/N-06
/17609
67P

NASA Contractor Report 194963

Synthetic Vision Display Evaluation Studies

David M. Regal and David H. Whittington
Boeing Commercial Airplane Group, Seattle, Washington

(NASA-CR-194963) SYNTHETIC VISION DISPLAY EVALUATION STUDIES Final Report (Boeing Commercial Airplane Co.) 67 p

N94-37658

Unclas

G3/06 0019609

Contract NAS1-19360
August 1994

National Aeronautics and
Space Administration
Langley Research Center
Hampton, Virginia 23681-0001

CONTENTS

ACKNOWLEDGMENTS

We would especially like to thank our contract monitor, Randall Harris, for his support and helpful suggestions throughout the course of the contract. Rodney Tjoelker's assistance with statistical analysis was very much appreciated. We would also like to express our gratitude to Russel Parrish and Richard Edwards for their reviews of the draft proposal. These added significantly to the quality of the document. And finally we express our appreciation to the pilots who took time out of their busy schedules to serve as subjects and share their ideas with us.

EXECUTIVE SUMMARY

The goal of the present contract was to investigate display requirements for a synthetic vision system to be used on the High Speed Civil Transport (HSCT). We describe a series of experiments that examine the usefulness of certain perceptual cues used in computer generated perspective displays. An extensive background section that sets the stage for this work is also provided.

The background section begins with a brief description the HSCT effort and considers some of the benefits a synthetic vision system would bring to the aircraft. Automation and how it effects the pilots' role in the HSCT is discussed. This is important as it influences the type of synthetic vision system to be implemented. Next, the different technological approaches to synthetic vision are described, including computer generated imagery which is the primary focus of the work reported here. This is followed by a detailed definition of spatial situational awareness since the ultimate goal of our work is to provide pilots with this awareness. Then, a comprehensive listing of the different perceptual cues that can influence spatial perception, and thus situational awareness, is furnished. The present work is directed toward determining which of these cues or combination of cues best communicates desired information to pilots. The background section concludes with a review of relevant literature.

In Experiment I we examined the usefulness of increasing the level of terrain texture in computer generated perspective scenes flown by pilots. As in all the experiments, we measured the ability of pilots to flare and land an aircraft using different display scenes. In this first experiment vertical velocity at touchdown was found to increase with increased levels of texture, while landing distance and flare initiation altitude both decreased as texture was increased. The results are considered ambiguous with respect to the usefulness of adding texture to displays. This finding was unexpected given the conclusions of relevant literature and the strong subjective preference expressed by pilots for the higher levels of texture.

Experiment II was largely an attempt to explain the unexpected results of Experiment I and a previous experiment (Regal and Whittington, 1993) that showed minimal performance advantages from adding enriched familiar size cues to displays. In an attempt to examine the possibility that a ceiling effect was responsible for the previous results, we increased the difficulty of the flying task (and thus the pilot's workload).

Experiment IIA examined the effects of familiar size cues on pilot performance. The results were inconclusive, with an increase in cue strength resulting in harder but shorter landings.

In Experiment IIB the effects of increased display texture were measured. This was a replication of Experiment I under higher workload conditions. Sink rate at touchdown was found to decrease with increased texture, the opposite of Experiment I findings. However, the shorter landings found in the first experiment did not show up here. The results are interpreted as indicating a performance advantage for increased display texture under higher workload conditions.

In Experiment IIC two different high resolution runway texture patterns were compared. There was a small, but significant, subjective pilot preference and higher flare initiation altitude for one of the textures. Otherwise, there were no significant performance differences resulting from the two displays.

Finally, we urge caution in interpreting the results of the present study. The fact that our results show only modest advantages for increased perceptual complexity does not justify the conclusion that this complexity is not important. There are many cues, combinations of cues, display formats, and piloting tasks that were not considered here. The process of discovering the perceptual building blocks that will allow displays to be as effective as natural vision is still in its early stages.

SYNTHETIC VISION DISPLAY EVALUATION STUDIES

David M. Regal and David H. Whittington
Boeing Commercial Airplane Group

INTRODUCTION

Boeing has been exploring the feasibility of incorporating a synthetic vision system (SVS) in the High Speed Civil Transport (HSCT). This system would eliminate the need to droop the aircraft's nose section and potentially provide expanded operations under low visibility conditions. One approach being investigated involves the use of a digitized terrain database from which is generated an out-the-window type scene. This approach provides the designer great flexibility in the design of display images. The increased flexibility results from technological advances that allow the generation of a broad range of high quality images. The question becomes one of determining what information to present, and in what form to present it, in order to best support the needs of the flight crew. The purpose of the present set of experiments is to explore the perceptual requirements necessary if perspective displays are to provide pilots with good spatial situational awareness. We feel it is important that these requirements be established in an empirical fashion and not simply based on subjective opinion. A number of specific areas were examined, including: the effects of display texture density on pilot flare and landing performance, the interaction between workload and perceptual cues for texture density and familiar size cues, and a comparison of two different surface texture patterns.

BACKGROUND

In this section we present background material that sets the stage for the experimental work to follow. We begin with a brief description of the HSCT effort and consider some of the benefits a SVS would bring to this aircraft. We discuss automation and how it affects the pilots' role in the HSCT. This is important as it influences the type of synthetic vision system to be implemented. Next, the different technological approaches to synthetic vision are described, including computer generated imagery, which is the primary focus of the work reported here. After this, we provide a detailed definition of spatial situational awareness since the ultimate goal of our work is to provide pilots with this awareness. Then, a comprehensive listing of the different perceptual cues that can influence spatial perception and thus situational awareness is presented. This is included to help provide context for the cues addressed in the present study. The background section concludes with a review of relevant literature.

THE HSCT AND SYNTHETIC VISION

The feasibility of a new generation supersonic transport is currently being studied by groups within both industry and government. If ongoing studies support its viability, this aircraft could be launched early in the next century. The current Boeing baseline aircraft will cruise at Mach 2.4, be over 300 feet long, have an initial range of between 5000 and 6000 NM, and carry approximately 300 passengers. On routes that are largely over water, it will cut travel time by more than half.

As mentioned, we are presently studying the feasibility of incorporating a synthetic vision system in the flight deck of the HSCT. The original motivation for this effort was our desire to not droop the nose of the aircraft during terminal area operations. The aerodynamics of a supersonic aircraft dictate a long pointed nose section. With the high angle of attack required for landing, the nose blocks the pilots' view of the runway. One solution to this problem, and that adopted by the Concorde, is to swing the nose down out of the pilot's field of view. There are a number of reasons to avoid this approach and build an aircraft without an articulating nose. Primary among these is the desire to save weight by eliminating the mechanism and supporting structure needed to droop the nose. This is especially important in the weight sensitive HSCT where the fuel required per pound-mile is greater than for subsonic travel. It is also desirable to eliminate the mechanical complexity of the drooping mechanism and the maintenance needs and potential for malfunction that accompany it.

A fixed nose configuration will also have some limited benefit in terms of reduced drag. There are two ways this can happen. A drag penalty will be experienced when the nose is in the down configuration, although this is minimal since it only occurs in the terminal area and at low speeds. The second drag advantage could be more significant, coming into play during all phases of flight. A drooped nose aircraft, with its nose in the up position, would still probably have a discontinuity in the external surface of the aircraft to provide some residual forward windows (as Concorde does). If a synthetic vision system replaced all forward windows, then it would be possible to eliminate this discontinuity in the external shape and gain an aerodynamic advantage. However, the inclusion of a limited-view forward window in a synthetic vision aircraft, as has been proposed by some, would likely eliminate this advantage.

Another important factor driving the development of a synthetic vision system is the desirability of enhancing low visibility operating capabilities during periods of poor weather. Weather is the major factor responsible for disruptions in scheduled flights. These delays and diversions produce a serious financial drain on airlines and are an ongoing problem for the flying public. Schedule disruption is especially serious for an HSCT. This is because the major selling point of the aircraft is that it saves travel time. Delays and diversions can quickly erode this advantage. They also result in extra time spent aloft which will disproportionately affect the HSCT with per-hour operating costs higher than those of subsonic aircraft.

AUTOMATION PHILOSOPHY

Since the sole purpose of a synthetic vision system is to provide information to pilots, the system design must be based on the pilots' requirements. These requirements, in turn, depend on the specific role to be played by the crew in operating the aircraft. We discuss automation philosophy at this point because it is important in defining the pilot's role, and thus influences the type of synthetic vision system to be implemented.

There are two general automation approaches for dealing with poor or nonexistent forward vision. One is to automate a large portion of the process and eliminate any involvement on the part of the pilots. The other is to retain the pilot and have a system that includes a combination of humans and automation, including human-centered automation.

It is our belief that it is not presently feasible to fully automate large and/or complex portions of aircraft operations (such as approach and landing) and achieve the levels of reliability necessary in commercial aviation. The key here is reliability. There is a big difference between building a system that works 95 percent of the time and one that experiences a failure only one time in a billion operations. To achieve this high level of reliability the automated system would need algorithms that allowed it to handle the novel situations that invariably occur and that humans are generally very skilled at managing. It is this inability to build systems that can match the human's ability to make decisions when confronted with novel events that is mainly responsible for our belief that final authority for aircraft operations will remain with the pilot in the near future.

Systems that combine humans and automation can take a variety of forms with substantial variations in how work is split between the two (see Riley, 1989; Kantowitz and Sorkin, 1987; Regal and Braune, 1992). The case most relevant to a synthetic vision system concerns a debate over the relative role of autoflight versus manual flight. The question has been raised as to whether autoland should be considered the primary landing system, with pilots playing a backup role, or whether we want to maintain current operating procedure with pilots as primary controllers and autoland used at the pilots' discretion. We argue (given one assumption) that the design of a synthetic vision system interface will not differ much between these two approaches. The assumption is that the quality of pilot performances when acting in the backup mode of the "autoland primary" approach must be as good as when operating in the "pilot primary" mode. That is, the pilot must be performing to the same accuracy and safety levels in both cases. The reason that the design will differ little is that even when called on only rarely, the pilots require the same display information and control capability as they do when performing the task more frequently. They must also be able to achieve the same level of situational awareness to guide their decision making. One can also argue, that with autoland it is still necessary for the pilots to monitor system operations and to do this properly requires displays that provide a full level of situational awareness. On the other hand, there are scenarios in which the above assumption of comparable safety levels will not hold. If the pilot/SVS combination is called on to supplement autoflight in only very rare instances, it might be acceptable to have it operate at a lower safety level, given that the overall system reliability reaches required values.

It should be noted that there is actually a danger in employing automation that takes the pilots out of the control loop during normal operations, turning them from participant into supervisor, but then calls them back when something goes wrong. Evidence suggests that the ability to perform the original control task is diminished in operators who become supervisors. This has been shown to be the case for subjective pilot judgments (McClumpha, James, Green, and Belyavin, 1991) and experimental work in process control (Bainbridge, 1983). Thus, establishing procedural scenarios in which pilots are required to intercede only rarely in the control or decision making process could result in performance decrements. Additional measures may be required to assure that pilots maintain the training and situational awareness necessary to handle these events. This problem is exacerbated by the fact that the pilots will generally be brought back into the process only when there is a problem and the performance demands are high.

The economics of automation must also be considered. New systems should be able to show, via a cost/benefit analysis, that they can buy their way onto the aircraft. A pilot who can do the job just as well as an automated system, and can fit it into his or her workload, is generally very cost effective given that he or she is already present on the flight deck.

SYNTHETIC VISION SYSTEM TECHNOLOGY APPROACHES

Our approach to synthetic vision is to provide pilots with a perspective out-the-window type scene. Perspective displays of this type are highly intuitive, being close to our natural mode of spatial information gathering. They provide advantages such as the minimization of training and the ability to quickly process large amounts of information. We assume that this display would be used in close conjunction with a navigation display, with the perspective display being used in a more tactical, and the Navigation display in a more strategic, fashion. During ground operations, where we are operating in a two-dimensional fashion, the choice between a perspective display and a display showing a plan view is less clear.

Four different techniques for creating perspective displays are under consideration. The first involves computer generated imagery (CGI) in which displays are generated from an on-board digitized terrain database. The major advantage of this approach is the great design flexibility it affords. The ability to display a wide range of information in a variety of ways means that displays can be tailored to fit the specific perceptual and information needs of the pilots. This approach benefits from recent technological advances. Terrain databases for most of the world are either available or soon will be, and the storage of this large amount of information is now a straightforward matter. With data compression techniques, the entire continental United States is now available on a single optical disk. High resolution color display technology is also advancing rapidly and we are confident that the large format cathode ray tubes (CRTs) we presently use will be replaced with flat panel displays in time for the HSCT. The recent advances in graphics processors are impressive with improvements expected to continue in the future. The power to run high quality interactive displays is currently available and we expect no problem in procuring highly capable flight-ready systems when needed.

The major disadvantage of the computer generated imagery approach is that the display is limited to information in the database. Transient events, such as another aircraft encroaching on the active runway are not depicted. For this reason it is probable that CGI will be used in conjunction with real time sensor information. A CGI system also requires the development of terrain and feature (including manmade features) databases and the regular updating and distribution of this information to the airlines. Certification of these databases must also be accomplished.

The second approach for displaying perspective information is to use imaging sensors [e.g., millimeter wave radar (MMWR); infrared (IR); light detection and ranging (Lidar); high definition television (HDTV)]. The primary advantage here is that the images displayed are real-time representations of the outside world. There are a number of disadvantages, however. The sensors that produce better image quality are those least capable of penetrating weather, and those that can see through it provide poor image quality. There are also the problems of variations in image quality with changes in environmental conditions, the potential for image artifacts, reliability questions, and hardware placement issues.

The third approach is that of image fusion, in which information from more than a single source is combined. The merging of information can be very simple, such as the simultaneous presentation of different images, or it may be complex, such as when the most important information from different images is recognized, extracted and then combined into a single hybrid display that is superior to any of the individual images contributing to it. It is our feeling that a final synthetic vision system is likely to involve some form of image fusion.

The fourth approach involves direct-view optical devices such as a periscope or an external mirror. This approach has the advantage of apparent simplicity but suffers from a number of disadvantages. It does not provide low visibility performance and suffers field-of-view problems resulting from the difficulty of looking around the aircraft's nose section.

SPATIAL SITUATIONAL AWARENESS

We have stressed that good situational awareness is necessary to support the pilot's decision making process. Since this contract deals with spatial situational awareness, which is part of overall situational awareness (see Regal, Rogers, and Boucek, 1989), it is important to understand just what we mean by this term.

A synthetic vision display, like flight deck windows, provides the pilot with near-term information about his or her environment. Note that we are not concerned with the more strategic aspects of long-term route planning at this point. The SVS should provide information that supports guidance and control of the aircraft in its immediate environment (see Abbott, 1993). This includes providing information that supports situation monitoring, situation assessment, decision making, and the execution of these decisions in the spatial environment.

Consider now the elements that make up spatial situational awareness. First, it is necessary to know the relative location of objects within the environment. This includes the position of our own aircraft (including attitude), the location of terrain and man-made obstacles on the ground, the position of other aircraft, and the identification of necessary reference points (e.g., the airport or other useful landmarks). The location of atmospheric conditions such as clear air turbulence, microbursts and thunderstorms would also be included in this category, as they can be considered to be obstacles that need to be avoided.

Second, given the current aircraft parameters (direction of motion, velocities, and accelerations) it is necessary to know how the position of our aircraft will change over time with respect to other objects in the environment. For example, we would like to know if we are on a collision course with anything in the environment and how much time remains before impact given no modification to our route.

Third, we need to know where in space we want the aircraft to be at different points in the future. This will generally require knowledge of an envelope in space that defines an acceptable range of positions. This envelope can be small, requiring precise aircraft positioning, or large, allowing the aircraft considerable freedom. It may be more or less well defined depending on the situation and how far into the future we are looking. The consequences of breaching the borders of the envelope may also vary. Accurate knowledge of envelope parameters is also part of situational awareness and can influence pilot performance and impact workload. Knowing the aircraft's present and predicted position along with its desired position provides the error signal that guides control inputs.

A fourth requirement is knowledge of the spatial performance parameters of the aircraft (e.g., its maneuverability and ability to change speed). This knowledge adds a necessary dimension to our understanding of the spatial environment, in that it allows us to scale distances in terms of our ability to position the aircraft. As an example, consider how differently the paddler of an ocean kayak and the captain of a super tanker view the significance of distances and objects in their environment. There are two aspects of an aircraft's performance parameters that it is desirable for a crew to understand. One is the desirable performance range, based on considerations such as passenger comfort, company policy, and regulations. The other, employed only very rarely, involves knowledge of the aircraft's outermost performance envelope. If required in an emergency, the pilot should be aware of the maximum performance that can be achieved.

CUES TO SPATIAL PERCEPTION

The purpose of this contract is to help determine which perceptual cues are important in the generation of computer generated out-the-window type scenes. To this end, it is useful to list and give a brief description of potential cues. These have been divided into three groups: static cues relevant to VDT (video display terminal) displays, dynamic cues relevant to VDT displays, and cues not relevant to VDT displays.

Static Cues.

The following cues provide spatial information to an observer viewing a perspective scene displayed on a VDT, and do not require dynamic information.

Familiar size. The visual angle subtended by an object of known size provides information regarding the absolute distance of that object.

Relative size. When objects are assumed to be of similar size, smaller objects are perceived as being further away.

Linear perspective. Parallel lines (such as runway edges) appear to converge as their distance from an observer increases. This convergence is a cue to distance. The angle formed by converging parallel lines is called the splay angle.

