Flying Quality Improvement - The Measure And The Payoff*. Paper presented at the AIAA Guidance, Navigation and Control Conference, Paper No. AIAA-1992-4327. Hilton Head Island, SC, Aug 10-12, 1992.
- Hoh, R.H. (1990): *Lessons-Learned Concerning the Interpretation of Subjective Handling Qualities Pilot Rating Data*. Paper presented at the AIAA Atmospheric Flight Mechanics Conference, AIAA Paper No. 90-2824, Portland, OR, Aug 20-22, 1990.
- Jaquet, B.M.; and Riley, D.R. (1965): *An Evaluation of Gemini Hand Controllers and Instruments for Docking*. NASA TM X-1066, 1965.
- Jarvis, C.R. (1967): *Flight-Test Evaluation Of An On-Off Rate Command Attitude Control System Of A Manned Lunar-Landing Research Vehicle*. NASA-TN-D-3903. 1967.
- Kidd, E.A., and Bull, G. (1963): *Handling Qualities Requirements As Influenced By Pilot Evaluation Time And Sample Size*. CAL Report No. TB.1444-F-1, Feb.1963.
- Kirk, R.E. (1982): *Experimental Design: Procedures for the Behavioral Sciences*. 2nd Edition, Brooks/Cole Publishing Co., Belmont. California, 1982
- Klumpp, A.R. (1971): *Apollo Lunar-Descent Guidance*. Report R-695, MIT Charles Stark Draper Laboratory, June 1971.
- Leggett, D.B. and Cord, T.J. (1994): *Flying Qualities Demonstration Maneuvers*. Paper presented at the 7th AIAA Biennial Flight Test Conference, AIAA-1994-2113, Colorado Springs, CO, June 20-23, 1994, p. 52-61
- Lindquist, O.H. (1968): *Development of a Manual Thrust-Vector Control Mode for Apollo*. Journal of Spacecraft. Vol. 5, No. 3, March 1968.
- Long, E.R., Jr.; Pennington, J.E.; and Deal, P.L. (1965): *Remote Pilot-Controlled Docking With Television*. NASA TN D-3044, October 1965.
- Matranga, G.J. and Walker, J.A. (1965): *An Investigation Of Terminal Lunar Landing With The Lunar Landing Research Vehicle*. NASA-TM-X-74475. 1965. p. 1-30.
- Matranga, G.J., et al. (1963): *Handling Qualities and Trajectory Requirements for Terminal Lunar Landing, As Determined from Analog Simulation*. NASA-TN-D- 1921. 1963.
- Martin, G.F., and Eggemeier, F.T. (1991): *Physiological Assessment of Workload in Multi-Task Environments*. In Multiple Task Performance, Diane L. Damos, Ed. Francis and Taylor, Published by CRC Press, 1991
- McDonnell, J.D. (1968): *Pilot Rating Techniques for the Estimation and Evaluation of Handling Qualities*. Air Force Flight Dynamics Laboratory AFFDL-TR-68-76, December 1968.

- McDonnell, J.D. (1969): *An Application of Measurement Methods to Improve the Quantitative Nature of Pilot Rating Scales*. IEEE Transactions On Man-Machine Systems, Vol. 10, No. 3, September 1969. pp. 81-92.
- McElroy, C.E. (1988): *FAA Handling Qualities Assessment – Methodology in Transition*. Thirty-second symposium proceedings of the Society of Experimental Test Pilots, Beverly Hills, CA, October 1988.
- Meintel, A.J., Jr., Garren, K.R., and Driscoll, N.R. (1966): *Manual Control Of Launch Abort High-Altitude Apollo*. NASA TN D-3433, June 1966.
- Mitchell, D.G., Arencibia, A., and Munoz, S. (2004): *Real-Time Detection of Pilot-Induced Oscillations*. Paper presented at the AIAA Atmospheric Flight Mechanics Conference and Exhibit, Paper No. AIAA-2004-4700. Providence, Rhode Island, Aug. 16-19, 2004.
