

11. DETERMINING THE TRANSFERABILITY OF FLIGHT SIMULATOR DATA

DAVID GREEN

The ability of rotorcraft pilots to hover and maneuver with agility in slow speed flight has placed unique and complex requirements on simulator manufacturers to demonstrate the authenticity of their product for the purpose of gaining rotorcraft training credit.

The FAA's evaluation of a simulator's capability is further complicated by the fact that the FAA does not have the resources to collect and compare the static and dynamic flying-qualities data that are required to conduct a comprehensive analysis. As a result, the FAA resorts to the practical approach of assigning qualified pilots to fly a flight simulator for the purpose of determining its value as a training device. Restated, pilots and engineers operate and otherwise evaluate flight simulators and render opinions about the adequacy of the simulator in terms of its proposed use and the credits requested. There are many other important objective measures of adequacy, but the importance of the subjective evaluation conducted by the pilot cannot be overstated.

This subjective portion of the evaluation may be enhanced by following the procedures suggested below. The details of a method for collecting and graphically correlating subjective ratings will be presented. The process has been tailored to aid engineers in their efforts to define the training value and limits of a given simulator with a substantially improved degree of confidence.

The FAA pilot's job is to define the simulator. Ideally, the pilot should be able to characterize the simulator in a format that can be understood by engineers and regulators. The evaluation pilot's insight into the real aircraft and its operational applications can be useful in helping engineers establish an appropriate scope of test to insure that the important flight phases and environmental conditions are considered.

The evaluation of rotorcraft flight simulator devices during up-and-away operations is seldom critical to the determination of overall suitability. This is because the aircraft is generally stable, and the quality of the visual

scene is often not critical to the learning experience. In contrast, the slow-speed regime is critical because most helicopter-unique training experiences occur in the slow-speed regime. In addition, the helicopter is least stable at these speeds, and the visual-motion system cues are most difficult to reproduce.

Relaxed slow-speed maneuvering high above the ground decreases the demand on the visual scene. In contrast, precision hover operations, low over a textured surface, place the greatest demand on the simulator's visual scene and motion system. In short, the evaluation pilot must investigate the authenticity of the simulator during a variety of maneuvers, including precision hover and during aggressive maneuvers, such as quick stops and inadvertent, uncommanded heading reversals (weather-cocking into a tailwind).

Although simulators are also very useful for teaching emergency procedures (such as tail-rotor failure), the validation of these events in a simulator dictates the use of quantitative data to determine reasonableness. A quantitative analysis is the only practical validation technique for such an event since there is normally little opportunity for pilots to build up an adequate (failure-mode) experience base in a real aircraft for use in an evaluation of the characteristics designed into a simulator.

The pilot assessment of suitability has historically been a key factor during the evaluation of aircraft by the FAA. The importance of this activity is difficult to overstate. Thus, before proceeding, it is useful to take a brief look at current procedures to establish a common point of departure.

Although research pilots and military test pilots tend to employ pilot rating scales, FAA pilots typically do not. The FAA pilot's task is to determine if the aircraft and its systems are safe. They make determinations about the adequacy or suitability of an aircraft for civil operations. There really is little call for pilot rating data per se. In addition, FAA pilots are primarily interested in workload,

and the basic pilot rating scale is not well suited to such an application. Finally, when the pilot ratings of several pilots are compared, they often do not agree, and such disagreements tend to bring the validity of the entire evaluation into question.

In short, the lack of a usable (FAA-oriented) pilot rating scale and the historical problems stemming from scatter in the data have produced deterrents to the general use of pilot ratings. These deterrents need to be eliminated before FAA pilots and engineers can be expected to embrace an evaluation method for flight simulators that involves pilot ratings.

There are many explanations for disagreements in pilot subjective ratings, and though some scatter in the data is normal, all evaluations should be conducted so as to minimize the scatter in the ratings. This presentation deals at great length with this issue and offers techniques to minimize scatter in the data when a number of pilots are employed on the same evaluation.

The method presented is based on the premise that if an engineer asks two equally qualified pilots the very same question, the result will be a common answer (pilot rating). A sloppy approach to staging a rating question to a number of pilots will in turn produce scatter in the results. That is, the proposed method introduces a discipline to the evaluation process.

Nevertheless, all scatter cannot be eliminated, nor should it be. Some apparent scatter in the data is not scatter at all, it is more data. For example, some disagreement in ratings may be explained by examining the background of the pilots. One pilot may be much more qualified in the aircraft than the others. Alternatively, one pilot may have used a different piloting technique and effectively changed the task. There is almost always a reason for apparent scatter that is not eliminated by the discipline to be proposed.

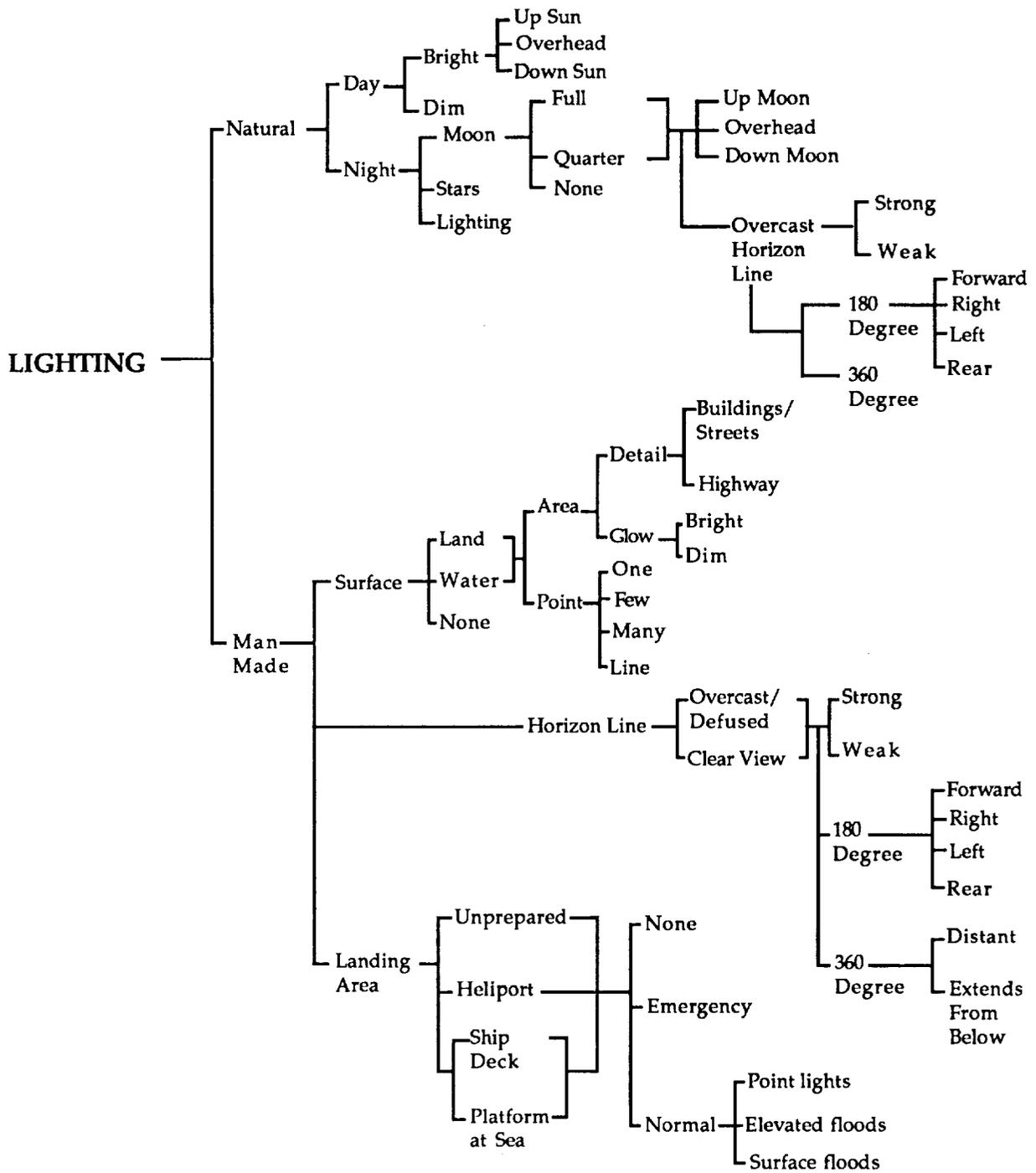
Pilots evaluate simulators by manipulating them as though they were flying a real aircraft in the conduct of a real mission task. Some operations are conducted single-pilot, some are two-pilot operations. Some flights are conducted with all systems operative, others are conducted with a variety of failures. Some tasks are very relaxed. Some relaxed flight tasks are made more difficult by the need to accomplish a number of secondary tasks at the same time. Other tasks require a great deal of precision interaction with the vehicle. Regardless of the basic circumstances, if the evaluation pilot is not required to work hard, there will be little potential for the kind of stress required to obtain a useful evaluation.