Texture gradient (Detail perspective). Texture in a uniform texture field grows finer and more compact as the distance from the observer increases. The relationship between texture elements in a uniform texture pattern follow all the same Euclidean principles that govern linear perspective. Gibson (1950) believed textural information an important cue for distance perception.

Interposition. An object that obstructs our view of a second object is perceived closer than the second object.

Shading. Shading can play a significant role in the perception of object shape.

Shadows. Shadows can play an important role in revealing the shapes and position of objects in our environment.

Aerial perspective. As objects become more distant haze tends to make contours more blurred. Objects tend to lose their coloration and become more bluish as they move further away. These cues are commonly used by artists in depicting distant scenes.

Relative brightness. The luminance of a lighted surface or object does not decrease with distance unless it is attenuated by particulate matter in the air (aerial perspective), with more distant objects being more affected. Thus, one tends to see the brighter of two as closer (given the presumption of equal brightness when viewed from equal distances).

Distance from horizon. Objects that appear closer to the horizon are often perceived as being further away.

Stereopsis. Our two eyes, viewing the world from different locations, see slightly-different views of an object or scene. Stereopsis is the sense of depth we have when the visual system resolves these different views into the appropriate three-dimensional representation. Stereopsis is a useful depth cue up to about 30 m (Arditi, 1986). Special equipment is needed to generate a stereoscopic display using a VDT.

Dynamic cues.

The following cues provide spatial information to an observer viewing a perspective VDT display, but require some form of display movement.

Motion parallax. When we move through space, static objects in the environment that are at different distances from us appear to move relative to one another. This is a cue to the relative distance between objects. If we already know the relative distance between objects, motion parallax can provide information about flight path. For example, if we want to know if an obstacle will be cleared, a comparison can be made of the relative motion between the top of the obstacle and its immediate background. If the background is moving up relative to the obstacle, the obstacle will be cleared (Langewiesche, 1944)

Kinetic depth effect. The three dimensional nature of some objects (often represented by impoverished cues) is not correctly perceived by observers until the objects move in such a way as to reveal the relationship between parts of the object.

Flow rate. The angular speed of terrain elements as one passes through a visual scene can provide spatial information about our environment. Flow rate increases as speed increases or as altitude decreases (Foyle, Kaiser, and Johnson, 1992).

Texture rate (edge rate). Texture rate is a measure of how many texture elements pass out of a scene per unit time (Foyle, Kaiser, and Johnson, 1992). This provides information on ground speed, but depends on fairly homogeneous spacing of texture elements.

Point of symmetrical expansion. A point about which a flow field is expanding when viewed from a moving vehicle indicates the vehicle's aim point. For example, if on final approach the intended landing spot moves downward in the field of view (with attitude held steady) the airplane will overshoot the spot.

Non-VDT cues.

The perceptual cues listed below have meaning when we view our natural environment, but are not applicable to a perspective scene displayed on a two-dimensional VDT screen.

Accommodation. The curvature of the eye's lens changes with the distance of the object being fixated to achieve sharp retinal focus. Accommodation can provide useful, if not very precise, information on distance between about 20 and 300 cm (Hochberg, 1971).

Convergence. The angle between the directions of gaze of the eyes when we fixate a point in space has the potential to provide distance information. There is some question as to how useful and reliable this cue is (Hochberg, 1971). Recent evidence indicates that convergence and accommodation together can provide accurate absolute distance information (Morrison and Whiteside, 1984). In any case, convergence will not provide any distance information for objects further away than about 10 ft.

LITERATURE REVIEW

We are interested in knowing which perceptual cues, when used in aircraft displays, will most enhance a pilot's spatial situational awareness. In this section we review those experimental studies that address this question within a context representative of aircraft operations.

The perceptual cues of interest are those that will provide an observer viewing a two-dimensional perspective display an understanding of the three-dimensional environment in which he or she is operating. These cues to spatial perception are listed and defined in the previous section. While there exists an extensive literature on spatial perception (see Boff, Kaufman, and Thomas, 1986; Carterette and Friedman, 1975; Kaufman, 1974; Graham, 1966), its focus is largely on the basic science of human vision, and while supplying a necessary theoretical framework, it does not provide us with sufficient information to design aircraft displays. Again, we review only that research dealing directly with the design of perspective displays for aircraft.

The quality and relevance of the existing studies vary greatly. Many different measures of subject performance have been employed. These range from the estimation of spatial parameters (e.g., heading or altitude) after passively viewing a display, to the empirical measurement of performance during an active flying task (e.g., flare and landing). There is also a wide variation in the quality of displays employed. These can range from simple grid or outline patterns to highly realistic looking scenes. Subject skill level varies from the ubiquitous college sophomore with no flying experience, to highly experienced professional pilots.

Increased scene complexity.

A number of studies indicate that an increase in scene complexity can result in improved pilot performance.

Barfield, Rosenberg, and Kraft (1990) showed subjects a video presentation of three computer generated approach to landing sequences. The three scenes consisted of a runway surrounded by a fairly homogeneous flat surface, farmlands, and farmlands with hills. Subject judgments, made during pauses in the video presentation, indicated that increased scene complexity allowed superior estimates of altitude and aim point.

Lintern, Thomley-Yates, Nelson, and Roscoe (1987) evaluated subject performance using three different scenes: multicolored fields with groups of buildings and isolated buildings (cubes), a white grid on a green background, and a river valley with buildings and 4000 ft. mountains on either side. The subject's task was to maintain a specified flight path and then dive from 8000 to 3000 ft. at a 30 degree. angle as if on a bombing run. Aircraft pitch and altitude errors were measured. Findings indicate that training and transfer of training were better with the more complex scenes than with the grid pattern.

Lintern and Walker (1991) examined the effects of variable scene content and familiar size cues on the performance of pilots flying landing approaches using a computer generated display. Scene content and runway breadth were varied. One scene consisted of sky, ground, and a runway with a centerline and aim point. The second scene divided the ground up into a small number of colored areas and added irregularly spaced cubes and cones to represent buildings and trees. An extended runway centerline was added. Runway widths of 18, 31, and 42 m were used. Approaches were found to be lower with reduced scene content or with narrower runways. Trial-to-trial variability was higher with reduced scene content.

Reardon (1988) showed subjects displays that consisted of a wire frame runway either unfilled, filled with a random dot pattern, filled with a grid pattern, or filled with a more complex pattern of X's. Subjects passively viewed a computer generated approach and made an estimate of the aircraft's aim point when the image was frozen at an altitude of 50 ft. Results indicted that estimates of aim point improved with increased scene complexity. The experiment also examined the effect of nested texture on pilot performance (i.e., additional texture was added as altitude decreased). Subjects showed no difference in performance when viewing displays with and without nested texture.

Buckland (1980) measured vertical speed at landing for different computer generated scenes. Four scenes had a grid pattern superimposed on the touchdown zone area with grid sizes of 4, 8, 16, and 25 feet. A runway with standard markings and a bare runway with only a centerline were also employed. A night runway scene was the final display. Twelve experienced pilots landed a simulated T-37 aircraft. Performance was best for the displays containing grid patterns with the more detailed grids producing the best performance. Next best performance was with the runway with standard markings followed by the bare runway. The night scene produced the poorest results. While the best performance (2.45 ft/sec) was with the smallest grid pattern, this was not as good as actual flying performance which was given as 0.53 ft/sec.

Kleiss and Hubbard (1993) looked at the effects of three types of computer generated scene detail on the ability of pilots to detect changes in altitude. They presented different densities of vertical objects, varied the realism of the presented objects, and examined the effects of texture mapping. Results indicate that the speed and accuracy of detecting altitude changes improved with an increase in the density of vertical objects. Adding detail to individual objects to increase their realism produced no consistent improvement in performance. Texture mapping added to the terrain surface proved more effective than when added to individual objects. Texture mapping, while effective, was not as beneficial as object density under the testing conditions used.

Lintern, Sheppard, Parker, Yates, and Nolan (1989) ran pilots on a transfer of training task involving bombing accuracy. In one part of their experiment they evaluated performance using more and less complex scenes. Results showed no differential transfer among the conditions.

Regal and Whittington (1993) examined the effect of varying the level of familiar size cues in a computer generated scene. Experienced pilots performed approach, flare, and landing maneuvers using three different levels of cues: a plane runway, a runway with standard markings, and a runway with standard markings and general aviation aircraft parked to one side. Performance measures showed no difference in vertical touchdown velocity between the three conditions. There was a tendency toward lower flare initiation altitudes and shorter landings when using the two displays containing a higher level of familiar size cues. The pilot's subjective reports indicated a strong dislike for the plain runway as compared to the runways with standard markings.

Display resolution

Mann (1987) looked at the effects of different resolution levels of MMW radar on pilot performance. The display resolution was varied by presenting three different vertical line counts (400, 174, or 47 lines). Different resolutions did not significantly effect pilot touchdown performance but did tend to cause a greater variability in flight path control during approach. There was also a trend toward an increase in subjective workload with decreasing resolution. The author believes that there is a tendency toward decreased performance with decreased resolution, but acknowledges that this is not conclusive finding.

Splay angle and texture density.

A number of studies have examined the effects of three specific display conditions on the ability of subjects to maintain a constant altitude. These conditions provide representations of the ground plane with (1) a set of parallel lines oriented parallel to the direction of motion, (2) a set of parallel lines perpendicular to the direction of motion with the distance between lines following the rules of perspective, or (3) a grid pattern made up of a combination of the first two patterns. The first display condition relies on changes in splay angle as a perceptual cue, the second on changes in texture density.

Wolpert, Owen, and Warren (1983) had subjects make passive judgments regarding loss of altitude using these three display conditions. Detection of altitude loss was better when a descent event took place over texture consisting of stripes parallel to the direction of travel than when the stripes were perpendicular to the direction of travel. Detection performance was intermediate when the grid pattern was used.

In a study that found results that differed from those reported above, Johnson, Tsang, Bennett, and Phatak (1989) measured the ability of subjects to maintain altitude when actively flying over the same three ground texture patterns. Performance was poorest when flying over a parallel texture pattern. The results for the other two patterns were similar to one another. Johnson, Bennett, O'Donnell, and Phatak (1988) found similar results when examining a helicopter hover task. The authors suggest that the difference between these findings and those of Wolpert, Owen, and Warren (1983) may be due to the use of local optical elements in the earlier study. These include display screen/display element intersection location and angle. The Johnson et al. study eliminated this potential artifact by including lateral wind disturbances that varied the aircraft's lateral position.

Wolpert (1988) added disturbances in altitude and roll to test the possibility of local edge artifacts influencing performance. His results, using the same three display conditions, confirmed the superiority of splay angle as a cue for altitude regulation.

Flach, Hagen, and Larish (1989) used the same stimuli as Johnson et al. (1988, 1989) and also found that splay angle (parallel grid pattern) produced better altitude control than a horizontal grid pattern. However, the authors point out that the optical flow rate (ground speed divided by altitude) used by Johnson et al. (1988, 1989) was less than that used by the other investigators. When Flach, Hagen, and Larish (1989) used a slower flow rate (although not as slow as that used by Johnson et al.) they found that the superiority of the parallel line display was significantly less than at higher rates.

Kelly, Flach, Garness, and Warren (1993) conducted still another study using these three stimuli. In an effort to determine whether the controversy over the superiority of vertical vs. horizontal texture was due to the different optical flow rates used in the different studies, they tested four different optical flow rates that covered a range inclusive of those used in the previous experiments. With subjects controlling aircraft altitude in the presence of pseudo-random wind disturbances, performance was superior for the texture pattern parallel to the direction of flight and the grid pattern at all flow rates Vs. the horizontal pattern. The error rate for the horizontal pattern did, however, decrease as the flow rate became lower.

In a related experiment Warren and McMillan (1984) had subjects maintain altitude in the presence of wind gusts. Stimuli were a roadway extending to the horizon, a random dot pattern, and a combination of the two. Best performance was found for the display containing only the roadway (splay information). Note that the addition of random dots to the roadway produced a decrement in performance over the roadway only.

Warren (1988) measured the ability of student subjects to maintain various altitudes during high speed, low altitude flight while viewing one of three displays: a roadway, a perspective view of a field of random dots, and a combination of the two. Performance was, in general, best with the roadway. The author concludes that rich cue, high-fidelity scenes may not be best for specific training on specific tasks, that the presence of irrelevant cues can interfere with performance, that different cues for the same referent are not all equally effective, and that the cue dominance hierarchy must be determined empirically.

A number of the above studies (Wolpert, Owen, and Warren, 1983; Warren and McMillan, 1984; Warren, 1988; Wolpert, 1988; Flach, Hagen, and Larish, 1989; Kelly, Flach, Garness, and Warren, 1993) indicate that increased scene complexity does not always result in improved performance on tasks involving judgments of altitude. This is true even when the simpler but superior display is a sub-element of the more complex display. It is also the case that the complex displays used in these studies do not subjectively appear to be cluttered displays. Kelly et al. (1993) speculate that the decreased performance may result from a decrease in the visual signal to noise ratio. The complex displays have information related to forward motion and wind buffeting (visual noise) in addition to information specific only to altitude change (signal).

Motion parallax vs. flow patterns.

Kaiser, Perrone, Andersen, Lappin, and Proffitt (1990) looked at cues for terrain slant perception. Surfaces were defined by point lights with a uniform random distribution, with slopes ranging from 15 to 120 degrees. Observers were better at estimating slope when they were traveling parallel to it than when moving toward it or viewing a static display. Performance was equivalent for the static scene and when moving toward the slope. The authors interpret these findings to indicate that motion parallax has greater utility than differential optical expansion rates.

Accommodation.

Parrish, Kahlbaum, and Steinmetz (1979) examined the effects of varying the accommodative distance of different parts of a display screen containing a simple computer generated scene. The pilot's ability to flare and land was measured. In one case the center of the display screen was at an accommodative distance of 4.9 m and the top and bottom of the screen at infinity. In the comparison case, the display screen was tilted so that the bottom of the screen appeared at an accommodative distance of 9.1 m and the top of the screen at infinity. This second case is much closer to representing the accommodative conditions of the real world. Objective data indicated no significant difference in performance between the two conditions.

Stereoscopic displays.

Andre and Johnson (1992) looked at the effects of different types of visual scene information on precision rotocraft hovering tasks under stereo and biocular viewing conditions. Three different scenes were used: a group of buildings, a single tree, and an open field, in addition to two levels of ground texture, a grid pattern and a no-texture condition. Performance was best when vertical cues (buildings and trees) were present. The grid texture aided hover stability only when other position-reference cues were absent from the scene. The stereo viewing condition showed no uniform advantage over the biocular condition. It should be noted that the generalizability of results from the domain of rotocraft hover operation to that of winged aircraft is not established.

A number of studies examined the effects of stereo viewing on a pilot's ability to fly a pathway-in-the-sky that was part of a perspective display. Nataupsky and Crittendon (1988) evaluated two different pathways, a simple monorail connected to the ground by posts and a pathway made up of rectangles, each mounted to the ground by a post. Stereo viewing improved responses with the monorail, but provided no benefit when the more complex tunnel pathway was flown. Reising, Barthelemy, and Hartstock (1989) had subjects fly a tunnel-in-the-sky incorporated in perspective display with numerous other non-stereo depth cues present. Results indicate that the addition of stereo viewing provided minimal benefit over the non-stereo condition, although there was a small enhancement when a more difficult flight path was flown. Way, Hobbs, Qualy-White and Gilmour (1989) and Way (1989) had pilots fly air-to-ground combat scenarios using a perspective display that included a pathway-in-the-sky. These investigators did not find any advantage for stereo viewing in these full mission simulations. Wickens, Todd, and Seidler (1989) hypothesize that in a dynamic environment, stereoscopic viewing can sometimes be beneficial, but primarily for its compensation for the absence of other pictorial and motion cues, rather than for any enhancement of the effectiveness of well implemented cues that are already present.

Zenyuh, Reising, Walchli, and Biers (1988) looked at the detection of aircraft presented in a perspective display with only impoverished two-dimensional depth cues. Yeh and Silverstein (in preparation) had subjects evaluate the spatial location of two objects suspended in space in static perspective display. In both studies stereo viewing enhanced pilot performance. Consistent with the Wickens, Todd, and Seidler (1989) hypothesis this positive effect of stereopsis may be the result of the non-dynamic displays and the scarcity of other depth cue information.

Velocity cues.

A number of studies have examined velocity perception during self-movement. Two sources of information have been proposed as contributing to the perception of ground speed, global flow rate and edge rate (Denton, 1980). Global flow rate reflects the angular speed with which ground texture passes a fixed reference angle relative to the observer. It varies directly with ground speed during level flight, but is disrupted by a change in altitude. Edge rate is defined by the number of texture elements being passed per unit time. It is not affected by altitude, but does vary directly with ground speed as long as there is no change in ground texture density. Investigators have examined the role of each of these cues on the perception of velocity.

Owen, Wolpert, and Warren (1983) examined subjects' sensitivity to displays showing changing speed during straight and level flight over flat terrain. Subjects were sensitive to both flow and edge rate changes but with much higher sensitivity to changes in edge rate. Larish and Flach (1990) confirmed that edge rate was a more powerful cue than flow rate. They showed that flow rate, while an inferior cue, did become relatively more useful as surface texture was reduced.

Awe, Johnson, and Schmitz (1989) examined subjects' ability to selectively attend to flow rate and edge rate while controlling their forward speed. They found that subjects persisted in using edge rate as the basis for speed control even when instructed to use flow rate and given feedback about their success in using it. The authors suggest that the use of edge rate information corresponds to a relative inflexibility in selectively attending to information for self-speed.