- Mitchell, D.G., and Aponso, B.L. (1990): *Reassessment And Extensions Of Pilot Ratings With New Data*. AIAA Paper No. 90-2823, paper presented at the 1990 AIAA Atmospheric Flight Mechanics Conference, Boston, MA, August 1990.
- Monagan, S.J., Smith, R.E., and Bailey, R.E. (1982): *Lateral Flying Qualities of Highly Augmented Fighter Aircraft*. AFWAL-81-3171, June 1982.
- Moorhouse, D.J. (1990): *On The Level 2 Ratings Of The Cooper-Harper Scale*. Journal of Guidance, Control, and Dynamics. Vol.13, No.1. 1990. pp. 189-191.
- Moroney, W.F., Biers, D.W., Eggemeier, F.T., and Mitchell, J. (1992): *A Comparison Of Two Scoring Procedures With The NASA Task Load Index In A Simulated Flight Task*. Paper presented at the 1992 IEEE National Aerospace and Electronics Conference, 1992. pp. 734-740.
- NASA (2003), *Human-Rating Requirements and Guidelines for Space Flight Systems, NASA Procedures and Guidelines*. NPR 8705.2, Washington, D.C.
- NASA (2005), *Human-Rating Requirements for Space Systems, NASA Procedural Requirements*. NPR 8705.2A, Washington, D.C.
- NASA (2007), *Constellation Program Human-Systems Integration Requirements, NASA Procedural Requirements*. CxP 70024 (rev. A), Washington, D.C.
- NASA (2008), *Human-Rating Requirements for Space Systems*. NASA Procedural Requirements, NPR 8705.2B, Washington, D.C.
- National Research Council (1997): *Aviation Safety and Pilot Control – Understanding and Preventing Unfavorable Pilot-Vehicle Interactions*. National Academy Press, Washington, D.C. 1997.
- Neal, T.P. and Smith, R.E. (1971): *A Flying Qualities Criterion for the Design of Fighter Flight-Control Systems*. Journal of Aircraft, Vol. 8, No. 10, 1971. pp. 803-809.
- North Atlantic Treaty Organization, Research And Technology Organization (2000): *Flight Control Design – Best Practices*. RTO Technical Report No. 29, Neuilly-Sur-Seine Cedex, France, Dec 2000.

- Pennington, J.E.; Hatch, H.G., Jr.; Long, E.R.; and Cobb, J.B. (1965): *Visual Aspects of a Full-Size Pilot-Controlled Simulation of the Gemini-Agena Docking*. NASA TN D-2632, 1965.
- Pennington, J.E., Hatch, H.G. Jr., and Driscoll, N.R. (1966): *A Full-Size Pilot Controlled Docking Simulation Of The Apollo Command And Service Module With The Lunar Module*. NASA TN D-3688, Dec 1966.
- Powers, B.G. (1986): *Space Shuttle Longitudinal Landing Flying Qualities*. Journal of Guidance, Vol. 9 No. 5, Sept-Oct 1986, pp. 566-572.
- Riley, D.R., Jaquet, B.M., Pennington, J.E., and Brissenden, R.F. (1966): *Comparison Of Results Of Two Simulations Employing Full-Size Visual Cues For Pilot-Controlled Gemini-Agena Docking*. NASA TN D-3687, Nov. 1966.
- Riley, D.R.; Jaquet, B.M.; and Cobb, J.B. (1966): *Effect of Target Angular Oscillations on Pilot-Controlled Gemini-Agena Docking*. NASA TN D-3403, 1966.
- Riley, D.R.; Jaquet, B.M.; Bardusch, R.E.; and Deal, P.L. (1965): *A Study of Gemini-Agena Docking Using a Fixed-Base Simulator Employing a Closed-Circuit Television System*. NASA TN D-3112, 1965.
- Riley, D.R. and Wilson, D.J. (1990): *More On Cooper-Harper Pilot Rating Variability*. Paper presented at the AIAA Atmospheric Flight Mechanics Conference, AIAA Paper No. 1990-2822, Portland, OR, Aug 20-22, 1990.