For example, a relaxed task such as a cross country flight, 1,000 feet above rolling terrain, bathed in bright sunlight, may not introduce sufficient workload to detect the shortcomings of a given simulator. Gusty winds will increase the workload. Decreasing visibility will also increase the workload. The introduction of factors that produce increasing levels of workload result in stress and enable pilots to find faults which allow them to become more discriminating in their assessments of a simulator's performance and related authenticity.

The fact is, pilots train to insure that they are able to cope with adversity in flight. They learn how to fly instrument approaches, and how to provide compensatory control inputs to suppress the gust response of their aircraft in the real world. Pilots must learn how to fly and deal with failure modes in a variety of environments. Anyone can quickly learn to fly almost any kind of aircraft on a clear day under calm conditions. Darkness, turbulence, and aircraft failure modes stress the pilot's ability to maintain safe flight conditions. It seems reasonable that one of the objectives of simulation should be to provide a pilot with the opportunity to experience a variety of adverse (stressful) combinations of flight environments and failure modes with the intended purpose of accelerating the learning process, aging the pilot to maturity in the least calendar time and at a minimum expense to the employer, and at the same time maintaining maximum safety by minimizing accident exposure in actual flight during abnormal and emergency operations.

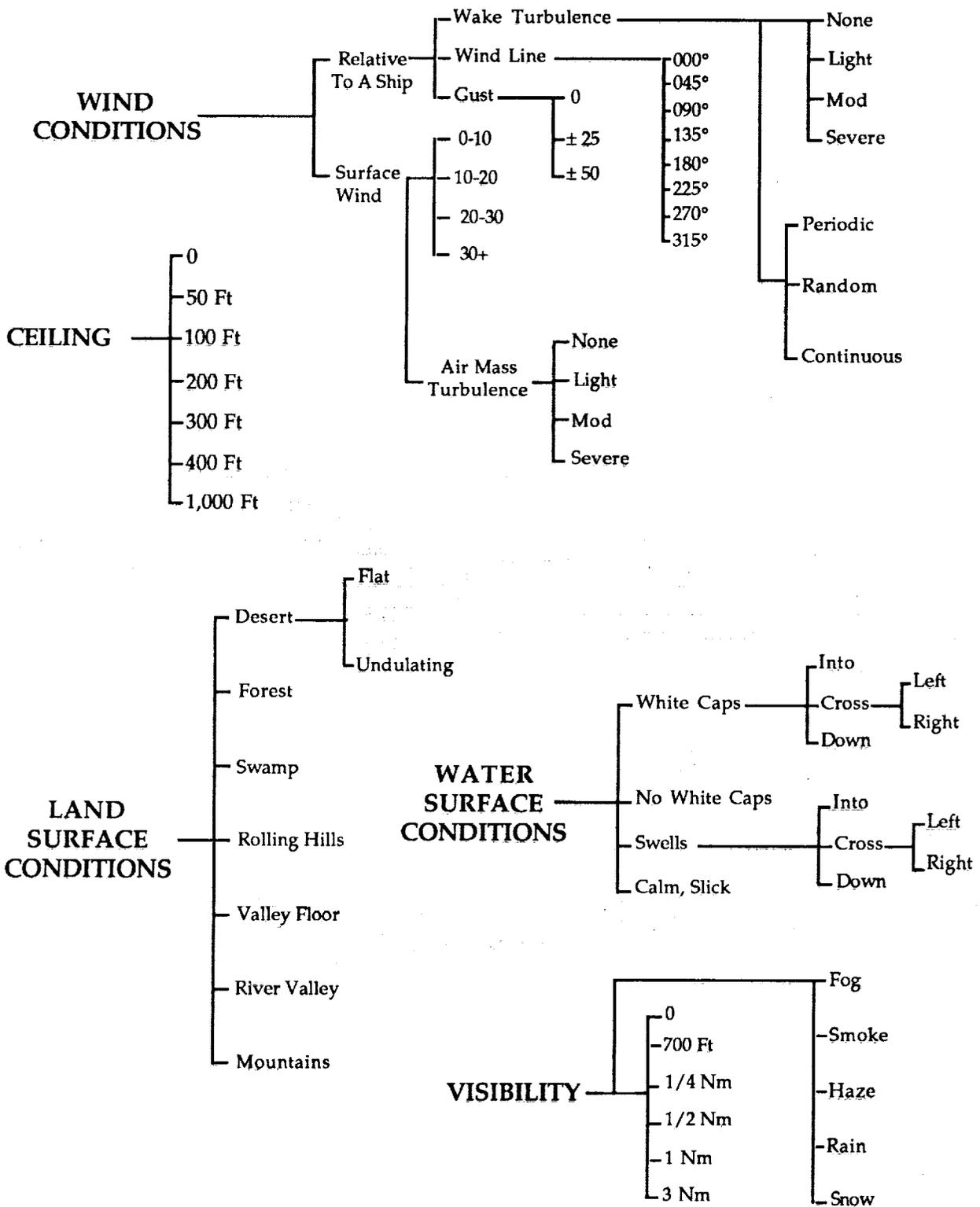
Figures 1(a) and 1(b) illustrate the variety of unique conditions which collectively define the environment within which a pilot can be expected to fly a rotorcraft. These environmental conditions can be used in a variety of visual conditions. The authentic duplication of these environments may dictate that a simulation device have a large repertoire of visual scenes. After some analysis, one might conclude that the availability of a large number of discrete visual scenes is not as important as the authenticity of the scenes available in the simulator. Repeatability of specific scenes in the simulator is also useful when analyzing the effect of variables such as pilot experience and training levels on the ability of crews to accomplish specific maneuvers. Waiting with a real aircraft for specific meteorological conditions (in the real world) to be repeated to derive similar data can be prohibitively expensive.

A moonless, starless flight over a dark sea is easy to simulate. The world is dark. Daylight scenes are more difficult. Images of trees, buildings, and runways as



(a) Lighting

Figure 1. Characteristics defining operational environment.



(b) Weather and terrain

Figure 1. Concluded.

observed through a haze may or may not be authentic; it is difficult to know. Maybe we don't even care if such scenes are authentic. The need for a sharp representation of microtexture during a low hover, on a bright day, is often very difficult to authentically simulate. This may be one of the most significant conditions to evaluate, for a failure to achieve the desired authenticity in the low-altitude, daylight environment may preclude the accomplishment of a precision hover training task.

The introduction of turbulence into this task (environment) can prevent a pilot from accomplishing a precision hover task in some real helicopters. Thus, the introduction of turbulence reduces the expectations of the pilot where he no longer expects to do well in the simulator either. Here the introduction of turbulence into a simulation event has the potential of masking some simulator problems because of decreased expectations. The point: one must be careful in the use of environmental variables. We will return to the environment later.