EXPERIMENT I - THE EFFECT OF DISPLAY TEXTURE ON PILOT PERFORMANCE

RATIONALE - EXPERIMENT I

As stated previously, the goal of our research effort is to design pictorial displays that best support pilot performance. We want to help determine what perceptual information to include in out-the-window type perspective displays, and discover how to best format this information. In a previous experiment (Regal and Whittington, 1993) we examined the usefulness of familiar size cues. To our surprise, this experiment revealed only limited improvement in pilot performance as the salience of this cue was increased across a series of displays. This was despite an increase in the pilot's subjective preference for the richer cue displays. In the present experiment we expand on our original study by examining a different perceptual cue, that of texture, considered to be an important cue for achieving good spatial perception (see Gibson, 1950).

METHODOLOGY - EXPERIMENT I

Displays.
Four different display conditions containing progressively greater amounts of texture information were flown by pilots. A no texture condition (Figure 1) contained a gray runway, 150 ft. wide and 8500 ft long, on a homogeneous green surround with a blue sky. The next condition had a texture pattern added to the runway surface (Figure 2). In the third condition, a texture pattern was also added to the area surrounding the runway (Figure 3). The fourth condition was the same as the third but with the addition of standard runway markings (AIM 1991) (Figure 4). Three dots to either side of the runway indicated the pilot's aim point (glideslope intersect, 1000 ft. from threshold). Thus, all display conditions, those without as well as those with runway markings, provided the pilot with an aim point. It was felt that to omit the aim point for some conditions would have added a new variable (aim point vs. no aim point) unrelated to the texture variable of interest and possibly have contaminated the results.

Apparatus.
Scenes were generated on a Silicon Graphics 320 VGXT workstation with an update rate of approximately 20 Hz. Screen resolution was 1280 by 1024 pixels. The subjects were positioned 20 in. from the display screen which resulted in a field of view of 40 degrees horizontally by 30 degrees vertically. A Measurement Systems Inc. sidestick controller was used to control flight path. No throttle or rudder pedals were available. A high fidelity 737-300 airplane model was used. This simulation included all code used for pilot testing in full mission cab simulations. It had all the aerodynamic effects used for flight test matching, including ground effects near touchdown. The aircraft's approach speed was 140 kt. and bled off in accordance with autothrottle procedure, with the engines going to idle over a period of six seconds beginning at an altitude of 27 ft. The engine model had all the appropriate spool up/down characteristics. Since our interest was in the evaluation of perceptual cues in a computer generated scene, and not aircraft performance characteristics, the use of a 737 airplane model rather than one for the HSCT should not adversely influence results. If anything, the use of a model more familiar to our pilots would eliminate the added variability that would usually accompany the operation of an aircraft on which they were less well trained.

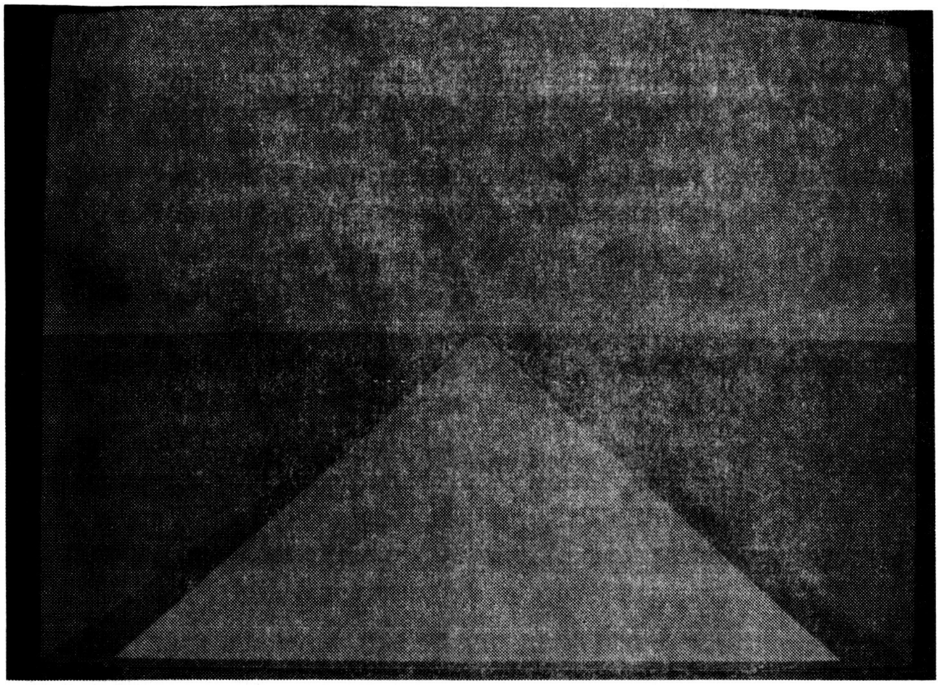

Figure 1. Display condition 1. Gray runway (150 ft. by 8500 ft.) with green surround and blue sky. The six white dots indicate the glideslope intercept (1000 ft. from threshold).

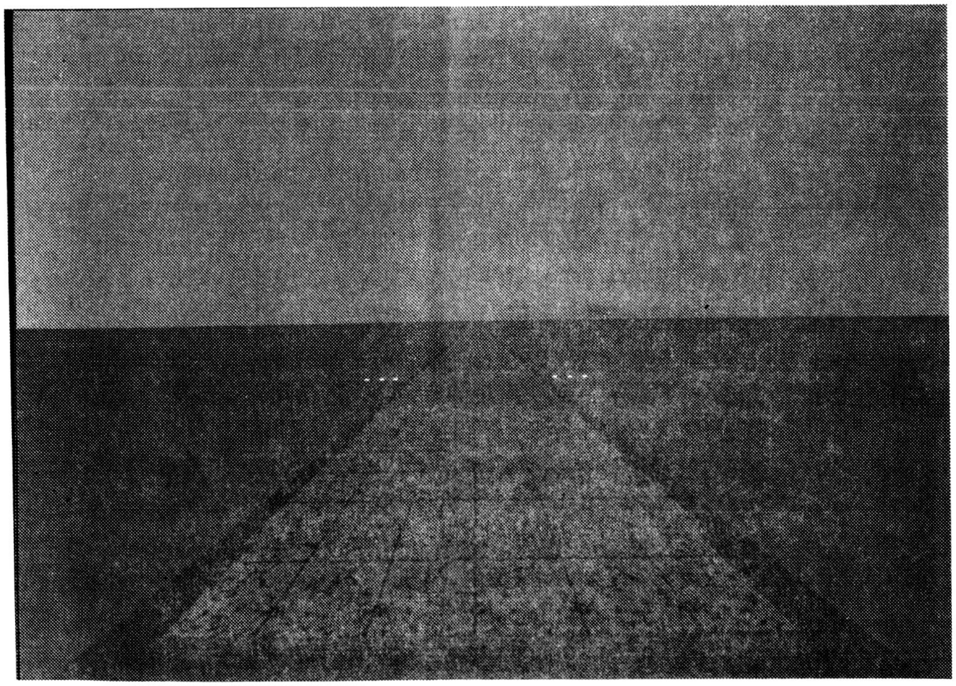

Figure 2. Display condition 2. Same as Figure 1, but with the addition of a texture pattern to runway.

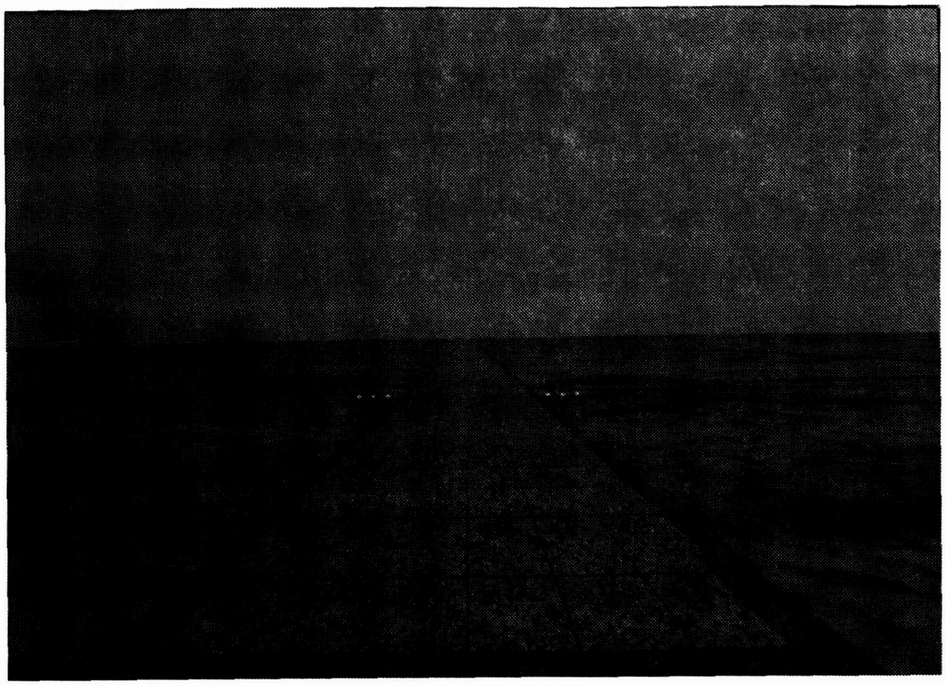

Figure 3. Display condition 3. Same as Figure 2, but with addition of texture pattern to surround.

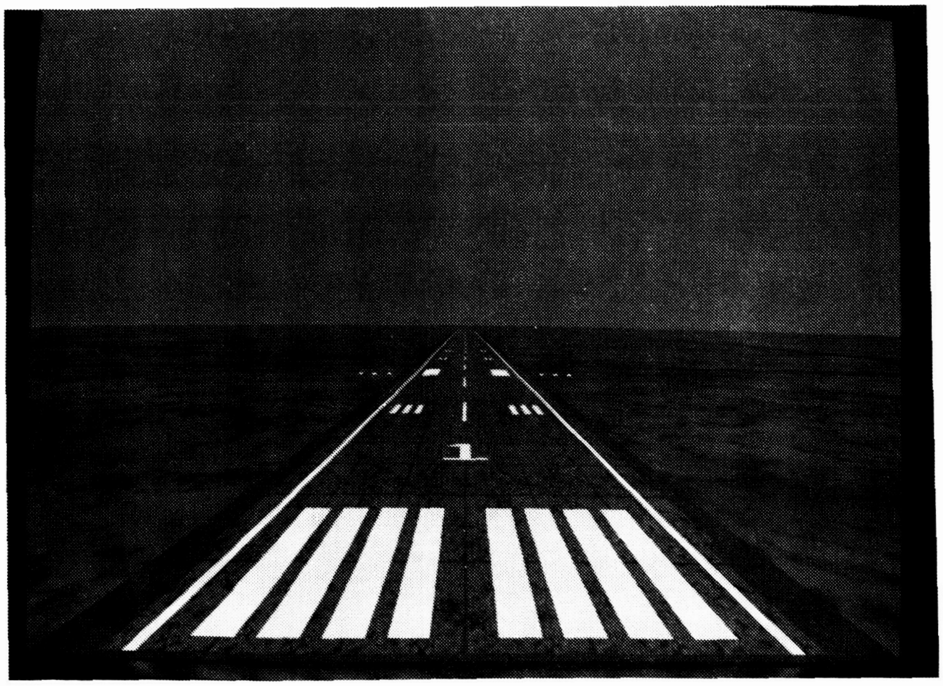

Figure 4. Display condition 4. Same as Figure 3, but with addition of standard runway markings.

Subjects.

Six subjects were tested. Five were Boeing pilots, one a general aviation pilot. All were experienced aviators active in their flying careers at the time of testing.

Procedure.

Each experimental trial began with the aircraft initialized on glide slope approximately 20 seconds from touchdown. It was the pilot's task to maintain glideslope and flare and land the aircraft. Pilots controlled the aircraft in the vertical dimension (through pitch control) but did not have control over lateral position. Lateral control was not provided because it significantly increased the difficulty of the flying task. It was felt that this might mask our attempt to differentiate between perceptual cues by having differences in manual control ability add variability to the results. (However, after examining the results of the present experiment it was decided to include lateral control in Experiment II. The reasons behind this decision are explained in the description of Experiment II.)

Each trial was started with the aircraft's lateral position aligned with the runway centerline or 2,4,6,8, or 10 feet right or left of centerline. The offset was implemented to discourage pilots from using cues generated from the interaction of the perspective scene with the edges of the display screen (e.g., initiating flare when the runway edges reach a certain height on the screen) (see Johnson et al., 1989). The initial lateral offset was maintained through touchdown and rollout. The pilots adjusted to this quickly and reported no difficulty operating under these conditions.

Glide slope and localizer deviation scales were shown on the display down to an altitude of 100 ft., at which time they were turned off. This assured that pilots would maintain a stabilized approach to a point approximately 10 seconds from touchdown. For these last ten seconds of flight, they had only the perceptual cues provided in the display to guide their performance.
The reason for assuring a stabilized approach was that, in this experiment, we were interested in measuring pilots' flare and landing performance, which are accomplished during the last few seconds of flight. If we had allowed them to drift into non-stabilized approaches we would have added another variable (degree of stabilization) which would have added to the difficulty of interpreting our results.

Pilots were instructed to touchdown within a reasonable distance of the aim point (glide slope intercept) and as softly as possible (minimize vertical velocity). We stressed to subjects our desire that they minimize vertical speed at landing because we felt this was a good way to evaluate their ability to use visual cues to control the aircraft. When flying actual aircraft, a pilot's goal in terms of vertical speed is generally to make a soft landing (defined by Grantham (1989) to be less than 3 ft/sec), but once in the "soft" category, not attend to further reductions of vertical velocity. Given that our pilots generally landed harder than 3 ft/sec, we do not feel that the instructions we provided created a task for pilots that was appreciably different than their normal task. Feedback was provided on vertical speed at touchdown and longitudinal point of touchdown after each trial. Practice trials, covering all conditions, were provided until subjects achieved a consistent level of performance. This required between 12 and 32 trials. Examination of the data indicated no additional learning effects during experimental trials. Twenty test trials were run on each of the four experimental conditions in a counterbalanced order. Performance measurements included altitude at flare initiation, vertical velocity at touchdown, and longitudinal touchdown location. Flare initiation altitude was defined as the altitude, below 100 feet, at which the pilots initiated a sidestick input of at least 1.5 degrees. At the halfway point and end of each session pilots were asked to provide a subjective evaluation of the four display conditions.

RESULTS AND DISCUSSION - EXPERIMENT I

Empirical results.

Vertical speed at touchdown (sink rate). The average vertical speed at touchdown, as a function of display condition, for each of the six pilots tested, is shown in Table 1. An analysis of variance indicated a main effect for vertical speed as a function of condition ($F(3,456) = 3.24$, $p < 0.02$), with speed increasing as the level of texture increased (as indicated in Table 1). A strong main effect for pilots ($F(5, 456) = 49.16$, $p < 0.001$) indicated a significant difference between subjects in terms of their vertical speed at touchdown performance. There also existed a significant interaction between pilots and conditions ($F(5, 456) = 2.78$, $p < 0.001$). That is, the pattern of responses to the different display conditions varied among pilots. This can be seen in Figure 5, in which the curves for different subjects do not parallel one another. An examination of the change in vertical speed between the conditions with the least and most texture (Conditions 1 and 4) indicates a slight increase for four pilots and a decrease for two pilots. Of these changes, those for pilots 3 and 5 reached significance (multiple t-test comparison with alpha set to 0.05) with both subjects showing an increase in vertical touchdown speed with an increase in texture. This interaction between pilots and conditions was not expected and the variables responsible for it are not known. It means that we must be careful in interpreting the main effect that shows an increase in vertical velocity with an increase in display texture. The data are not sufficient to allow us to conclude that sink rate increases with display texture for the pilot population in general. It does appear, however, that there is at least a sub-population of pilots who do show this relationship under the existing experimental conditions.

VERTICAL SPEED AT TOUCHDOWN				
SUBJECT	**CONDITION 1** (NO TEXTURE)	**CONDITION 2** (R/W TEXTURE)	**CONDITION 3** (R/W & SURR. TEXTURE)	**CONDITION 4** (R/W MARKINGS ADDED)
1	**5.1 (1.4)**	**4.8 (1.7)**	**5.1 (2.7)**	**4.5 (1.8)**
2	**3.1 (1.0)**	**4.7 (2.4)**	**3.8 (2.4)**	**4.2 (1.9)**
3	**1.8 (0.9)**	**2.5 (1.3)**	**2.3 (1.2)**	**2.9 (1.2)**
4	**6.3 (2.4)**	**5.1 (2.2)**	**6.4 (2.3)**	**7.3 (2.4)**
5	**3.4 (1.3)**	**4.2 (1.9)**	**5.1 (1.9)**	**6.0 (1.8)**
6	**3.1 (1.0)**	**2.8 (1.0)**	**2.9 (1.0)**	**2.5 (1.4)**
Average	**3.8**	**4.0**	**4.3**	**4.6**

Table 1. Average vertical speeds at touchdown (ft./sec) as a function of display condition for each of the six pilots. Standard deviations are shown in parentheses.