- Schultz, W.C., Newell, F.D., and Whitbeck, R.F. (1970): *A Study of Relationships between Aircraft System Performance and Pilot Ratings*. NASA CR-1643, July 1970.
- Stafford, T.P, Armstrong, N.A., Collins, M., and Cooper, L.G. (1969): *Apollo: Past, Present, and Future*. 13th Annual Symposium Proceedings of the Society of Experimental Test Pilots, Volume 9, No. 4, 25-27 Sept 1969.
- Stengel, R.F. (1970): *Manual Attitude Control of the Lunar Module*. Journal of Spacecraft, Vol. 7, No. 8, August 1970.
- Suchomel, C.F. (1996): *Automatic Rating Technique (ART) For Measured Flying Quality Ratings*. Paper presented at the AIAA Atmospheric Flight Mechanics Conference, Paper No. 1996-3377. San Diego, CA, July 29-31, 1996.
- Tseng, C., Gupta, P., and Schumann, J. (2006): *Performance Analysis Using A Fuzzy Rule Base Representation Of The Cooper-Harper Rating*. 2006 IEEE Aerospace Conference, July 2006.
- Twisdale, T.R. and Franklin, D.L. (1975): *Tracking Task Techniques for Handling Qualities Evaluation*. AFFTC-TD-75-1, May 1975.
- U.S. Army (2000), *Aeronautical Design Standard, Performance Specification: Handling Qualities Requirements for Military Rotorcraft*, ADS-33E-PRF, Redstone Arsenal, AL.
- Wilson, D.J. and Riley, D.R. (1989): *Cooper-Harper Pilot Rating Variability*. Paper presented at the AIAA Atmospheric Flight Mechanics Conference, AIAA Paper No. 89-3358, Boston, MA, Aug 14-16, 1989.

Wingrove, R.C., Stinnett, G.W., and Innis, R.C. (1964): *A Study Of The Pilot's Ability To Control An Apollo Type Vehicle During Atmosphere Entry*. NASA TN D-2467, August 1964.

REPORT DOCUMENTATION PAGE				Form Approved OMB No. 0704-0188	
<p>The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.  <b>PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.</b></p>					
1. REPORT DATE (DD-MM-YYYY) 01-06-2009		2. REPORT TYPE Technical Memorandum		3. DATES COVERED (From - To)	
4. TITLE AND SUBTITLE Cooper-Harper Experience Report for Spacecraft Handling Qualities Applications				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S) Bailey, Randall E.; Jackson, E. Bruce; Bilimoria, Karl D.; Mueller, Eric R.; Frost, Chad R.; Alderete, Thomas S.				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER 644423.02.36.15.04.50	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) NASA Langley Research Center Hampton, VA 23681-2199				8. PERFORMING ORGANIZATION REPORT NUMBER  L-19695	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) National Aeronautics and Space Administration Washington, DC 20546-0001				10. SPONSOR/MONITOR'S ACRONYM(S)  NASA	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)  NASA/TM-2009-215767	
12. DISTRIBUTION/AVAILABILITY STATEMENT Unclassified - Unlimited Subject Category 16 Availability: NASA CASI (443) 757-5802					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT A synopsis of experience from the fixed-wing and rotary-wing aircraft communities in handling qualities development and the use of the Cooper-Harper pilot rating scale is presented as background for spacecraft handling qualities research, development, test, and evaluation (RDT&E). In addition, handling qualities experiences and lessons-learned from previous United States (US) spacecraft developments are reviewed. This report is intended to provide a central location for references, best practices, and lessons-learned to guide current and future spacecraft handling qualities RDT&E.					
15. SUBJECT TERMS Spacecraft design; Handling qualities; Flight control; Pilot ratings; Cooper-Harper ratings					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			STI Help Desk (email: help@sti.nasa.gov)
U	U	U	UU	72	19b. TELEPHONE NUMBER (Include area code) (443) 757-5802