Systematic reports of subjective evaluations typically employ pilot rating scales. The most popular pilot rating scale is referred to as the Cooper-Harper pilot rating scale (see fig. 2). With ratings ranging from 1 to 10, it is the basic scale for most aircraft flying-qualities research work accomplished today. This an excellent scale, supported by 40 or more years of experience, but it lacks the detailed definition required for the evaluation of simulation devices. The range of this scale extends beyond the scope (or typical needs) of most FAA evaluations of simulation devices.

It is conceivable that the pilot of a certified civil helicopter may experience a situation to which a rating of 7 could be assigned, but even 7s should be rare. A rating of 7 means that the pilot was in control, but that the pilot was working as hard as possible, and that the resulting performance was inadequate.

At the other extreme of the scale, the pilot rating of 1 is reserved for highly automated flight-control systems or extremely relaxed tasks. In summary, pilots actively controlling certificated aircraft (with no system failures) in normal operational environments are expected to assign ratings that range between 2 and 5.5. Pilots evaluating automated flight-path control may assign 1 and 1.5. Serious flight-control failures, or very adverse operating environments, or difficult combinations of failure mode and bad environments, may produce pilot ratings of 6 or more.

Figure 3 shows a scale that has been expanded to meet the needs of the FAA for the evaluation of civil rotorcraft operations. This rating scale is only a sugges-

tion; it has not been endorsed by the FAA and there is every reason to expect that it can and should be improved. Nevertheless, the added detail is intended to help a group of pilots produce more consistent results by minimizing the opportunity for scatter in the data caused by individual interpretation of the Cooper-Harper scale.

When you compare the scale in figure 2 with the scale in figure 3, be advised that they are the same scale. The words in figure 3 are meant to expand upon the words in figure 2. They are intended to provide pilots with a better understanding of the meaning of the very brief statements in figure 2. Also note that the expanded scale provides definitions for ratings of 1.5, 2.5, 3.5, etc., whereas figure 2 does not. These additional half-ratings are not the invention of the author; they have been used from the beginning of time. The use of half-ratings is required, because most ratings range between 2 and 5. Experience has shown that the rating scale has been used as a kind of shorthand for pilots to communicate with engineers and other pilots. It is used to report the results of research that involves many, many variations in the evaluation task or characteristics of the aircraft. The half-numbers increase the number of "quality steps" available within a given small range of ratings to allow pilots to achieve the desired discrimination or hierarchic ranking of evaluation situations. These additional quality steps also allow the pilot to more accurately report the effect of variations in the environment on pilot-aircraft performance.

Pilots should not be required to commit the scale to memory, but pilots should make an effort to develop an awareness of the scale. They then should be allowed to look at the scale during the debriefing period following a flight evaluation. At that time, the pilot should rate the simulator experiences. This process will be developed in detail later.

Assume that a team of four pilots has been selected to evaluate a simulator. Their first step is to refresh their knowledge of the aircraft. If they are very familiar and current in that respect, this step is accomplished from memory. But for this example, assume that all of these pilots need to fly the aircraft. The first pilot, Green, conducts the hover-landing task described on the "Pilot Data Card" under the four conditions identified in figure 4 as A, B, C, and D.

Each time a pilot conducts the task, the factors that define the situation are recorded. Next, an assessment is entered for each situation. In this example, the assessments have ranged from a rating of 2 for a "clear day, calm air" to a 6 for an "overcast nighttime" situation. The

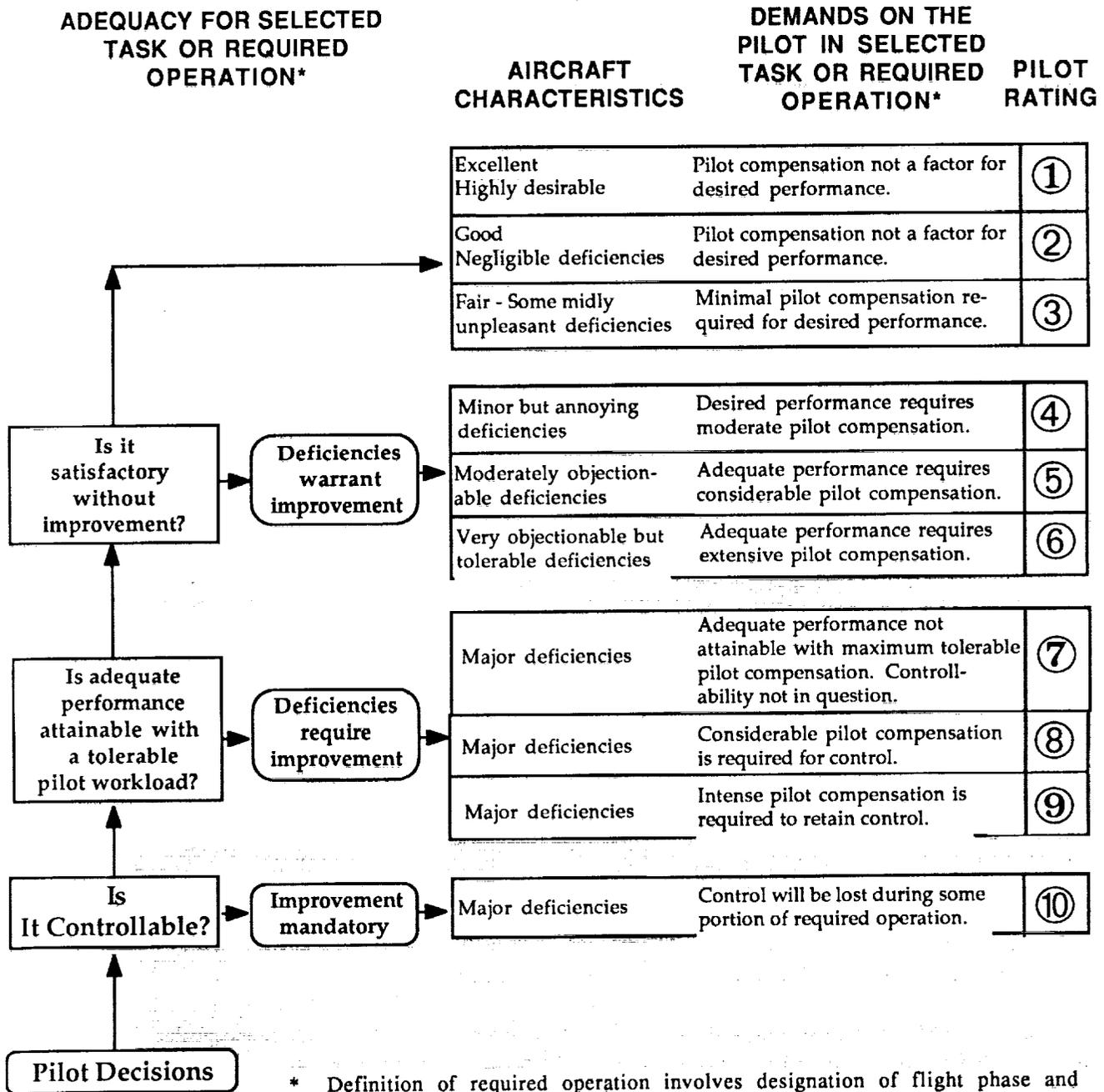


Figure 2. Cooper-Harper pilot rating scale.