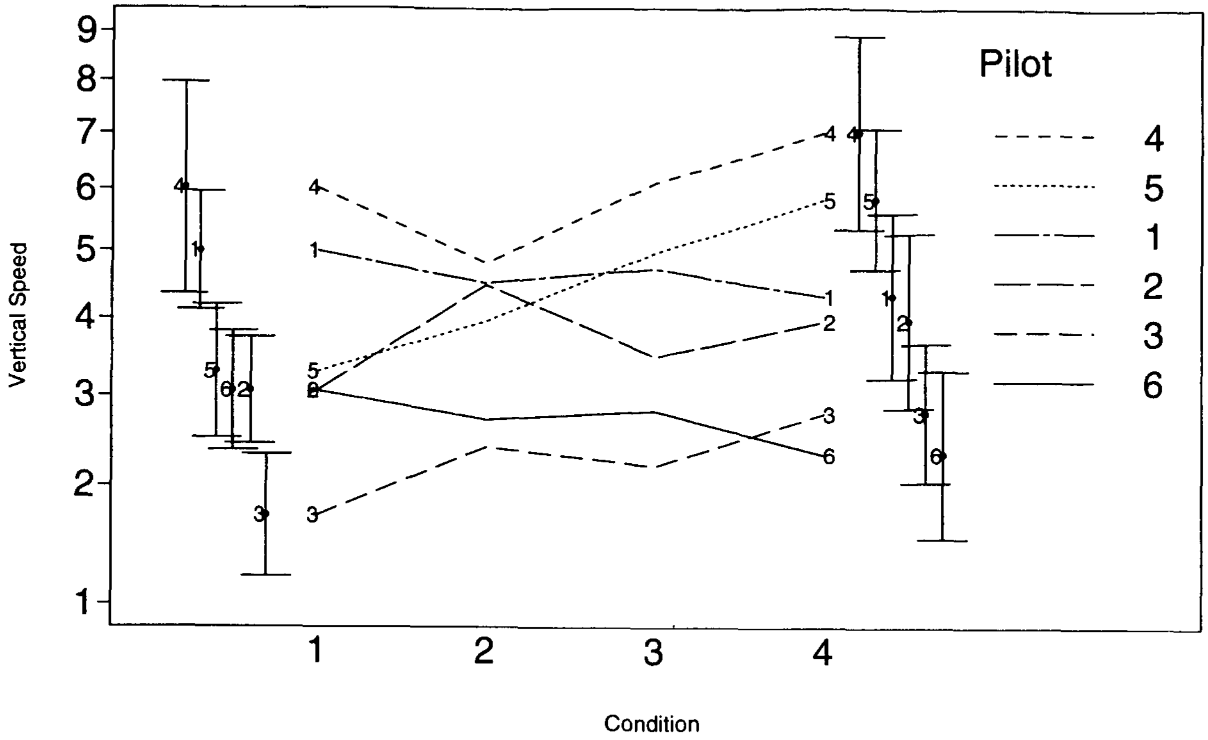

Figure 5. Average sink rate (ft./sec.) for each pilot as a function of display condition. Error bars show the standard error of the mean for conditions 1 and 4.

Figure 6 shows the distribution of trials as a function of touchdown vertical speed values for the different display conditions. We see a small shift of the distribution to higher vertical speeds and an increase in standard deviation between conditions 1 and 4.

Compared to actual aircraft landings, the vertical speeds at touchdown found in this study were high. Jewel and Stickle (1992) looked at the vertical velocity at touchdown of a variety of aircraft. Actual landings were measured. Findings indicated harder landings with turbojet than turboprop or piston-engined airplanes. The median vertical velocity for piston-engined aircraft was about 1.2 ft/sec while that for turbojets was about 1.5 ft/sec. The turbojets also had greater variability in touchdown sink rate. The hardest landing that could be expected in each 100 landings was estimated as 2.7 ft/sec for piston-engine aircraft and 4.2 ft/sec for turbojets. Other average vertical velocities at landing for propeller driven aircraft include 1.38 ft/sec (Silsby, 1955), 1.27 ft/sec at a sea level runway (Silsby and Livingston, 1959), and 0.92 ft/sec a high altitude (5300 ft) airport (Silsby and Livingston, 1959). In another evaluation of turbojet transport aircraft, Stickle and Silsby (1960) found average touchdown velocities of 1.62 ft/sec with one landing in 100 likely to exceed 4.0 ft/sec. Grantham (1989) compared landing performance using a full motion simulator to a six-degree-of-freedom in-flight simulator (USAF Total In-Flight Simulator, TIFS). He found 90 percent of the TIFS landings to be in the "smooth" (0-3 ft/sec) category, 4 percent in the "solid" (3-6 ft/sec) category, and 6 percent in the "hard" (over 6 ft/sec) category. Using the simulator, the results were 15, 70, and 15 percent for

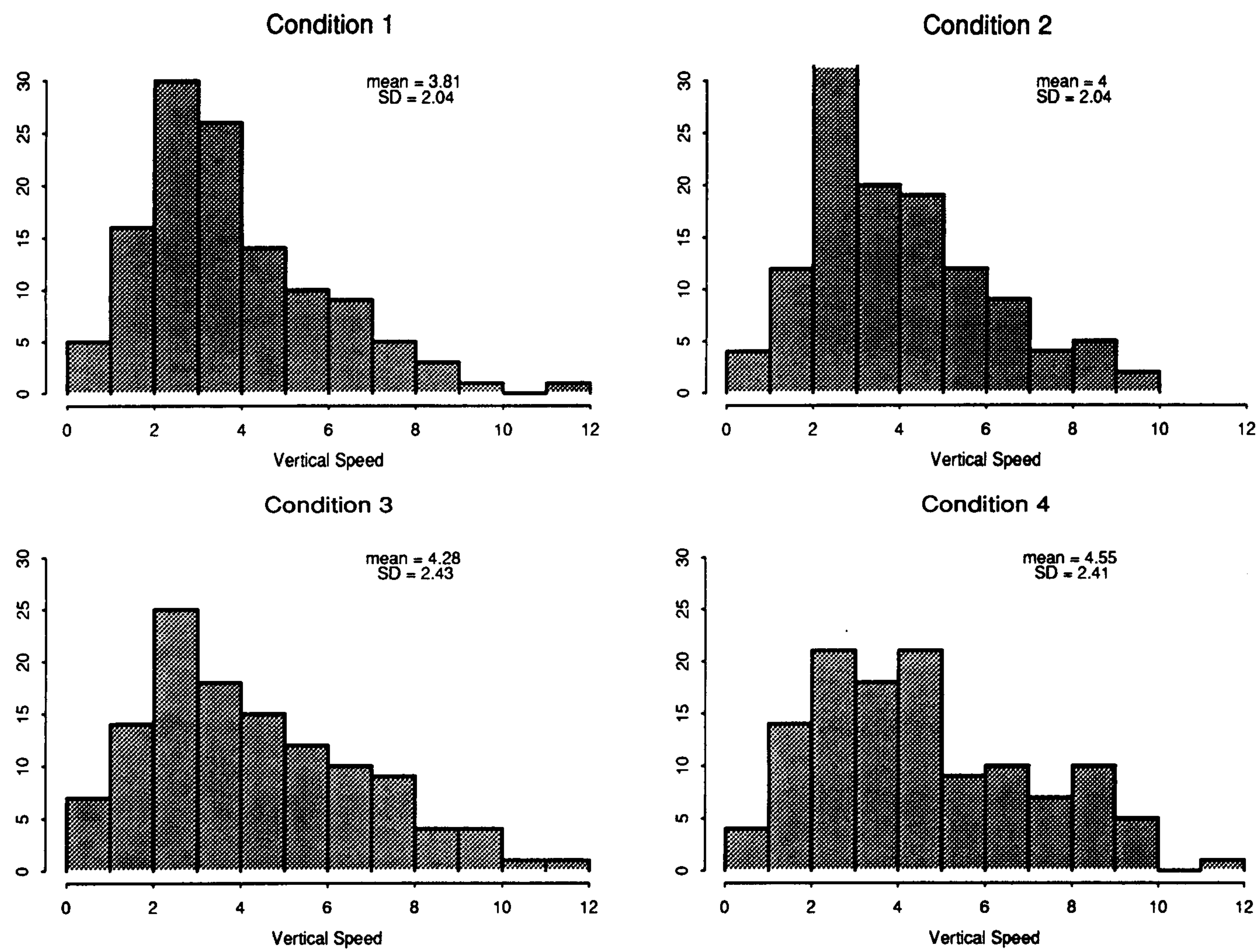

Figure 6. Distribution of touchdown vertical speeds, combined across subjects, for the four display conditions.

the smooth, solid, and hard categories respectively. The average touchdown sink rates were 1.7 ft/sec for the TIFS aircraft and 5.1 ft/sec for the simulator. Applying the categories defined by Grantham to the present study we found 27 percent of landings to fall in the smooth category, 58 percent in the solid category, and 13 percent that were hard. These results were not as good as TIFS aircraft landings found by Grantham (1989), but better than those obtained when they had pilots fly their full motion ground based simulator.

There was not a substantial difference between the performance of the general aviation (GA) pilot and the Boeing pilots on this or any of the performance measures in Experiment I, with the GA pilot actually performing above the group mean.

Landing distance. The average longitudinal landing distance from the runway threshold for each of the six pilots is shown in Table 2. An analysis of variance indicated a main effect for landing distance as a function of display condition ($F(3,456) = 6.25$, $p < 0.001$), with landing distance decreasing as the level of texture increased (see Table 2). The analysis also showed a strong main effect for pilots ($F(5, 456) = 42.08$, $p < 0.001$) indicating differential performance among pilots. An interaction between pilots and conditions ($F(5, 456) = 4.05$, $P < 0.001$) indicated that the differential performance among pilots was qualitative as well as quantitative. This can be seen in the variations in shape of the individual pilot performance curves shown in Figure 7. Subjects 1, 3, and 5 showed a fairly regular decline in landing distance with increasing scene complexity. Subjects 2 and 4 showed a decrease in landing distance between Condition 1 and Condition 4, but some substantial swings in performance for Conditions 2 and 3. Only subject 6 showed an increase in landing distance with an increase in texture. A statistical comparison of the performance of individual pilots on the lowest and highest texture conditions (conditions 1 and 4) indicated that in no case did the difference reach significance (t-test for multiple comparisons with a 0.05 alpha level). We conclude that, while there is a clear trend toward shorter landing distances with increased display texture, the data are not strong enough to support the conclusion that this finding can be generalized to the full pilot population.

LONGITUDINAL LANDING DISTANCE				
SUBJECT	**CONDITION 1** (NO TEXTURE)	**CONDITION 2** (R/W TEXTURE)	**CONDITION 3** (R/W & SURR. TEXTURE)	**CONDITION 4** (R/W MARKINGS ADDED)
1	**1534 (356)**	**1487 (510)**	**1291 (495)**	**1365 (507)**
2	**2046 (305)**	**1249 (355)**	**1544 (549)**	**1899 (365)**
3	**1933 (364)**	**1678 (498)**	**1677 (357)**	**1541 (421)**
4	**1131 (582)**	**863 (351)**	**1168 (555)**	**770 (107)**
5	**2004 (287)**	**1820 (354)**	**1786 (283)**	**1757 (486)**
6	**1441 (472)**	**1789 (404)**	**1659 (557)**	**1603 (407)**
Average	**1682**	**1481**	**1521**	**1489**

Table 2. Average longitudinal landing distance (ft.) from threshold for each of the six pilots. Standard deviations are shown in parentheses.

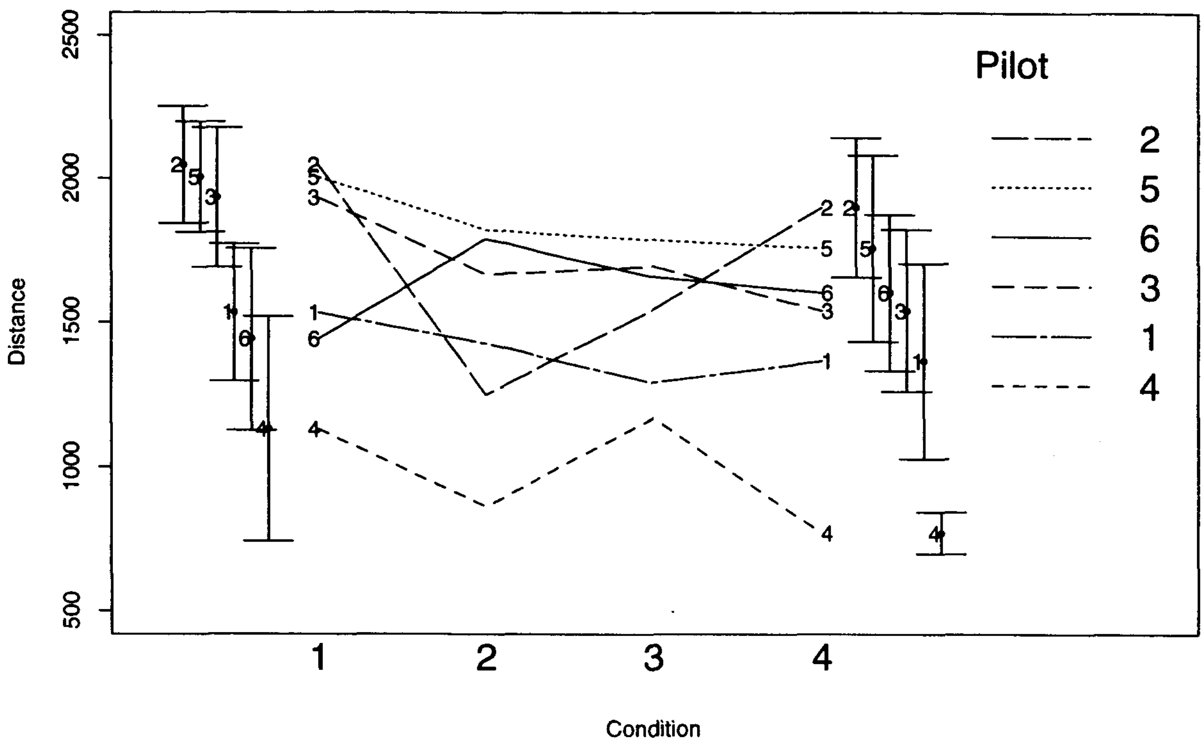

Figure 7. Average longitudinal landing distance from threshold (ft.) for each pilot as a function of display condition. Error bars show the standard error of the mean for conditions 1 and 4.

Flare initiation altitude. The average altitude at flare initiation as a function of display condition for each of the six pilots is shown in Table 3. An analysis of variance indicated a main effect for flare initiation altitude as a function of display condition ($F(3,456) = 12.13$, $p < 0.001$). Table 3 indicates that the direction of this effect is toward lower flare initiation altitude with increasing levels of texture. The analysis also showed a strong main effect for pilots ($F(5, 456) = 56.00$, $p < 0.001$) indicating differential performance among them. A significant interaction between pilots and conditions ($F(5, 456) = 2.44$, $p < 0.002$) indicated that the way in which performance varied with conditions was different for different pilots. The extent of this interaction can be seen in Figure 8. While the shapes of the curves in the figure vary, an examination of performance changes between the conditions with the least and most amount of texture (Conditions 1 and 4) indicated a decrease in flare initiation altitude with increasing texture for five of the six subjects, with the sixth subject showing no change. Of these five changes, two reached the level of significance: subjects 2 and 5 (t-test for multiple comparisons with alpha set to 0.05). A Numan Keuls test for differences between conditions produced a number of significant differences: condition 1 > condition 3, condition 1 > condition 4, condition 2 > condition 4, and condition 3 > condition 4. We conclude that there is a condition effect, indicating a decrease in flare initiation altitude with increasing levels of texture. It should be noted, however, that some pilots responded more strongly than others.

FLARE INITIATION ALTITUDE				
SUBJECT	CONDITION 1 (NO TEXTURE)	CONDITION 2 (R/W TEXTURE)	CONDITION 3 (R/W & SURR. TEXTURE)	CONDITION 4 (R/W MARKINGS ADDED)
1	36.1 (4.2)	32.5 (4.0)	30.3 (5.5)	31.7 (4.9)
2	45.1 (4.8)	38.3 (10.3)	38.5 (5.8)	36.6 (3.9)
3	39.3 (9.3)	41.9 (5.0)	40.7 (5.1)	37.2 (5.6)
4	27.9 (8.4)	28.8 (3.9)	27.5 (5.4)	25.0 (4.4)
5	42.7 (7.6)	39.4 (5.8)	37.1 (6.5)	33.7 (5.1)
6	37.0 (6.4)	39.4 (5.0)	39.3 (5.6)	37.2 (4.3)
Average	38.0	36.7	35.6	33.6

Table 3. Average altitudes at flare initiation (ft.) as a function of display condition for each of the six pilots. Standard deviations in parentheses.