| | | |
|-----|---|---|
| 1.0 | Excellent Highly Desirable | From time to time, the pilot may instruct the autopilot. System achieves long and short term objective with no pilot input directly to the conventional flight controls; inputs are selected via secondary (electronic) controls. The quality of flight path performance is self-monitored and alerts are provided to the pilot when he needs to take over; first and second failures are fail operate. Automatic mode shifting is provided (i.e., cruise to glideslope or glideslope to go around). |
| 1.5 | | From time to time, the pilot may instruct the autopilot. System achieves long and short term objective with no pilot input directly to the conventional flight controls; inputs are selected via secondary (electronic) controls. The quality of flight path performance is self-monitored and alerts are provided to the pilot when he needs to take over; first failure is fail operate: second or third failure one fail passive. Pilot is required to make occasional long term trim adjustments in one or two controls during transitional flight or during mode shifts. |
| 2.0 | Good | System achieves long term and short term gust suppression objectives with little or no pilot input directly to the conventional flight controls; inputs are often accomplished via secondary (electronic) controls. The quality of flight path performance is self-monitored and alerts are provided to the pilot when he needs to take over. Monitoring of short and long term response continuous but relaxed. Pilot may be required to occasionally adjust one axis/parameter during the performance of precision maneuvers or during major flight path changes. |
| 2.5 | | The pilot is continually involved in monitoring the short and long term performance of the aircraft. Deviations develop slowly and in a predictable way, and can be eliminated quickly with relaxed control techniques. Errors generally develop along or about one axis at a time. |
| 3.0 | Fair, Some Mildly Unpleasant Characteristics | The pilot is continually involved in the short-term control of the aircraft. Two or more controls are typically displaced in a sequential pattern. The aircraft can be trimmed with no more than one parameter/control needing attention at any given time. Control techniques are relaxed and pilot compensation is predictable and easy but requires continuous involvement. |
| 3.5 | | There is a characteristic that occasionally requires heightened attention, potentially disrupting the pilot's scan or control technique and momentarily taking precedent over other tasks. The aircraft is just a bit less predictable, possible because of problems trimming or due to an inconsistent response to gusting winds. |
| 4.0 | Minor, But Annoying Characteristics | Moderate pilot compensation is required. For relaxed flight phases, the control activity required is clearly achievable, but the effort produces impatience with the task and fatigue. Adjusting one control may require adjustments in other controls. For precision tasks, the workload contributes to occasional errors and excessive deviation. |
| 4.5 | | Moderate pilot compensation is required to achieve desired performance. There are one or more clearly annoying characteristics that make relaxed control clearly unachievable. On occasion, the desired performance is not achieved without considerable pilot compensation. |

Figure 3. Expanded evaluation scale for evaluation of civil rotorcraft.

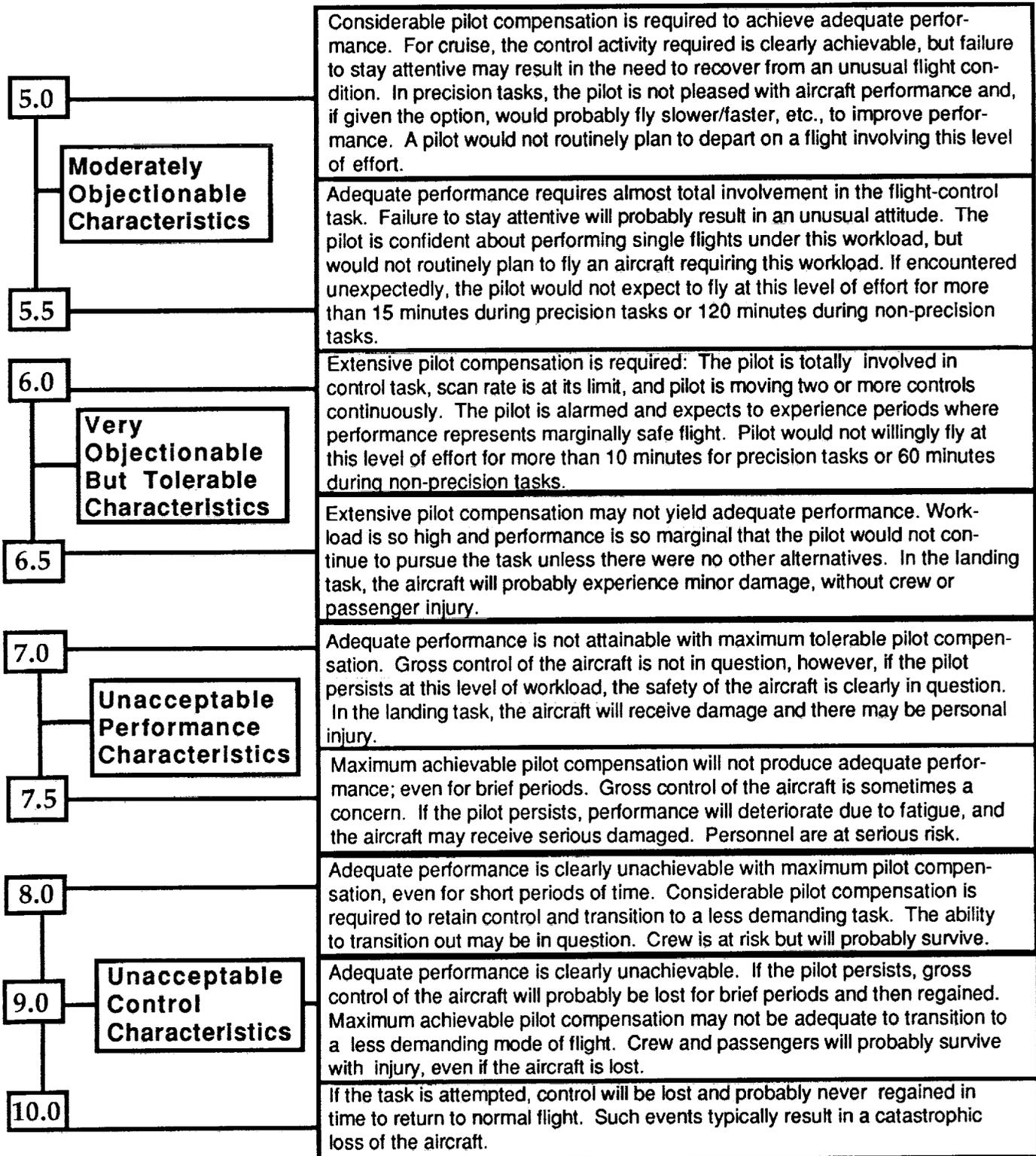


Figure 3. Concluded.

| TASK SHORT TITLE | | PILOT DATA CARD | | SIM FLT <input type="checkbox"/> |
|---|--|--------------------------|--|---|
| PLATFORM HOVER-LANDING | | Pilot Name: <u>GREEN</u> | | A/C FLT <input checked="" type="checkbox"/> |
| TASK: <i>Low hover in confined area. Landing on a platform one hundred feet above a water surface. Obstructions are present ahead and to the right. Upon landing rotor clearance is 30 feet to closest obstruction. Steel structure rises ahead.</i> | | | | |
| SITUATION ID CODE | FACTORS DEFINING THE TASK ENVIRONMENT SITUATION | | | PILOT ASSESSMENT (RATING) |
| A | Clear Day, Calm Air. | | | 2 |
| B | Clear Day, 10 KT RT Cross Wind. | | | 2.5 |
| C | Clear Day, 10 KT RT Cross Wind, Gusting to 17 KT. | | | 5 |
| D | Night, Overcast, no surface lights, single landing LT, 10 KT RT Cross Wind, Gusting to 17 KT, (see Note 1) | | | 6 |
| OPERATING STATE: Normal CONFIGURATION: Mid wt, mid C.G., Doors closed | | | | |
| Note 1: Tower obstruction lights, landing pad edge lights. | | | | |

Figure 4. Pilot rating data card for assessing one aircraft for the accomplishment of one task under four environmental conditions.

pilot's task involves a final flare and hover-landing to a platform on an oil rig in the open sea. The platform landing is considered a confined landing area involving the need for precision operations to avoid obstructions and to properly position the aircraft on the platform.

To continue this example, assume that three more pilots fly the same task under the same conditions and that they individually complete a data card. Their findings are summarized in figure 5. It is obvious that these four pilots did not totally agree, but when we analyze the results, we find the data are quite usable. First, we observe that the weather is never as constant or homogeneous as we would hope. As a result, all pilots probably operated the aircraft under slightly different conditions. Second, it is interesting to discover that pilot Black is most familiar with the aircraft and has extensive experience operating from platforms and ships at sea, day and night. Conversely, Brown has the least experience with the aircraft and the task-environmental situations evaluated.

The ratings in figure 6 are then the sum results of four pilots evaluating their personal "pilot-machine" performance under four task-environment situations. It must be understood that the rating process is personal. It refers to the performance that the evaluation pilot has achieved in flight. This performance evaluation is then something of a

self-appraisal and is the product of the pilot's skill level at the time, as well as the personal experience accrued by the pilot prior to the flight event that produced the recorded pilot rating.

This is the way the process should work. Some flying-qualities analysts ask pilots to establish a rating which they feel would reflect how the average pilot would evaluate a task. Such an approach is not applicable here. For this method to work, pilots must rate their personal performance.

The results summarized in figure 5 have been plotted in figure 6. This plot illustrates the preferred data presentation format for most comparative analyses. The format has been designed to be easily understood, and a shaded band has been added to figure 5 to emphasize the lack of scatter.

As noted before, there is some scatter in the data, but not a great deal. Experience has shown that the scatter will increase as the environment becomes extremely adverse. A larger scatter band is also possible when pilots are asked to evaluate degraded modes that they do not have a great deal of experience with. Both situations seem to suggest that a lack of pilot familiarity with the task or environment can produce scatter. This apparent uncertainty is both understandable and acceptable.

| TASK: <i>Normal Flare, Hover-landing onto Confined Elevated Platform Area.</i> | | | | | |
|---|--------------|--------------|--------------|--------------|--|
| ID CODE | PILOT | | | | |
| | GREEN | BLACK | BROWN | WHITE | |
| A | 2 | 1.5 | 2.5 | 2 | |
| B | 2.5 | 2 | 3 | 2.5 | |
| C | 5 | 4 | 5.5 | 4.5 | |
| D | 6 | 5 | 6.5 | 5.5 | |

Figure 5. Summary of pilot assessment data.

Figure 7 illustrates the next step in the method. For this illustration, pilot Green has been asked to evaluate the same hover-landing task for three additional and slightly different environmental situations (E, F, and G). The aircraft is not to be flown specifically to evaluate these situations. Instead, the pilot is asked to draw on experience. Green can relate well to two of these situations because he has personally experienced them in flight. We are not sure exactly when, but in any event, he relates well to these conditions and is easily able to provide an assessment of how well he can fly the aircraft. One situation, G, he has not experienced in the aircraft being evaluated, but he has flown other aircraft onto similar platforms under conditions approaching those identified with G. Thus we characterize G as a projected assessment. It is in effect an extrapolation. This extrapolation technique is not new; it is widely used during early assessments of military aircraft, every time development testing is initiated.

Here again, a certain amount of scatter in the data can be expected when the assessments of two or more pilots are compared. Projected ratings are subject to the greatest scatter, but even that can typically be explained and it is normally of little consequence. The scatter in projected ratings of operations involving violent weather at night can be expected to produce scatter of the order of ± 2 pilot ratings. On the other hand, the data from an extremely qualified pilot will often fall along the mean of the scatter in the projected data developed by less-qualified pilots. The data developed by pilots who do not understand the

pilot rating process are normally in conflict with the group and can be easily identified as such, and discounted.

Figure 8 illustrates one way that pilot ratings can be plotted for analysis. Note that the sets of conditions have been ordered across the chart in a way that allows the rating to ascend from left to right. This results in a situation where the sets of environmental factors are becoming more adverse left to right. This arrangement enhances data analysis and helps the evaluator insure that a complete spectrum of task complexity has been considered.

A simulator can be evaluated by one pilot or by a team of pilots. To simplify this next discussion, one pilot, Green, will be considered. Remember that the data in figure 8 represent the best characterization of the real aircraft that Green was able to establish. Assume for the moment that the data provided by the remaining pilots would have nominally agreed with Green's data. This confirms that Green's ratings of the seven different operating environments is sufficiently accurate to use in the evaluation of a simulator. In addition, an inspection of the seven operational environments used in flight confirms that they probably provide an adequate spectrum of situations to use as simulation environments for evaluating a simulator. That is, a simulator operator can be asked to electronically program the simulator to present the evaluation pilot with a set of winds, turbulence, and visual scene factors that collectively represent each of the environmental conditions relating to each of the situations defined in figure 8.

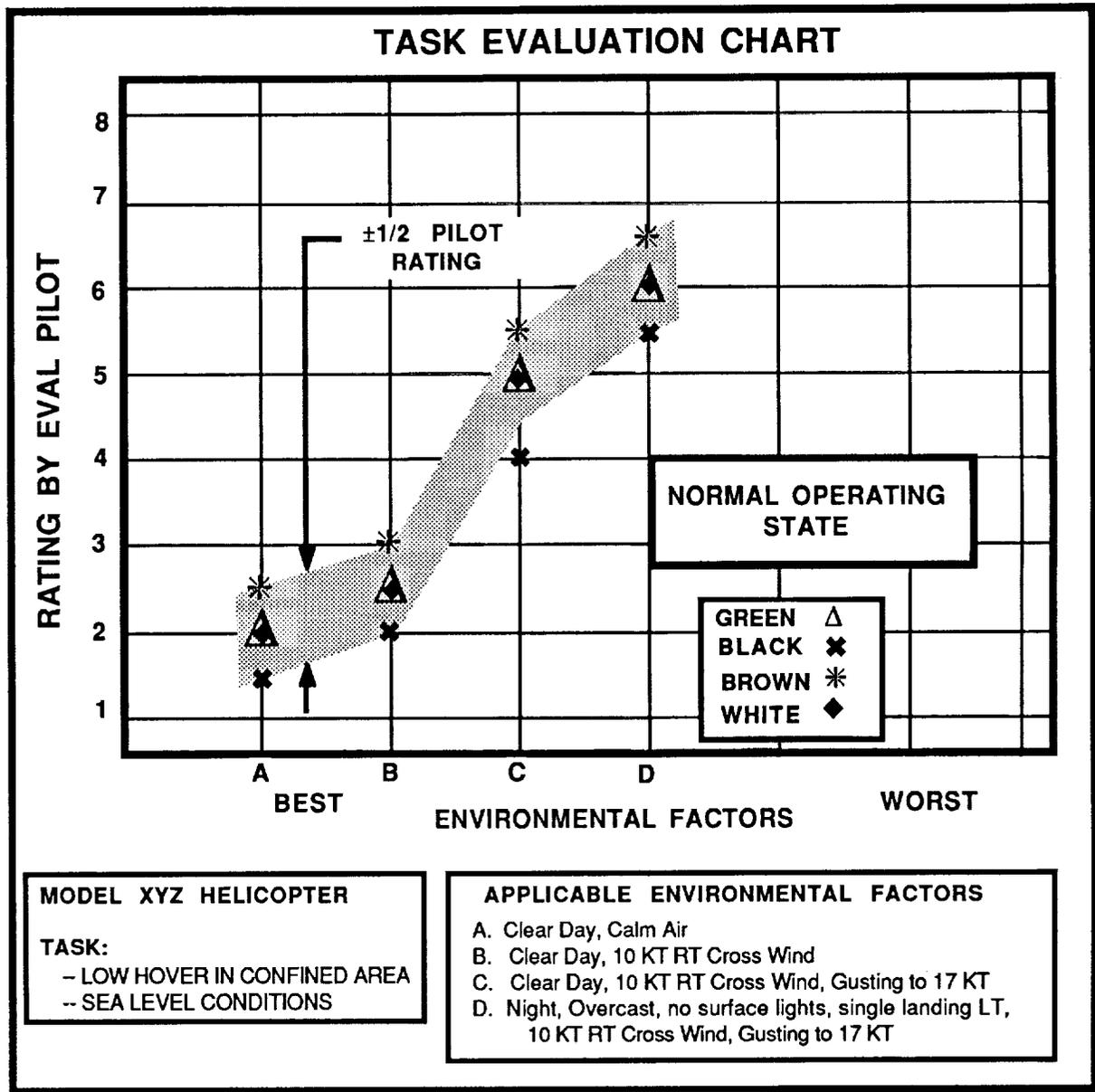


Figure 6. Charting pilot assessment data.