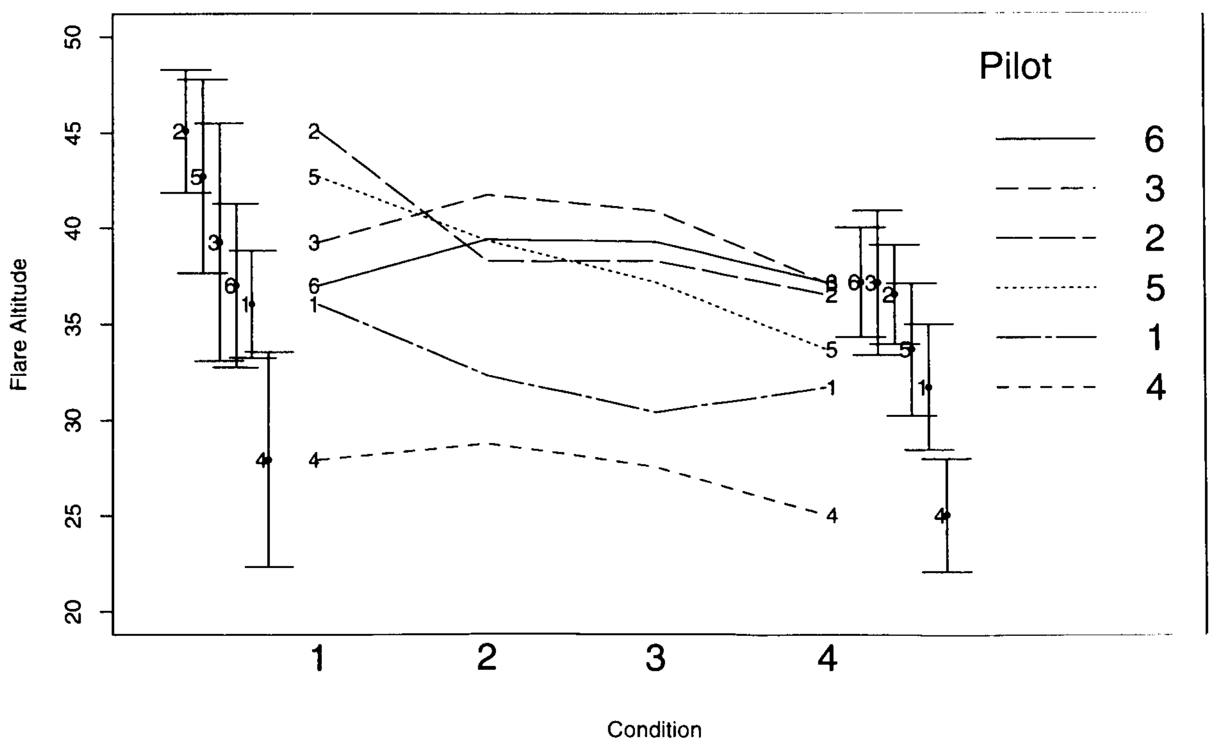

Figure 8. Average altitude of flare initiation (ft.) for individual subjects as a function of display condition. Error bars show the standard error of the mean for conditions one and four.

Interactions between variables. It is of interest to examine possible interactions among the dependent variables of sink rate, distance, and flare initiation altitude.

Figure 9 shows the relationship between sink rate at touchdown and length of landing for each of the four display conditions. A fairly consistent pattern is seen across the four conditions with clearly harder landings at the shortest landing distances followed by a significant improvement as landing distance gets longer. The softest landings occurred between roughly 900 and 1400 ft. (conditions 2,3, and 4) and in some cases for longer landings (conditions 1, 2, and 3). There also appear to be fewer landings in the 1000-1400 ft. range (see Fig. 10). One possible explanation for the harder landings at short distances is that it is a more difficult manual control task to flare and set the aircraft down gently in a short distance than to slowly reduce vertical speed over a longer distance (i.e., longer time). It is possible that the harder landings out in the 2000 foot range may have resulted from pilots realizing they were floating the aircraft and coming up on the 2500 foot mark that would result in a rejected trial. A tendency was noticed for some pilots to become more aggressive in putting the aircraft on the ground at this point and accepting the harder touchdown. This does not, however, explain the fact that in three of the four display conditions there was a decrease in vertical speed between 2000 and 2500 ft.

Figure 10 shows the relationship between landing distance and flare initiation altitude as a function of display condition. It can be seen that for each condition there was a tendency for landing distance to increase as the height of flare initiation increased. It appears that there may be a bimodal distribution. When flare initiation altitude was below 30 feet there was a strong tendency to land short. When it is above this altitude the range of landing distances spread out over a much greater range. An analysis of covariance produced the following correlations: condition 1 = 0.56, condition 2 = 0.60, condition 3 = 0.55, and condition 4 = 0.61. There are no significant differences between these correlations.

Figure 11 shows the relationship between vertical speed at touchdown and flare initiation altitude. An analysis of covariance to test for a relationship between these variables produced the following correlation coefficients: condition 1 = -0.35, condition 2 = -0.39, condition 3 = -0.35, and condition 4 = -0.53. A comparison of these correlations (using a z-statistic for differences in correlations) indicates that the correlation in condition 4 is significantly greater than those for the other conditions.

Subjective Results.

Pilots were asked to provide a subjective rating of the four display conditions, indicating their preferences in light of the approach, flare and landing tasks they were performing. The results for each of the six subjects is shown in Table 4. Rankings go from most preferred (score of 1) to least preferred (score of 4). A strong preference was indicated for increasing amounts of texture by all pilots. The no texture condition (Condition 1) was considered strongly inferior to any other condition by all pilots. A comparison of the runway texture condition (Condition 2) and runway plus surround texture condition (Condition 3) indicate a preference for Condition 3, but this was not a strong preference. Subject 1 rated these two conditions equal, and subject 6 preferred Condition 2. The other four pilots indicated a minimal difference. A textured scene with runway markings (Condition 4) was preferred by all pilots with the magnitude of the preference being small to moderate compared to the next most preferred condition.

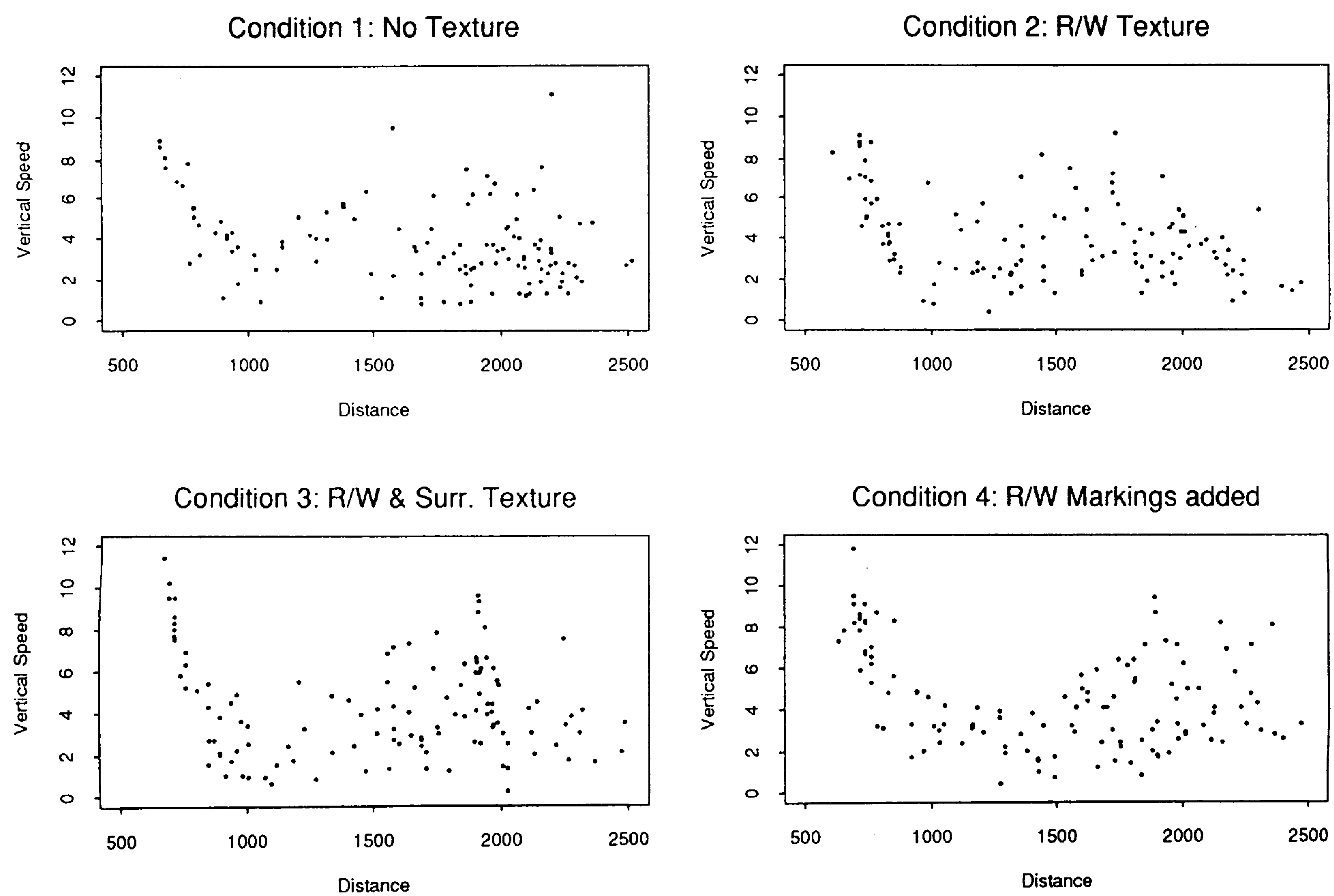

Figure 9. Vertical speed at touchdown (ft./sec.) as a function of landing distance for each of the four experimental conditions. Data points represent individual landings. The data for all six subjects is included in each graph.

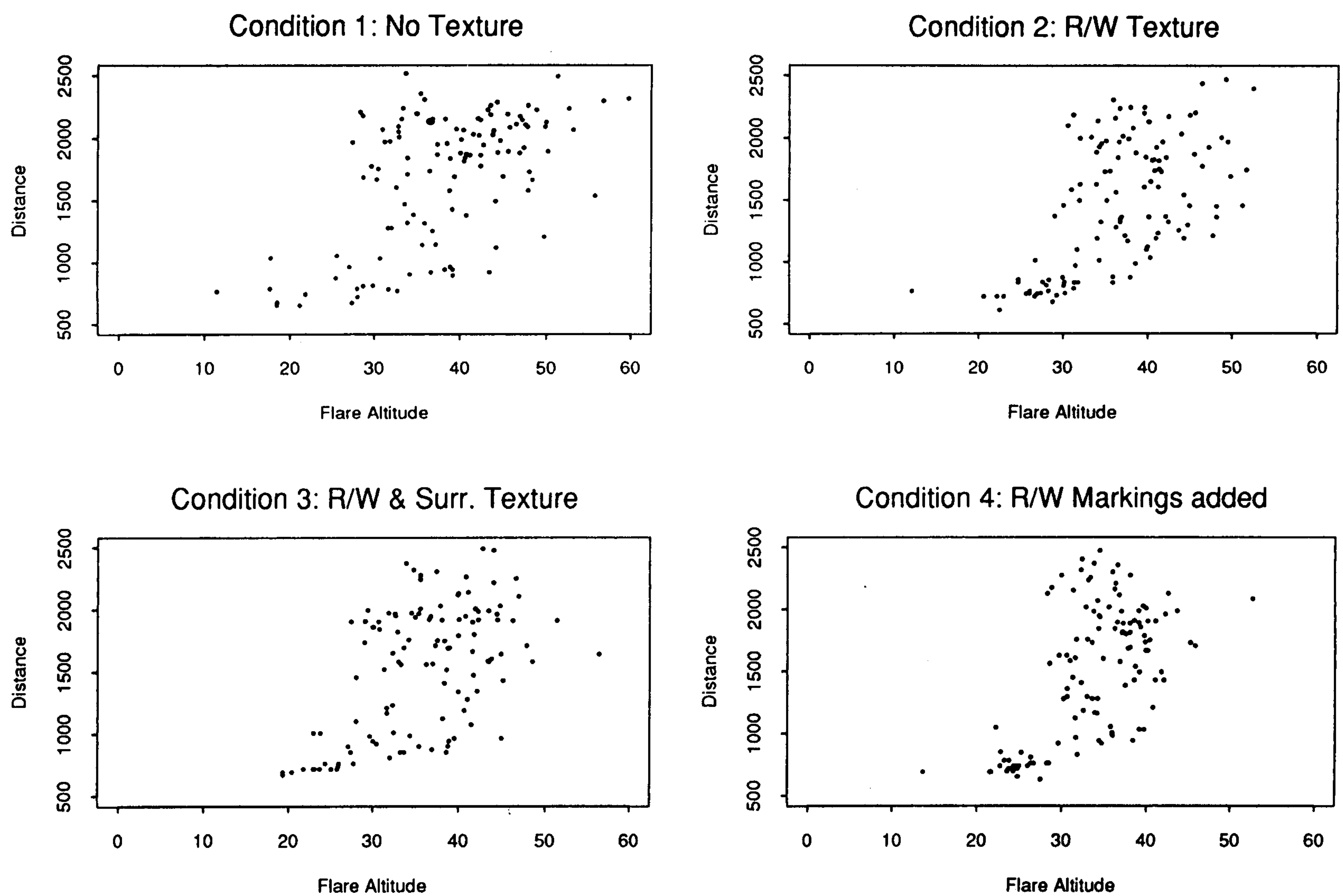

Figure 10. Landing distance at touchdown (ft.) as a function of flare initiation altitude (ft.) for each of the four experimental conditions. Data points represent individual landings. The data for all six subjects is included in each graph.

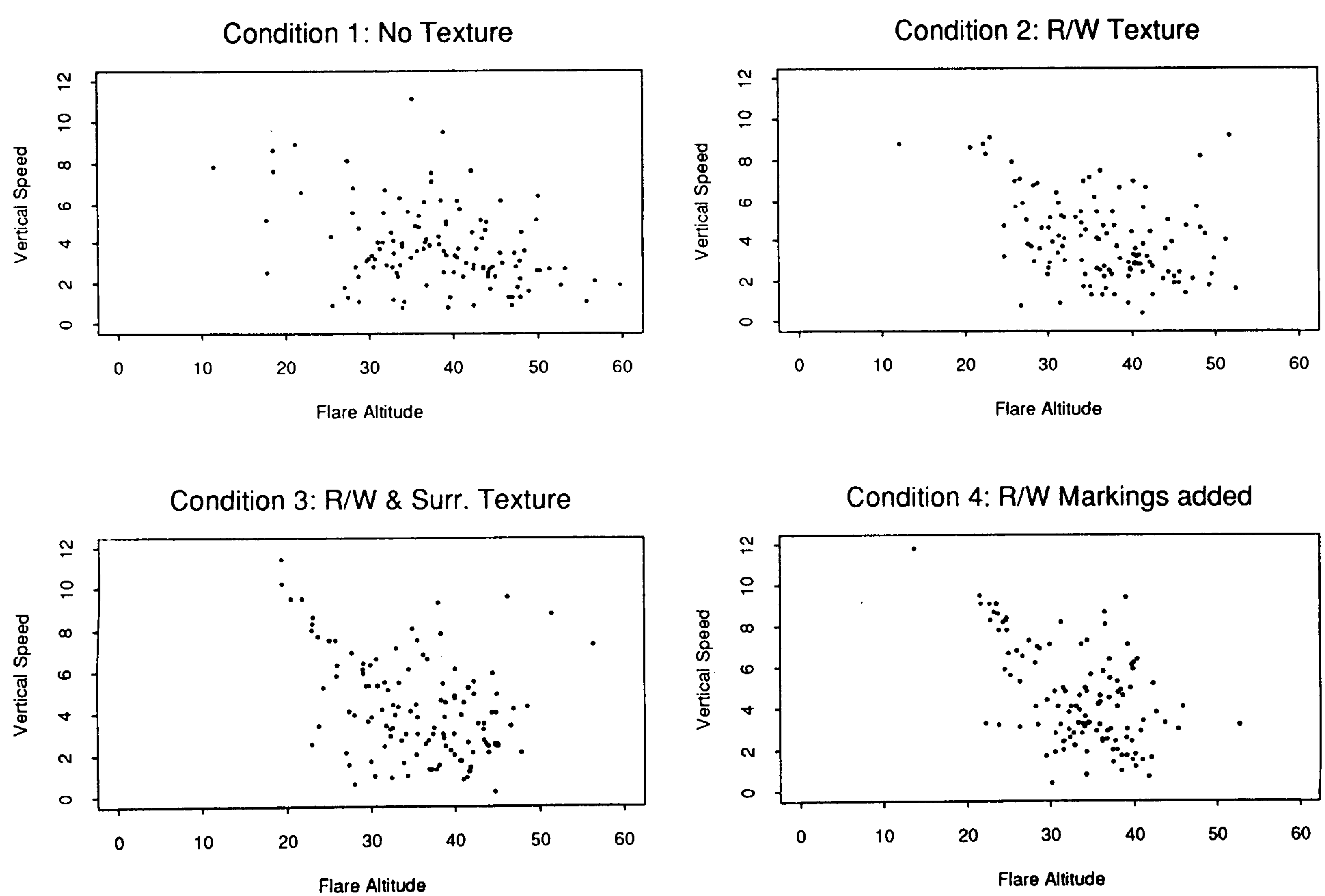

Figure 11. Vertical speed at touchdown (ft./sec.) as a function of flare initiation altitude (ft.) for each of the four experimental conditions. Data points represent individual landings. The data for all six subjects is included in each graph.

PILOT SUBJECTIVE PREFERENCE				
SUBJECT	CONDITION 1 (NO TEXTURE)	CONDITION 2 (R/W TEXTURE)	CONDITION 3 (R/W & SURR. TEXTURE)	CONDITION 4 (R/W MARKINGS ADDED)
1	4	2	2	1
2	4	3	2	1
3	4	3	2	1
4	4	3	2	1
5	4	3	2	1
6	4	2	3	1

Table 4. Pilot's subjective preference for different display conditions in Experiment I. Rank ordering goes from 1 to 4 (most preferred to least preferred).

ADDITIONAL DISCUSSION - EXPERIMENT I.