| TASK SHORT TITLE PLATFORM HOVER-LANDING | | PILOT DATA CARD Pilot Name: <u>GREEN</u> | | SIM FLT <input type="checkbox"/> | | A/C FLT <input checked="" type="checkbox"/> | |
|---|--|--|--|----------------------------------|---------------------------|---|-----------|
| TASK: <i>Low hover in confined area. Landing on a platform one hundred feet above a water surface. Obstructions are present ahead and to the right. Upon landing rotor clearance is 30 feet to closest obstruction. Steel structure rises ahead.</i> | | | | | | | |
| SITUATION ID CODE | | FACTORS DEFINING THE TASK ENVIRONMENT SITUATION | | | PILOT ASSESSMENT (RATING) | | |
| A | | Clear Day, Calm Air. | | | 2 | | |
| B | | Clear Day, 10 KT RT Cross Wind. | | | 2.5 | | |
| C | | Clear Day, 10 KT RT Cross Wind, Gusting to 17 KT. | | | 5 | | |
| D | | Night, Overcast, no surface lights, single landing LT, 10 KT RT Cross Wind, Gusting to 17 KT | | | 6 | | |
| E | | Night, Full Moon, Stars, Hover Lights, 10 KT RT Cross Wind, Gusting to 17 KT. | | | 3.5 | | |
| F | | Night, 1/4 Moon, Single Landing LT, 10 KT RT Cross Wind, Gusting to 17 KT. | | | 5.5 | | |
| G | | Night, Thunderstorm, 20 KT Wind, Gust to 30 KT. | | | 7.5 | | |
| OPERATING STATE: Normal | | | | | CURRENT OBSERVED | PAST OBSERVED | PROJECTED |
| CONFIGURATION: Mid wt, mid C.G., Doors closed | | | | | | | |
| Note 1: Tower obstruction lights, landing pad edge lights. | | | | | | | |

Figure 7. Pilot rating card for flight evaluation of an aircraft.

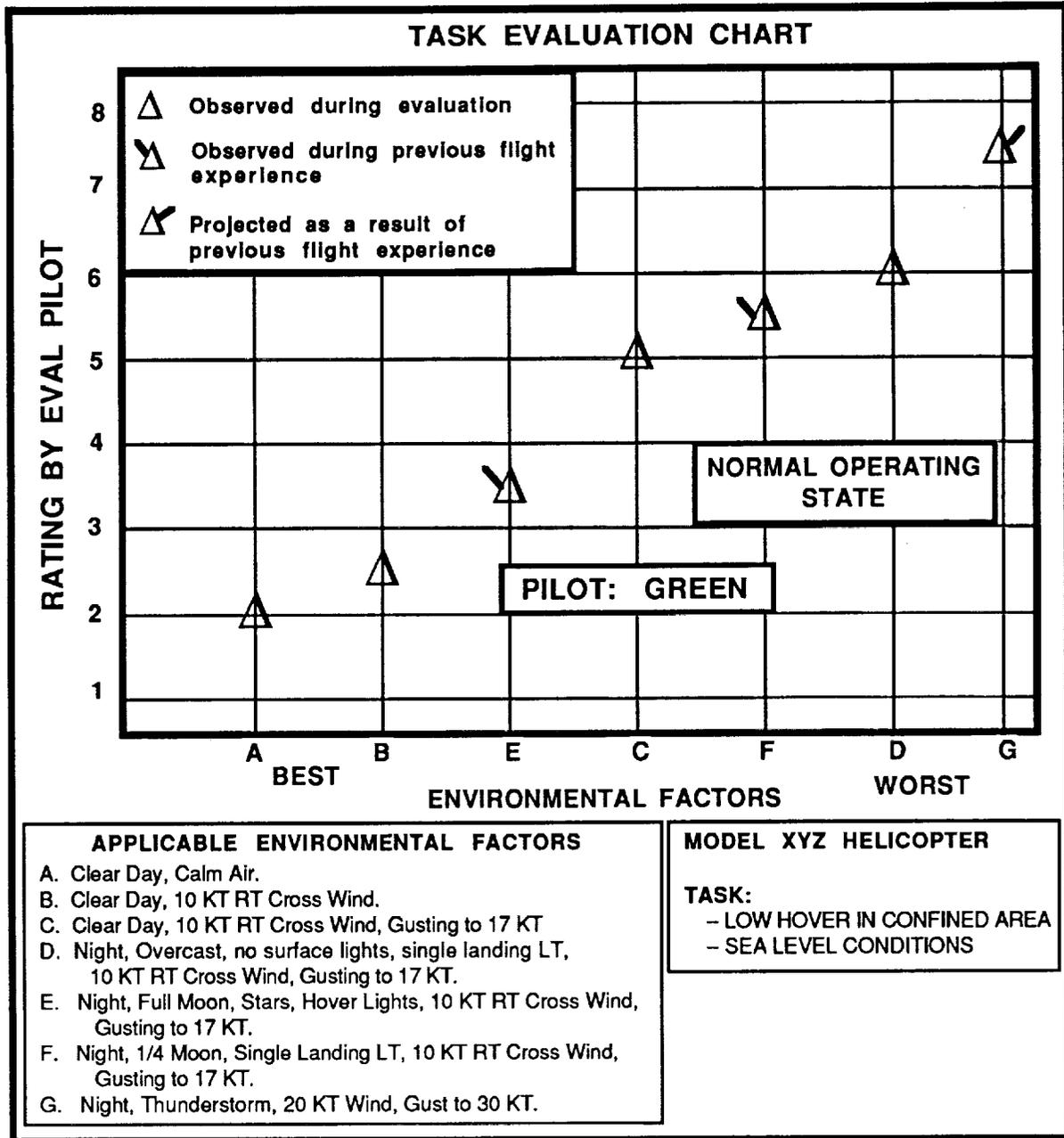


Figure 8. Building a more complete characterization.

Assume that these situations are simulated one by one and that the pilot establishes an assessment (rating) for each and enters this rating on a pilot data card as illustrated in figure 9. Now pilot Green has generated two sets of ratings trying to accomplish the very same task. One set responds to his experience in the real aircraft and one responds to his evaluation of the representation of the aircraft and visual scene provided by the flight simulator. The pilot has in fact rated his ability to achieve a given

task with a specific degree of precision (performance) at a given level of effort. It should therefore be possible to plot both sets of data on one chart to determine the degree to which the data agree or disagree.

This has been done and the results are presented here as figure 10. Figure 10 shows that the three pilot ratings established during "daylight" operations in the simulator are roughly two pilot ratings higher than the trend band which bounds the data defined for flight in the real aircraft

| TASK SHORT TITLE PLATFORM HOVER-LANDING | | PILOT DATA CARD | | SIM FLT <input checked="" type="checkbox"/> | |
|---|--|---------------------------|---------------|---|--|
| | | Pilot Name: <u>GREEN</u> | | A/C FLT <input type="checkbox"/> | |
| TASK: <i>Low hover in confined area. Landing on a platform one hundred feet above a water surface. Obstructions are present ahead and to the right. Upon landing rotor clearance is 30 feet to closest obstruction. Steel structure rises ahead.</i> | | | | | |
| SITUATION ID CODE | FACTORS DEFINING THE TASK ENVIRONMENT SITUATION | PILOT ASSESSMENT (RATING) | | | |
| A | Clear Day, Calm Air. | 4 | | | |
| B | Clear Day, 10 KT RT Cross Wind. | 4.5 | | | |
| C | Clear Day, 10 KT RT Cross Wind, Gusting to 17 KT. | 6.5 | | | |
| D | Night, Overcast, no surface lights, single landing LT, 10 KT RT Cross Wind, Gusting to 17 KT | 6.5 | | | |
| E | Night, Full Moon, Stars, Hover Lights, 10 KT RT Cross Wind, Gusting to 17 KT. | 4 | | | |
| F | Night, 1/4 Moon, Single Landing LT, 10 KT RT Cross Wind, Gusting to 17 KT. | 6 | | | |
| G | Night, Thunderstorm, 20 KT Wind, Gust to 30 KT. | 7 | | | |
| OPERATING STATE: Normal | | CURRENT OBSERVED | PAST OBSERVED | PROJECTED | |
| CONFIGURATION: Mid wt, mid C.G., Doors closed | | | | | |

Figure 9. Pilot rating card for simulator evaluation.

during similar conditions. In contrast, pilot ratings assigned for simulated night operations are in reasonable agreement with the pilot's earlier characterization of the real aircraft.