Interpretation of results

The primary dependent variable we examined was vertical speed at touchdown. This is the variable we specifically instructed pilots to attend to. The results showing an increase in sink rate with increased texture was unexpected and one we cannot explain at this time. It is significant that while the sink rates we obtained were higher than those expected for landing actual aircraft, they were less than those obtained in a much higher fidelity simulator (Grantham, 1989).

We found a decrease in touchdown landing distance with an increase in display texture. However, determination of what constitutes desirable performance is not so straight forward as with sink rate. The subjects were not specifically asked to land as close to the aim point as possible, but rather within a "reasonable" distance of it. It is also the case that the landing distances for the different display conditions are all within an acceptable range. Still, it is considered advantageous to land close to the designated aim point, allowing the aircraft to be stopped in a shorter distance or with less breaking. We consider the results on landing distance to indicate at least a limited advantage for increased display texture.

Results indicated a decrease in flare initiation altitude with increases in the level of texture. Again, the interpretation of these results is not straight forward. The pilots were not instructed as to when to initiate flare, and received no feedback regarding their performance. In general, however, lower flare initiation altitudes are preferable as they tend to lead to shorter landings and thus shorter stops. Indeed there was a positive correlation between lower flare initiation altitudes and shorter landings. On the other hand, there was also a positive correlation between lower flare altitudes and harder landings, although this second correlation was substantially weaker. Lower flare initiation altitudes also move pilot performance closer to the 15-20 feet (gear to ground) altitude recommended for manual flare initiation in 737 aircraft. Our feeling is that a lower flare initiation altitude (within the range covered by out testing) is indicative of desirable performance on the part of pilots, and as such indicates a positive effect of increased levels of display texture. It is possible that the increase in display texture provides pilots with additional situational awareness that allows them to feel comfortable flaring at a lower altitude.

The subjective reports provided by pilots indicated a clear preference for increased texture. We consider this an important factor in support of display texture.

The results of Experiment I follow the general pattern of the Regal and Whittington (1993) experiment in which we looked at the effects of "familiar size" cues on pilot performance. Landing distance, flare initiation altitude, and subjective reports indicate an advantage for displays with richer perceptual cues. Results for sink rate failed to indicate any benefits for richer cue conditions and actually indicated a disadvantage in the present study.

In conclusion, the benefits of increased display texture, as employed in this study, are ambiguous. Performance was improved on three measures (landing distance, flare initiation altitude, and subjective preference), but made worse on the important measure of vertical velocity at touchdown.

EXPERIMENT II - THE EFFECT OF INCREASED WORKLOAD ON THE USEFULNESS OF PERCEPTUAL CUES.

RATIONALE - EXPERIMENT II

Experiment II was conceived largely as an attempt to explain the unexpected results of Experiment I and the related Regal and Whittington (1993) experiment. Both these studies indicated only a minimal performance advantage form adding enriched perceptual cues to the displays being flown. These results were not as strong as would be expected from either the literature or the subjective preferences expressed by the pilots participating in the studies.

We do not have a satisfactory explanation for these findings, and believed it important to further investigate the issue. One possibility is that the results reflect a ceiling effect that has caused the differences between conditions to wash out. To test this hypothesis we chose to increase the pilot's workload and re-evaluate three factors: familiar size cues, texture gradients, and different texture patterns. To keep the duration of experimental sessions within acceptable limits, and still test three factors in a single experiment, we were able to test only two conditions per factor, but this broad brush approach was considered appropriate to our attempt to explain the unexpected results of the previous experiments.

METHODOLOGY - EXPERIMENT II

Displays.

Four different display conditions were tested. The first was a no cue condition that contained a gray runway on a homogeneous green surround with a blue sky (Figure 12). The next condition added standard runway markings (AIM 1991) to the display (Figure 13). The third display had texture added to both the runway and surrounding area in addition to runway markings (Figure 14). The fourth condition was similar to the third except that a different texture pattern was employed on the runway (Figure 15). Two of these displays (Figures 12 and 13) represent different levels of familiar size cues and are similar to displays used in the Regal and Whittington (1993) experiment evaluating familiar size cues. Other pairs (Figures 12 and 14 or 12 and 15) represent two levels of texture and are similar to displays employed in Experiment I of this contract in which we looked at the effects of texture. The final comparison was between two different types of texture. These are shown in Figures 14 and 15. As in Experiment I, the three dots to either side of the runway at the 1000 ft. from threshold help identify the pilot's aim point (glideslope intersect).

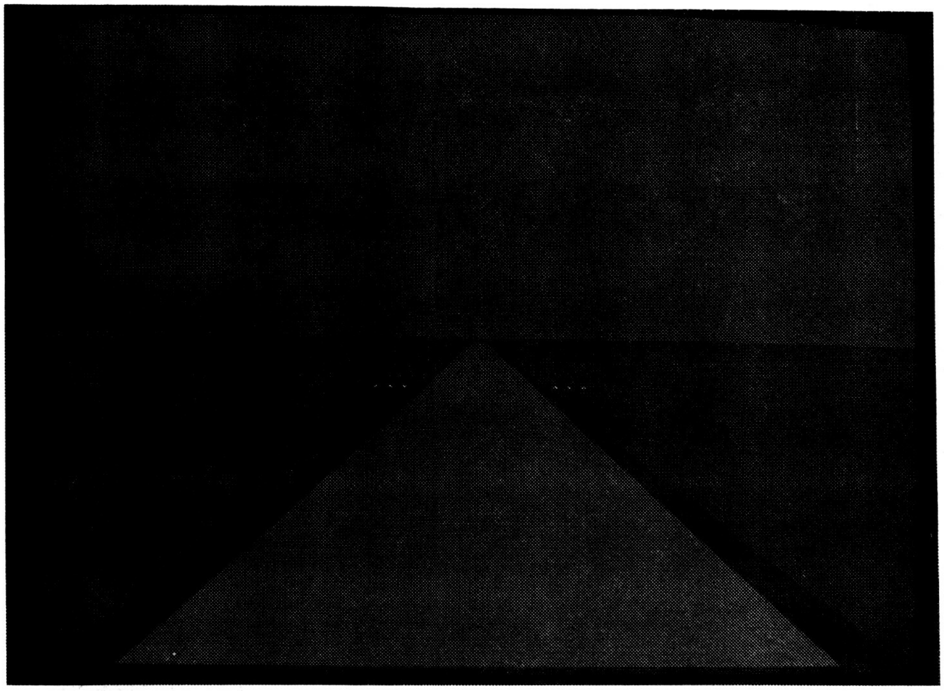

Figure 12. Experiment II, Display condition 1. A gray runway on a homogeneous green surround with a blue sky. The white dots indicate the glideslope intercept point (1000 ft. from threshold). This figure is the same as that used in Experiment I, condition 1.

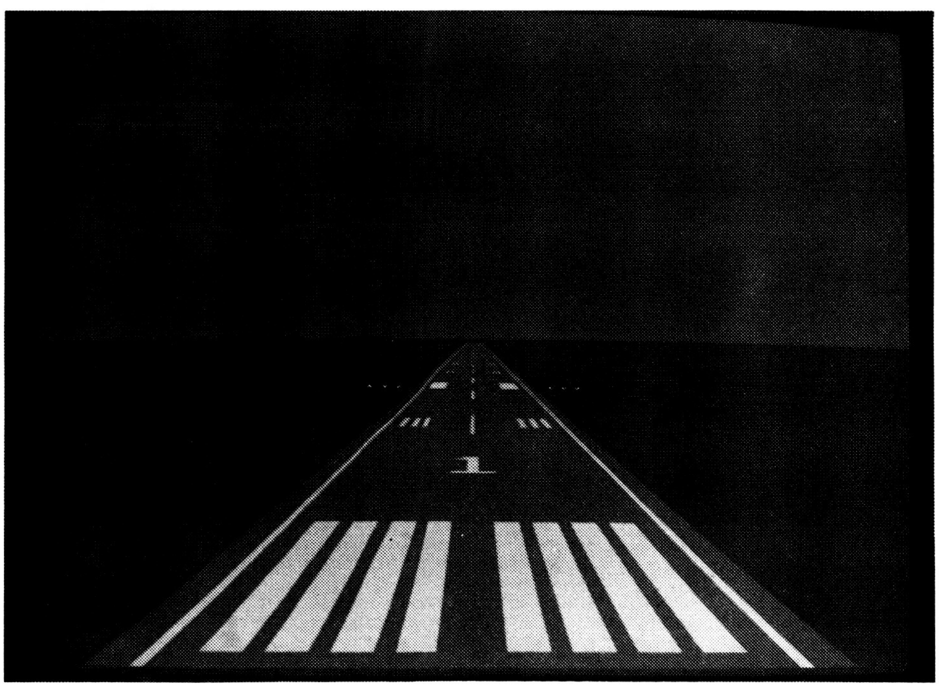

Figure 13. Experiment II, Display condition 2. This display is the same as used in display condition 1, but with the addition of standard runway markings.

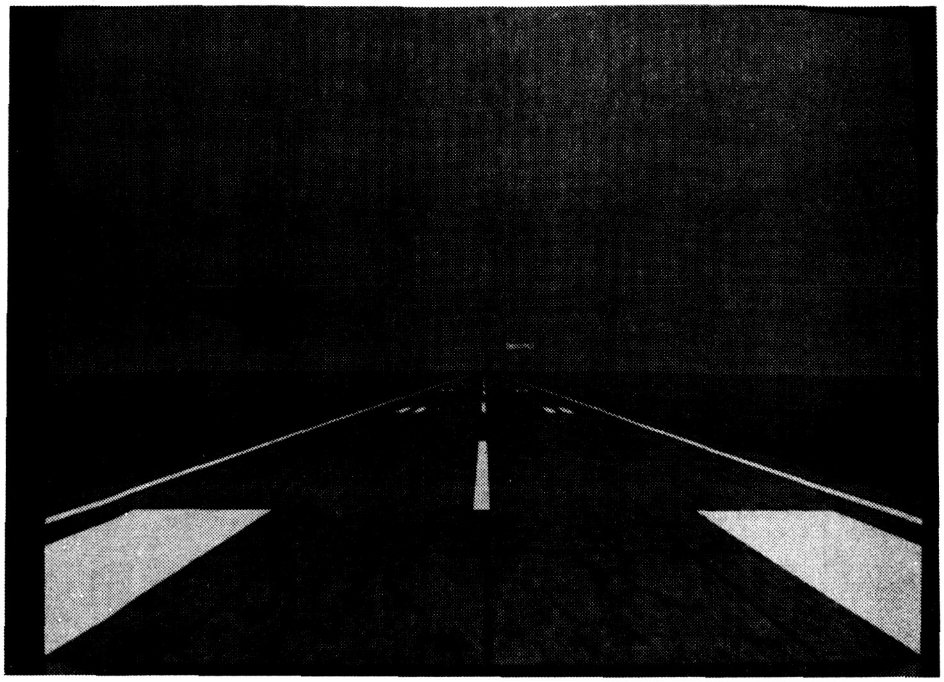

Figure 14. Experiment II, Display condition 3. Similar to condition 2, but with the addition of texture patterns to the runway and the surround. The figure is the same as that used in Experiment I, condition 4.

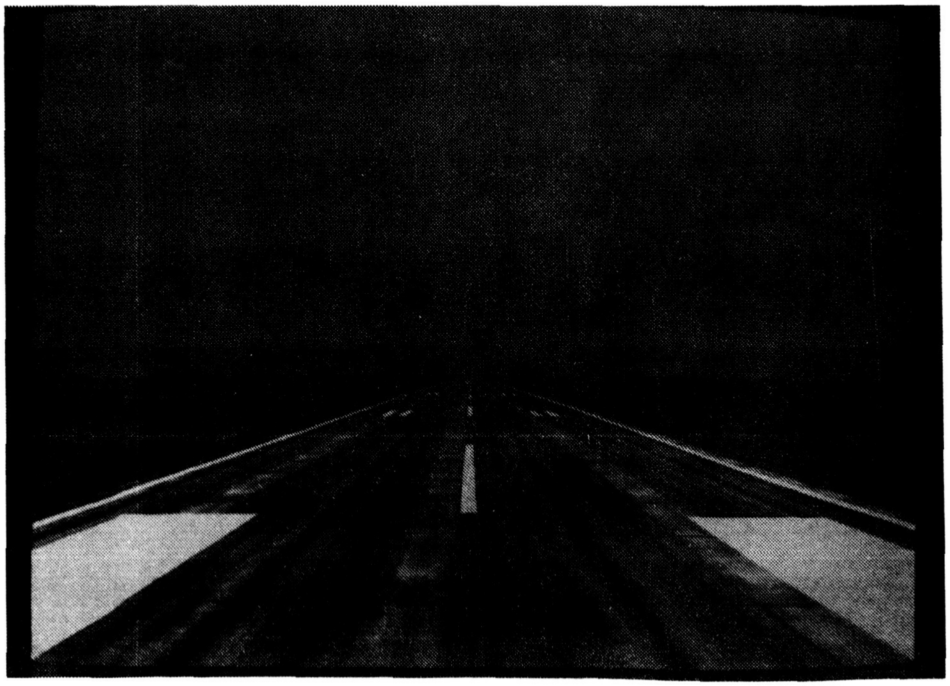

Figure 15. Experiment II, Display condition 4. Similar to Figure 14, but with a different texture pattern used on the runway.

Apparatus.

Scenes were generated on a Silicon Graphics Onyx/RE2 workstation. This is a dual R4400 (50 MHz) 3D graphics machine with a high performance graphics pipeline capable of several thousand textured polygons per frame at an update rate of 20 Hz. The picture is fully antialiased. The subjects were positioned 20 inches from the display screen which resulted in a field of view 40 degrees wide by 30 degrees high. A Measurement Systems Inc. sidestick controller was used to control flight path. A high fidelity 737-500 airplane model was used. This simulation included all code used for pilot testing in full mission cab simulations. It has all the aerodynamic effects used for flight test matching, including ground effects near touchdown. Engine settings were controlled by autothrottle. The aircraft's approach speed was 140 kt. and bled off in accordance with autothrottle procedure, with the engines going to idle over a period of six seconds beginning at an altitude of 27 ft. The engine model had all appropriate spool up/down characteristics. The aircraft model was executed on a Harris Nighthawk 5820 processor separate from the graphics unit.

Subjects.

Eight subjects were tested. Six were experienced Boeing pilots and two experienced general aviation pilots. All were active in their flying careers at the time of testing.

Procedure.

Experimental trials began with the aircraft approximately 20 sec from touchdown. The aircraft was initialized on a three degree glide slope and either 25 or 45 feet right or left of a vertical plane through the runway centerline. The initial lateral deviation condition was varied in a pseudo-random fashion. It was the pilots task to maintain a stabilized approach and flare and land the aircraft. Glide slope and localizer deviation scales were provided down to an altitude of 100 feet or approximately 10 seconds from touchdown. Thus, for the last ten seconds of flight, the pilots had only the perceptual cues provided in the display to guide their performance. They were instructed to touch down as softly as possible, as close to the centerline as possible, and within a reasonable distance of the glide slope intercept point. Feedback was provided on vertical speed, lateral offset, and the longitudinal point of touchdown after each trial. Practice trials, covering all conditions, were provided until subjects achieved a consistent level of performance. Twenty test trials were run on each of the four experimental conditions with trials run in a pseudo-random order. Performance measurements included altitude at flare initiation, vertical velocity at touchdown, lateral offset at touchdown, roll angle at touchdown, track angle at touchdown, and longitudinal touchdown location. Pilots were asked to provide a subjective evaluation of the four display conditions including order of preference and magnitude of the difference between conditions.

RESULTS AND DISCUSSION - EXPERIMENT II

The results for Experiment II will be covered in three separate sections corresponding to the three variables being manipulated: familiar size cues, texture level, and texture type. For each of these variables we consider only those display conditions that are applicable. The performance parameters examined include those considered in Experiment I: vertical speed at touchdown, landing distance, and flare initiation altitude, as well as additional measures: lateral offset at touchdown, track error at touchdown, and bank angle at touchdown.

Experiment IIA - Familiar Size Cue Experiment.

Empirical Results.

A comparison of display conditions 1 and 2 (Figs. 12 and 13) are relevant to this experiment. These are a subset of the display conditions used in the Regal and Whittington (1993) experiment on familiar size cues.

Vertical speed at touchdown. The average vertical speed at touchdown for each of eight experimental subjects is shown in Table 5. A graphical representation of this data is shown as part of Figure 16. An analysis of variance indicated a main effect for vertical speed at touchdown ($F(1, 304) = 6.9$, $p < 0.01$) with landings being significantly harder for the display containing runway markings (condition 2) as compared to the display showing a uniform runway (condition 1). Differences between pilots were also highly significant ($F(7, 304) = 16.6$, $p < 0.001$). This difference is qualitative as well as quantitative as indicated by a significant pilot/condition interaction ($F(7,304) = 2.3$, $p < 0.03$) and can be seen in the differences in the shapes of the curves in Figure 16. While six of the eight subjects showed an increase in touchdown sink rate in condition 2 compared to condition 1, a comparison of the performance for individual pilots (using a multiple t-test comparison with alpha set at the 0.05 level) indicated no significant differences for any of the subjects.

Discussion. These findings confirm the trend toward harder landings with added familiar size cues found by Regal and Whittington (1993). This result was and is surprising in that one would expect better performance in the richer cue condition (condition 2). We do not have a satisfactory explanation for this finding, but it may be correlated in some fashion with the significantly longer landing distances found for display condition 1.

SUBJECT	CONDITION 1 (runway only)	CONDITION 2 (r/w with markings)
1	2.5 (1.4)	2.8 (1.8)
2	4.8 (1.9)	5.5 (2.1)
3	3.3 (1.3)	4.2 (1.9)
4	5.3 (2.2)	4.9 (1.9)
5	7.4 (3.0)	6.1 (2.6)
6	6.0 (2.2)	6.8 (2.6)
7	3.1 (1.7)	4.8 (2.2)
8	4.3 (2.9)	6.3 (2.2)
Average	4.59	5.18

Table 5. Average vertical speed (ft./sec.) at touchdown for conditions 1 and 2 for each test subject. Numbers in parentheses indicate standard deviations. Conditions 1 and 2 employed the displays shown in Figs. 12 and 13 respectively.

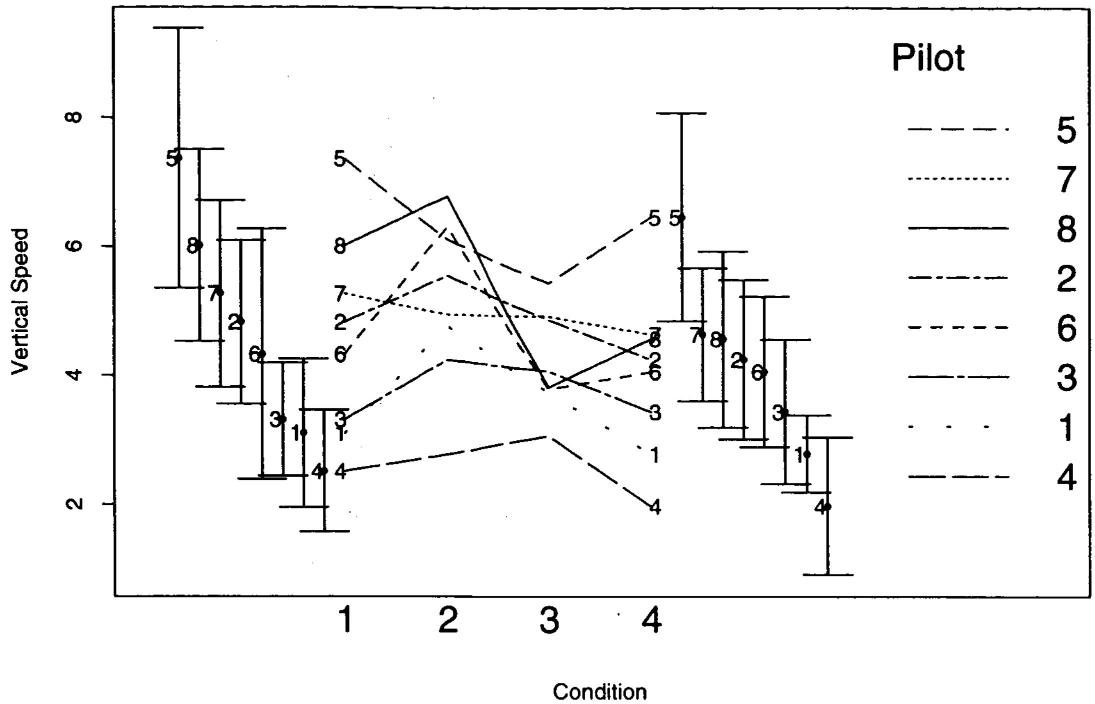

Figure 16. Vertical speed at touchdown plotted as a function of display condition for each test subject. Bars representing standard error of the mean are shown for conditions 1 and 4.

Landing distance. The average landing distance at touchdown for each subject, as a function of display condition, is shown in Table 6. A graphical representation of this data is shown as part of Figure 17. An analysis of variance showed a main effect for landing distance ($F(1, 304) = 19.9$, $p < 0.001$) with significantly shorter landings when the display contained greater familiar size cues (condition 2) . Significantly different performance levels were found between pilots ($F(7, 304) = 12.5$, $p < 0.001$). A pilot by condition interaction was also observed ($F(7,304) = 2.3$, $p < 0.01$) indicating that the decrease in landing distance with display condition 2 may not apply equally to all pilots. Of the seven out of eight pilots who did show this effect, analysis (using a multiple t-test comparison with alpha set at the 0.05 level) indicated a significant differences only for subject 8.

Discussion. The results of this experiment show the same significant decrease in landing distance with increased scene complexity as found in the Regal and Whittington (1993) experiment. We are not sure of the explanation for these findings in terms of the perceptual cue content of the scenes, but it may be that the reduced cues in display condition 1 caused a more cautious and thus longer landing.

Flare initiation altitude. The average flare initiation altitude for each subject is provided in Table 7. A graphical representation of this data is shown as part of Figure 18. An analysis of variance model does not fit this data well enough to provide meaningful comparisons - there are problems with unequal variance and non-normal residuals. However, an examination of the data does not indicate a strong effect for flare initiation altitude as a function of the level of familiar

SUBJECT	CONDITION 1 (runway only)	CONDITION 2 (r/w with markings)
1	1807 (363)	1589 (443)
2	1765 (410)	1516 (421)
3	1724 (478)	1444 (432)
4	1283 (419)	1174 (364)
5	1055 (317)	1190 (257)
6	1275 (526)	1033 (274)
7	1317 (379)	1288 (448)
8	1654 (502)	1074 (250)
Average	1485	1289

Table 6. Average landing distance (ft.) at touchdown for conditions 1 and 2 for each test subject. Numbers in parentheses indicate standard deviation. Conditions 1 and 2 employed the displays shown in Figs. 12 and 13 respectively.

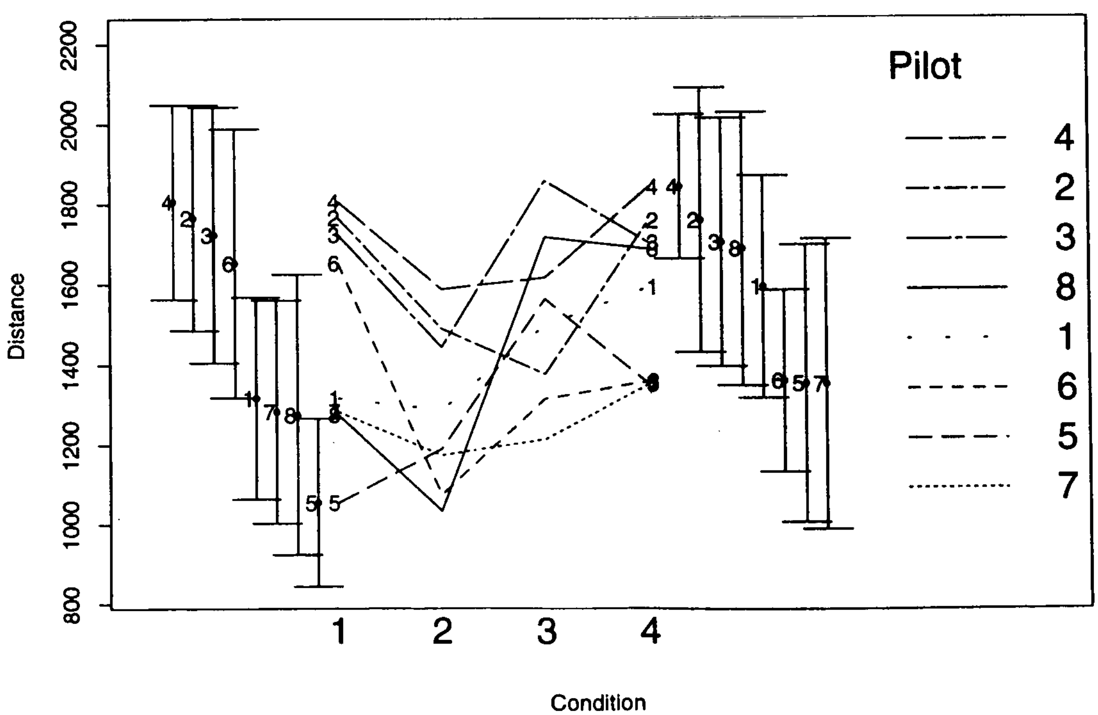

Figure 17. Average landing distance at touchdown for each pilot as a function of display condition. Bars representing standard error of the mean are shown for conditions 1 and 4.

SUBJECT	CONDITION 1 (runway only)	**CONDITION 2 (r/w with markings)**
1	48.0 (7.8)	50.4 (10.2)
2	49.3 (11.2)	54.9 (16.1)
3	54.4 (9.0)	49.3 (11.7)
4	61.4 (14.8)	64.4 (17.0)
5	44.2 (17.4)	56.1 (23.1)
6	47.6 (12.5)	38.6 (7.0)
7	50.0 (17.9)	49.9 (14.9)
8	54.0 (20.7)	58.9 (19.1)
Average	51.1	52.8

Table 7. Average flare initiation altitude (ft.) for each test subject. Numbers in parentheses indicate standard deviations.

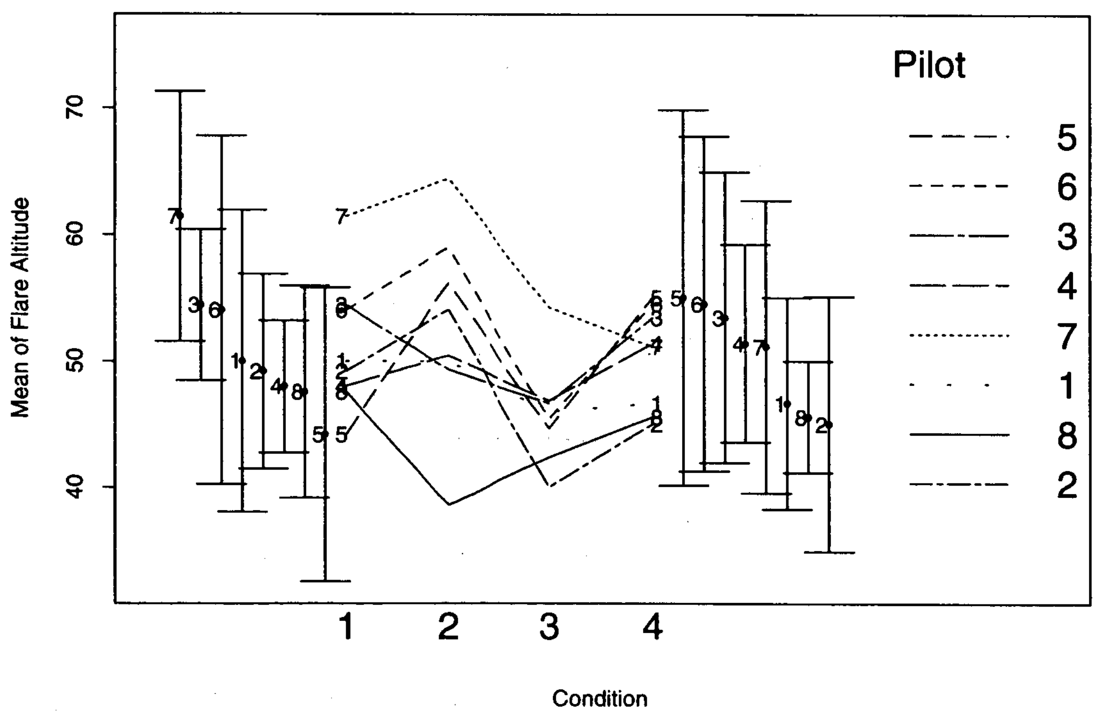

Figure 18. Average flare initiation altitude for each test subject is plotted as a function of condition. Bars representing standard error of the mean are shown for conditions 1 and 4.

size cue present in the displays. Five of the eight subjects showed a higher flare initiation altitude in condition 2 than condition 1, but an analysis of these individual performances (using a multiple t-test comparison set at the 0.05 level) indicated no significant differences for any of the subjects.

Discussion: Regal and Whittington (1993) found a significant decrease in flare initiation altitude for condition 2. The present study did not.

With the flare initiation altitude measure we are also interested in within subject variance, given the presumption that more precise spatial situational awareness will lead to reduced variability. Neither the present experiment nor the Regal and Whittington (1993) study found a significant difference in the variability of flare initiation altitudes between conditions 1 and 2.

Lateral offset. The average RMS lateral offset at landing for each subject is provided in Table 8. A graphical representation of this data is shown as part of Figure 19. An analysis of variance indicated a strong main effect for lateral offset as a function of condition ($F(1, 256) = 13.9$, $p < 0.001$), with a decrease in lateral offset found for the display condition containing runway markings. Differences in performance levels between pilots was also found ($F(7, 256) = 6.1$, $p < 0.001$). An analysis of the performance of individual pilots (using a multiple t-test comparison set at the 0.05 level) indicated no significant differences between conditions 1 and 2 for any of the subjects.

Discussion: As would be expected in comparing a runway with and without a centerline, pilots had a significantly lower lateral offset with the centerline available (condition 2).

SUBJECT	CONDITION 1 (runway only)	CONDITION 2 (r/w with markings)
1	8.9 (-3.0)	7.2 (+2.7)
2	23.7 (-5.4)	19.2 (-2.9)
3	24.8 (-13.4)	16.9 (-6.0)
4	23.0 (-5.6)	18.0 (+2.4)
5	26.7 (+2.2)	17.5 (-11.7)
6	25.6 (-17.8)	22.0 (-14.1)
7	20.3 (-7.5)	18.5 (-2.9)
8	28.1 (-8.6)	18.4 (-4.4)
Average	22.6	17.2

Table 8. The average RMS lateral offset (ft.) at touchdown for each test subject for display conditions 1 and 2. The numbers in parentheses represent the mean offset.

Track error at touchdown. The average RMS track error at touchdown for each subject is shown in Table 9. A graphical representation of this data is provided as part of Figure 20. An analysis of variance indicated a strong main effect for track error as a function of display condition

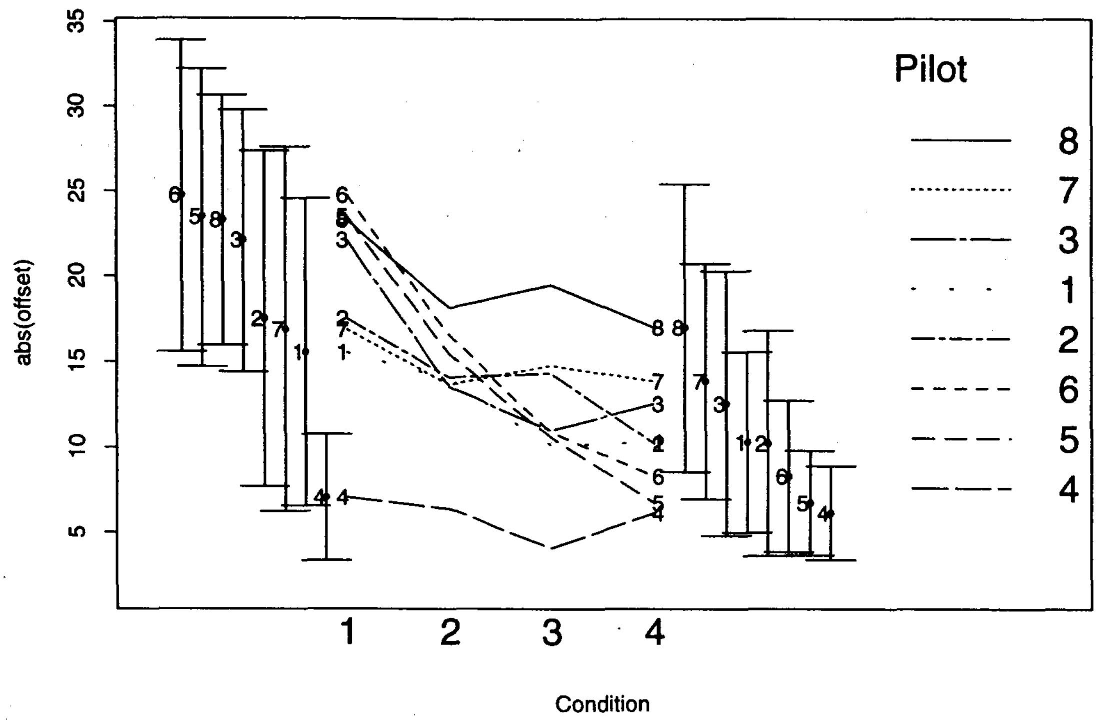

Figure 19. Average RMS values for lateral offset (ft.) at touchdown for each test subject as a function of display condition. Bars representing standard error of the mean are shown for conditions 1 and 4.

SUBJECT	CONDITION 1 (runway only)	CONDITION 2 (r/w with markings)
1	0.5 (-0.3)	0.4 (-0.1)
2	0.9 (-0.3)	0.8 (-0.4)
3	0.7 (-0.1)	0.6 (+0.2)
4	0.8 (-0.1)	0.6 (0.0)
5	0.9 (0.0)	0.8 (+0.2)
6	1.1 (-0.7)	1.1 (-0.4)
7	0.9 (-0.6)	1.1 (-0.3)
8	1.4 (+0.3)	0.8 (+0.3)
Average	0.90	0.78

Table 9. Average RMS track error (deg.) at touchdown for each test subject for display conditions 1 and 2. The numbers in parentheses represent the mean offset.

($F(1, 256) = 6.6$, $p < 0.05$), with a decrease in track error found for the display condition containing runway markings. Differences in performance levels between pilots was also found ($F(7, 256) = 5.6$, $p < 0.001$). Of the eight subjects tested, six produced a lower average track error under condition 2 than condition 1, however, an analysis of these individual results (using a multiple t-test comparison set at the 0.05 level) indicated no significant differences.