On first analysis, these data suggest that the pilot found the simulated daylight-visual task to be substantially more difficult than he found the task of operating the real aircraft in the real world. Continuing with this line of thought, the increased difficulty is probably a result of some lack of authenticity in the visual scene. The agreement between aircraft and simulator experience at night suggests the pilots did not detect any shortcoming in the simulator when the simulated scene contained only a modest amount of microtexture. That is, the authenticity of the visual scene became less important during situations in which poor definition was involved.

The evaluation-charting process can be used to evaluate the authenticity of flying qualities as well. The data in

figure 11 provide such an example. The data plot indicates the real aircraft was much more difficult to fly than the simulator. This disagreement in ratings may have been caused by simulator control characteristics (being too good) or by the simulator model being less sensitive to turbulence than it should have been. It is also possible that the wind/turbulence model is in error. Regardless, the data trends are consistent and have meaning.

This process can be repeated for (1) failure modes, (2) tasks that require gross-aggressive maneuvering, and (3) instrument flight where all reference is to cockpit displays. The results should allow the evaluation team to accurately determine the utility of the simulator. Most important, the process will help everyone gain a better understanding of the subject aircraft and of the procedures and techniques pilots employ during its operation. If everyone agrees about the way the aircraft should be flown, and if they all evaluate the simulator using these

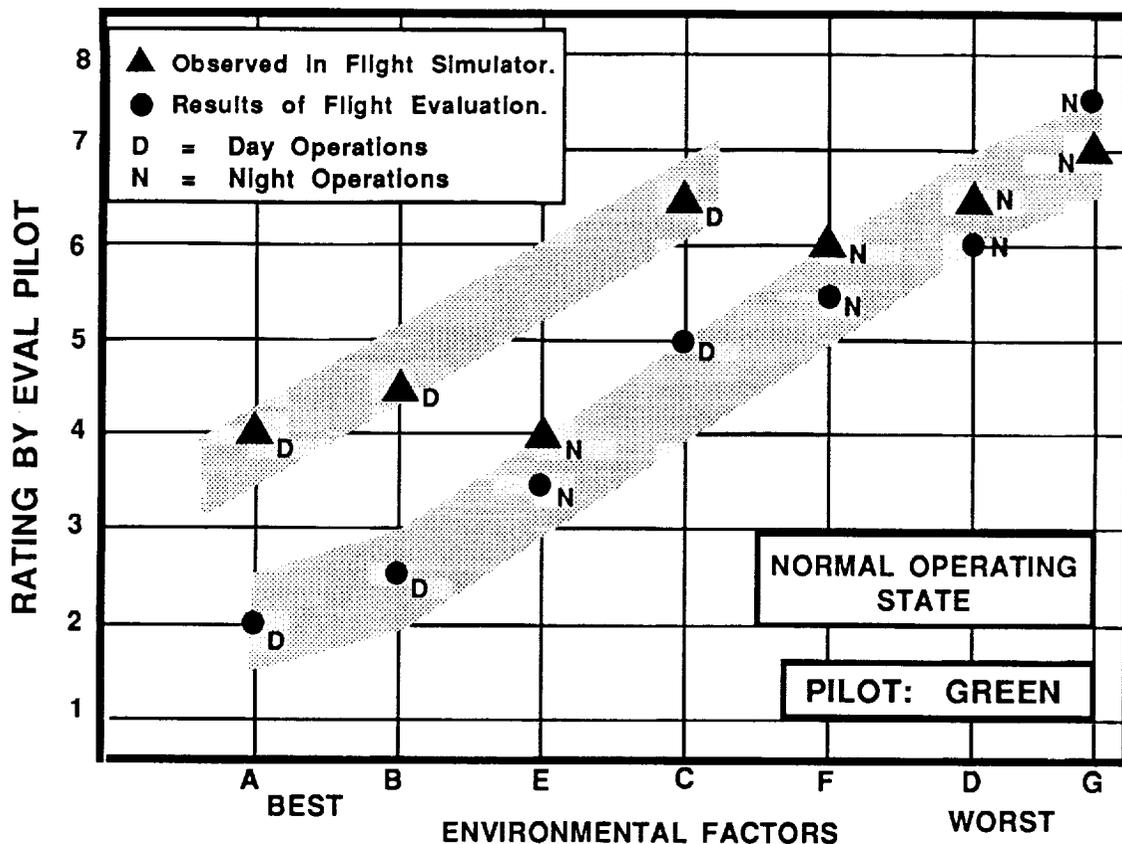


Figure 10. Comparing simulation assessments and real aircraft experience.

common methods, the evaluation will most likely produce results to which most pilot-evaluators will be able to ascribe. Agreement in these areas will help preclude misunderstandings regarding simulator value and applicability.

Finally, charts should be established for a family of flight phases. Failure modes should be examined for each flight phase considered to be critical to the crew training capability of the simulator.

A final set of graphics, figures 12(a) and 12(b) has been included to illustrate how a real pilot evaluated two real but very different aircraft during the accomplishment of a real task. Observe in figure 12(a) that the ratings dropped from 4.5 for C to 4 for D for the single-rotor helicopter, and that there was no change in the pilot's ratings for the tandem-rotor helicopter under these two different environmental situations. This means that, in the case of the single-rotor aircraft, the condition established by C was more stressful than the condition established by D. That is, the crosswind was important to the single-rotor helicopter, but insignificant to the tandem-rotor helicopter. In fact, the loss of the crosswind was more impor-

tant in reducing workload than the loss of daylight was to increasing workload.

Thus the environments should be reordered so that they are progressively more severe from left to right. This has been accomplished in figure 12(b) and the result is a more orderly plot, one which is easier to compare and analyze by the general public.

The scope of this presentation did not allow a complete treatment of the data collection-presentation methods that have been developed by Starmark. I encourage you to tailor and expand the concepts presented here to fit your individual needs.

There are many ways to achieve further reductions in scatter and ways to determine the importance of a given failure mode to the training experience. Many of these additional attributes became obvious to the evaluation engineer as experience is gained during application of the process discussed here.

Everyone who elects to use this material as a guide is encouraged to concentrate on the task of defining the combinations of environmental factors that (1) pilots have personally experienced and that can best define the

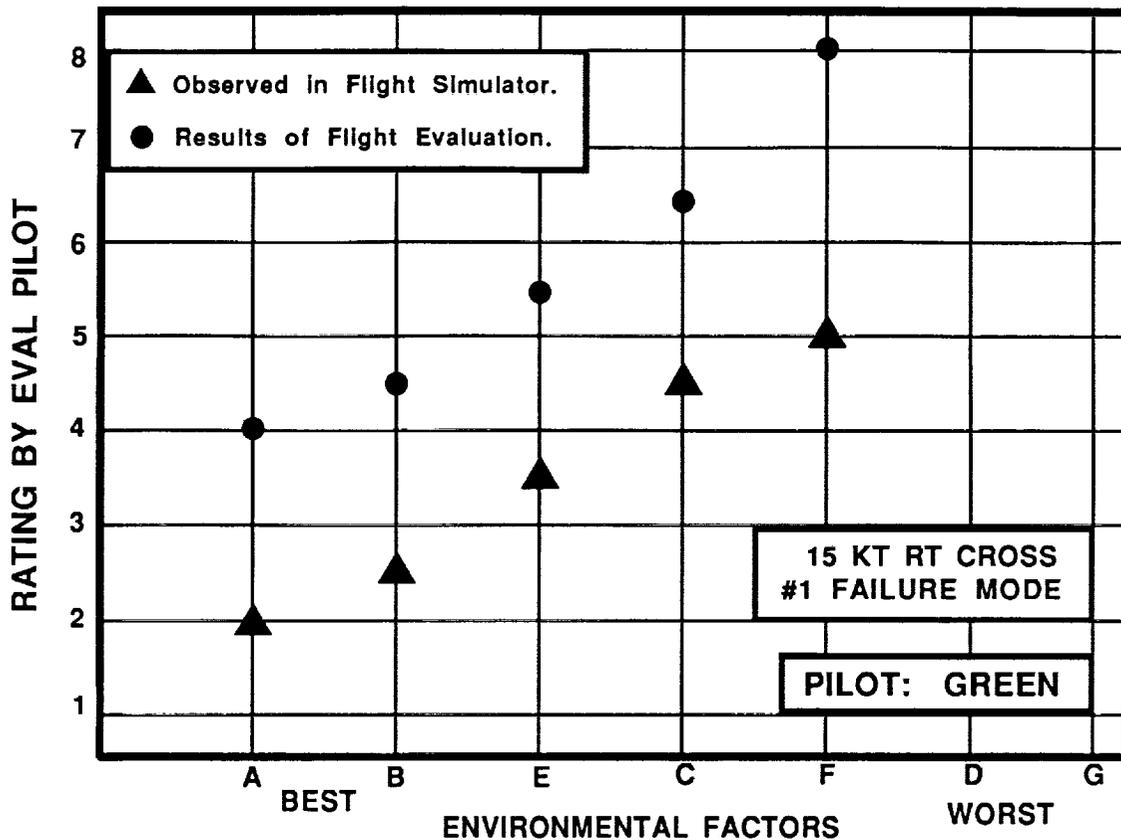


Figure 11. Data indicating that simulator is too easy to fly.

normal operating envelope, and (2) allows pilots to feel they can also best define the extremes of the operational envelope. If the simulation device can provide an adequate, authentic training experience under both situations, the usefulness of the simulator will have been validated in terms of handling qualities and visual scene representations.

MR. WARTH: It is good to see there is a life after flying. How close do the ratings have to be to be considered a good match in the Cooper-Harper figures?

MR. GREEN: I am saying when you write down a definition or expand the definition to meet your needs, just try to keep it in the perspective of Cooper-Harper. There are references that you can use. Did I answer your question?

MR. WARTH: How close do the numbers have to be?

MR. GREEN: You mean scattering of the data?

MR. WARTH: I mean between the simulator and the aircraft.

MR. GREEN: Well, see, that is a whole other discussion. I think just as a very quick answer, that if you could get within a pilot rating and a half, you would think you had died and gone to heaven, and you would want it to be a little more difficult in the real aircraft, I would guess. But what I would do is slip the whole scale to the right. In other words, my visual is wrong. I would say my visual is wrong or something else is wrong, just as long as we don't give the pilot a misimpression of the handling qualities of the aircraft, or misinform him somehow.

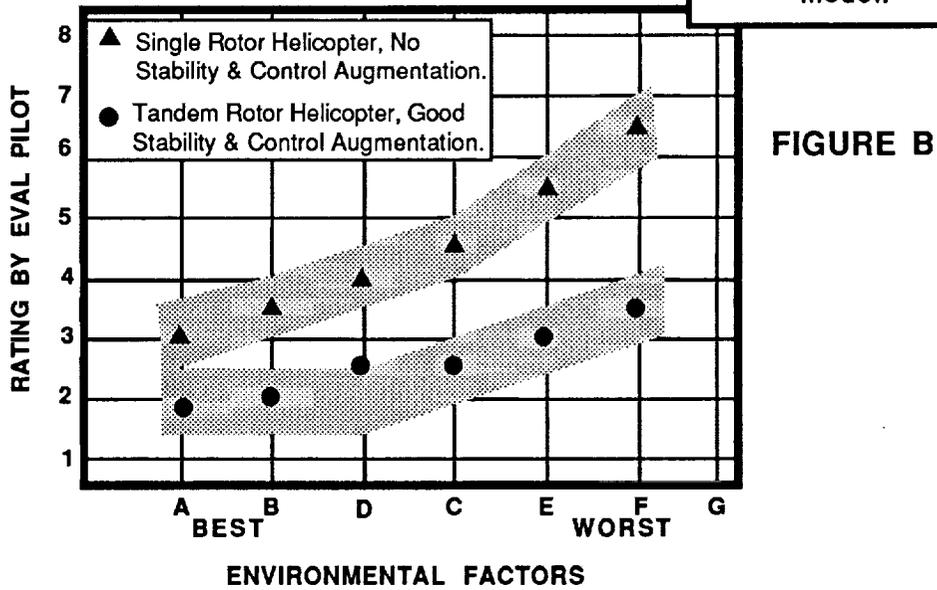
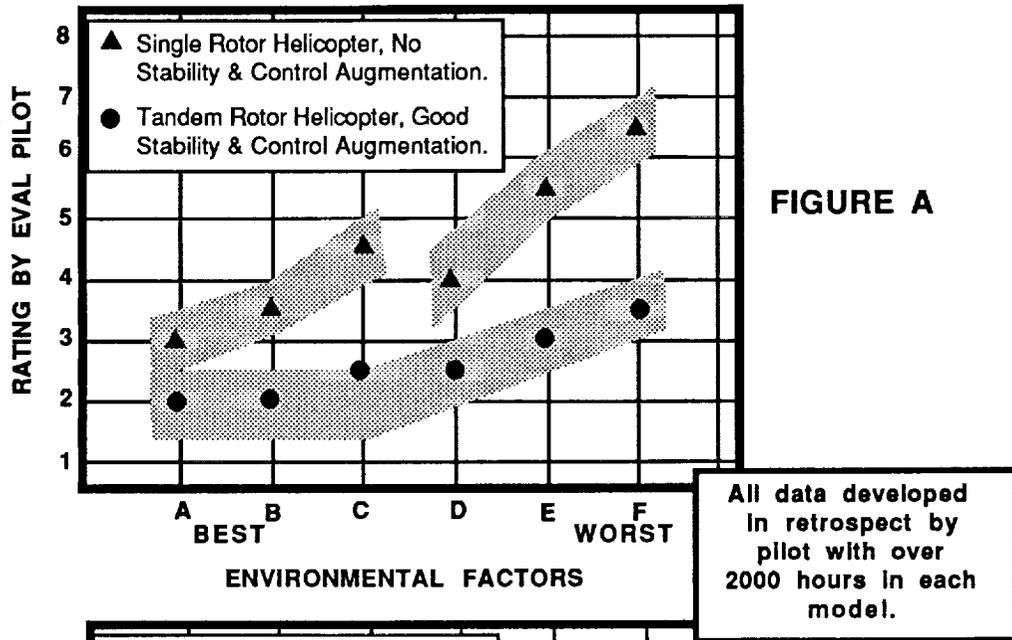


Figure 12. Pilot rating data for single- and tandem-rotor helicopter conducting precision hover. (a) Original sequence of environmental factors, (b) reorder sequence of environmental factors.



David L. Green is president of Starmark Corporation. He is widely recognized as an aviation safety expert and author. Mr. Green is a graduate of the U.S. Naval Academy and of the Navy Test Pilot School, and is the author of the first helicopter stability and control flight test manual. He managed rotorcraft flight tests for Fairchild-Hiller and was vice president of Pacer Systems, Inc., prior to his employment at Starmark. Mr. Green has been involved in FAA rotorcraft evaluation projects since 1968, and has flown 73 models of helicopters, V/STOL, and propeller and turbojet aircraft. He is now an adjunct associate professor of aviation systems at the University of Tennessee.