Discussion. It is likely that the reduction in RMS track error with display condition 2 is due to the existence of a runway centerline that allows the pilots to line up their approach more accurately.

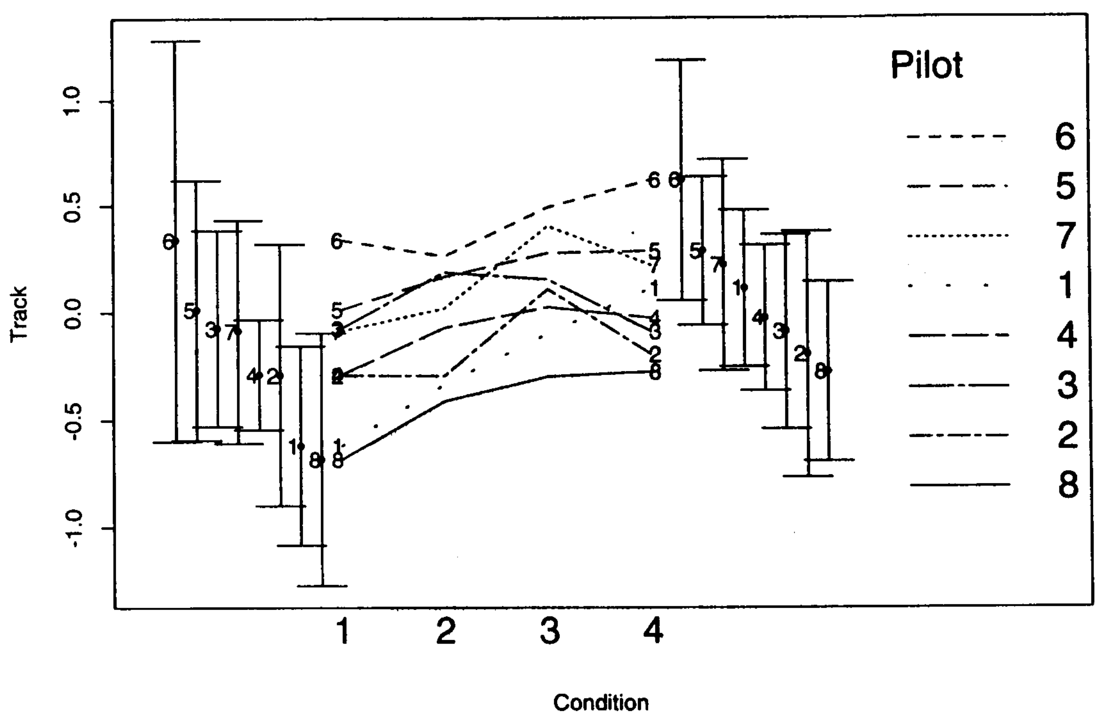

Figure 20. Average RMS track error (deg.) at touchdown for each of the eight subjects as a function of display condition. Bars representing standard error of the mean are shown for conditions 1 and 4.

Bank angle at touchdown. The average RMS bank angle at touchdown for each subject as a function of the two test conditions is shown in Table 10. A graphical representation of this data is provided as part of Figure 21. An analysis of variance indicated a main effect for bank angle as a function of display condition ($F(1, 256) = 6.0$, $p < 0.02$), with an increase in bank angle found for the display condition containing runway markings. Differences in performance levels between pilots was also found ($F(7, 256) = 2.5$, $p < 0.02$). An analysis of the performance of individual pilots (using a multiple t-test comparison set at the 0.05 level) indicated no significant differences between conditions 1 and 2 for any of the subjects.

SUBJECT	CONDITION 1 (runway only)	CONDITION 2 (r/w with markings)
1	0.9 (-0.3)	1.0 (-0.5)
2	1.0 (-0.5)	2.3 (+0.9)
3	2.1 (-0.5)	1.8 (-0.3)
4	2.1 (-0.4)	1.2 (-0.2)
5	1.0 (0.1)	2.0 (0.2)
6	2.3 (-0.4)	3.5 (0.0)
7	1.0 (-0.5)	2.0 (-0.5)
8	3.0 (-0.8)	3.4 (-1.3)
Average	1.68	2.15

Table 10. Average RMS bank angle (degrees) at touchdown for the eight test subjects as a function of display condition. Numbers in parentheses indicate mean bank angle.

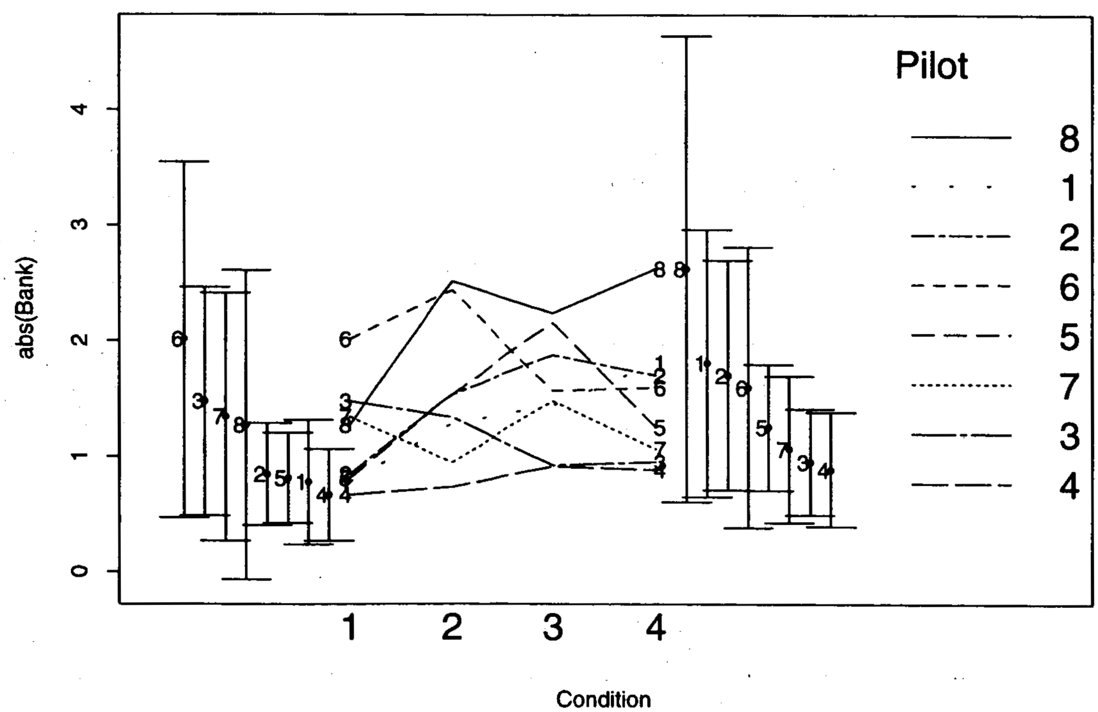

Figure 21. Average RMS bank angle (deg.) at touchdown for each test subject as a function of display condition. The vertical bars represent the standard error of the mean for conditions 1 and 4.

Discussion. The tendency for pilots to land with a greater bank angle in the richer cue condition (condition 2) may be due to the existence of a runway centerline in this display which allows pilots to recognize small lateral errors shortly before touchdown. A last minute attempt to correct these errors may be the cause of a residual bank angle. It is interesting to note that if this last minute correction is occurring, it is not sufficiently strong to throw off the track angle at touchdown which is better for condition 2 than condition 1 (see Table 9).

Interactions between variables. We looked at a number of different interactions between variables: vertical speed vs. landing distance (Fig. 22), vertical speed vs. flare initiation altitude (Fig. 23), landing distance vs. flare altitude (Fig. 24), and lateral offset vs. distance (Fig. 25).

Vertical speed vs. landing distance. Looking at Figure 22 we see that vertical speed seems to be higher for shorter landings for both condition 1 and 2. An analysis of covariance indicated a correlation coefficient of -0.41 for condition 1 and -0.53 for condition 2 with the difference between the two correlations being significant.

Vertical speed vs. flare initiation altitude. Because of the way we were measuring flare initiation altitude in this part of the experiment it is likely that data points with a flare altitude above about 70-80 ft. are representative of earlier pilot maneuvers and do not represent the true initiation of flare. Examination of the scatter plots for conditions 1 and 2 in Figure 23 indicates a minimal interaction. An analysis of covariance, considering only points below 70 ft., resulted in a correlation coefficient of -0.27 for condition 1, and -0.31 for condition 2.

Landing distance vs. flare altitude. As explained in the previous comparison, it is necessary to ignore data points above about 70 ft. flare altitude. For the display with runway markings (condition 2, Fig. 24) we see an increase in landing distance with an increase in flare altitude with this relationship being especially strong for the lower flare altitudes. Condition 1 appears to show this same trend, but to a considerably lesser extent. An analysis of covariance, considering only points below 70 ft., indicated a correlation coefficient of 0.28 for condition 1 and 0.47 for condition 2, with the difference between the correlations being significant.

Lateral offset vs. landing distance. Figure 25 clearly shows a tighter, more controlled pattern of lateral offset performance in condition 2 as compared to condition 1. Looking at the variability in lateral offset as a function of landing distance indicates a slight trend toward lower variability with longer landing distances.

Subjective Results - Experiment IIA.

Pilots were asked to indicate their preferences for the four display conditions. The rank order scores are presented in Table 11. For the two displays used in the familiar size experiment (Conditions 1 and 2) all but one pilot preferred the display with runway markings over the runway-only display (2 over 1). The subjects were also asked to indicate the magnitude of the differences between conditions on a "strong," "medium," or "slight" scale. Condition 2 was "strongly" preferred over condition 1.

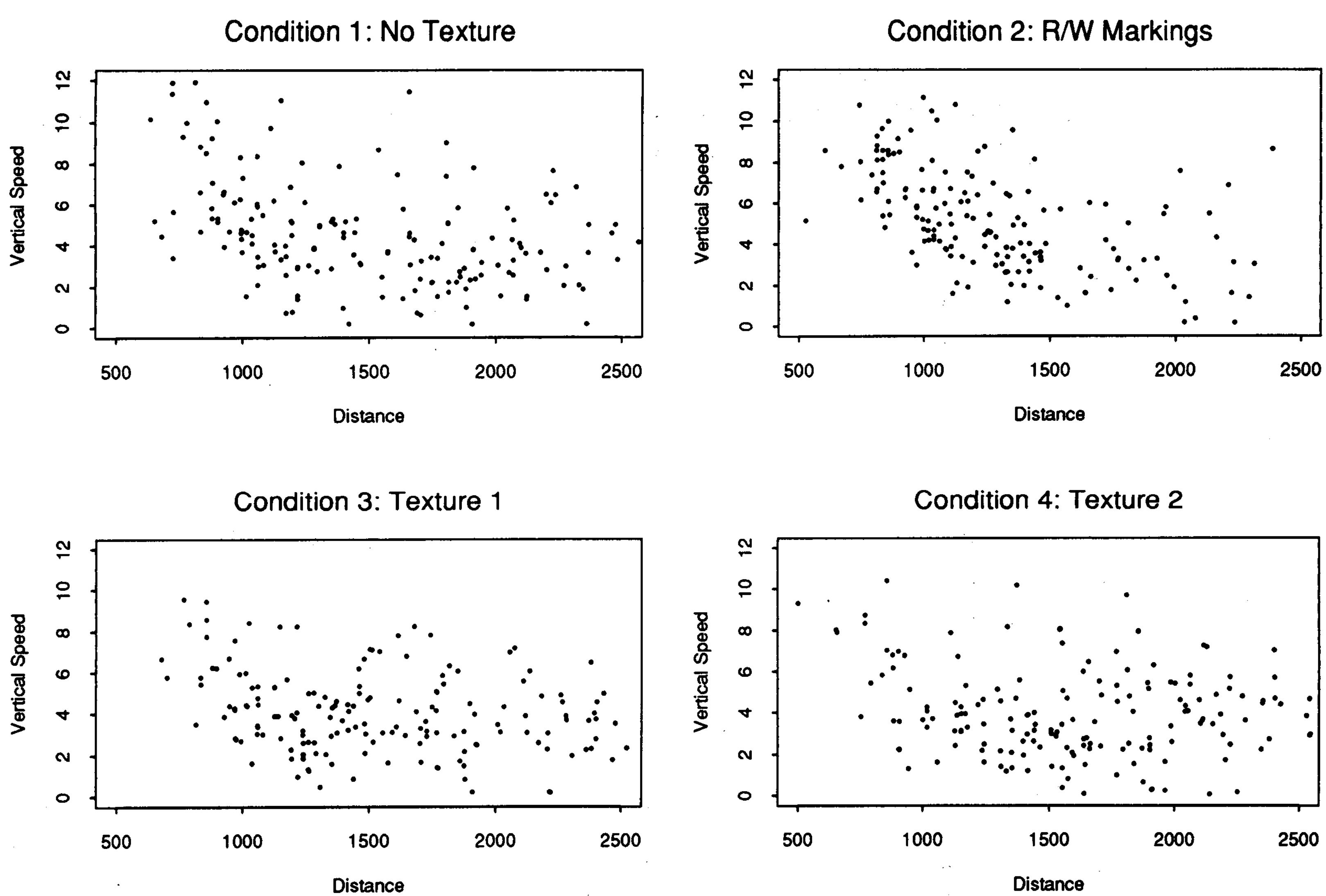

Figure 22. Vertical speed at touchdown (ft./sec.) as a function of landing distance (ft.) at touchdown for each of the four display conditions. Data points represent all landings for all pilots for a total of 160 landing per condition.

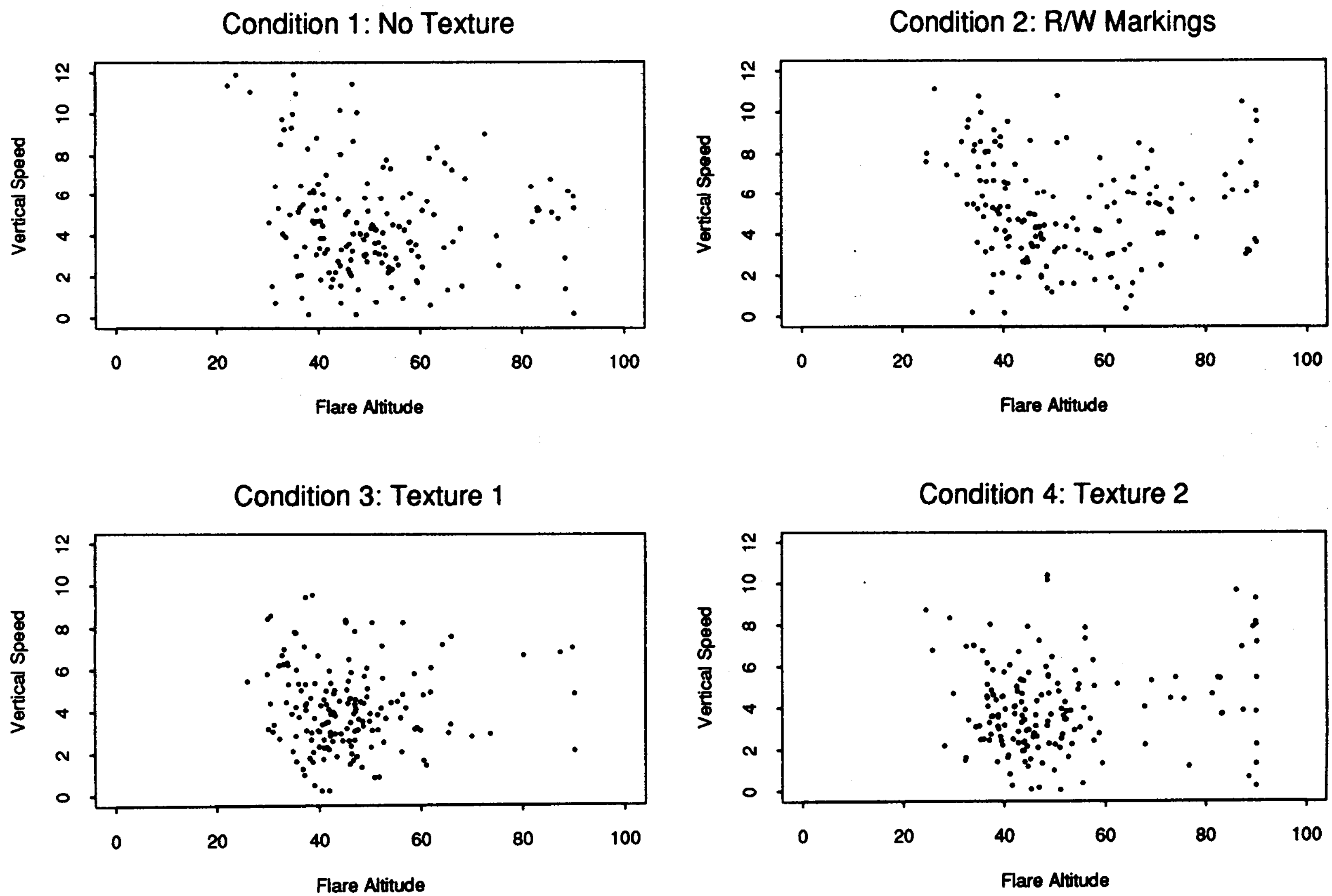

Figure 23. Vertical speed at touchdown (ft./sec.) as a function of flare altitude (ft.) at touchdown for each of the four display conditions. Data points represent all landings for all pilots for a total of 160 landing per condition.

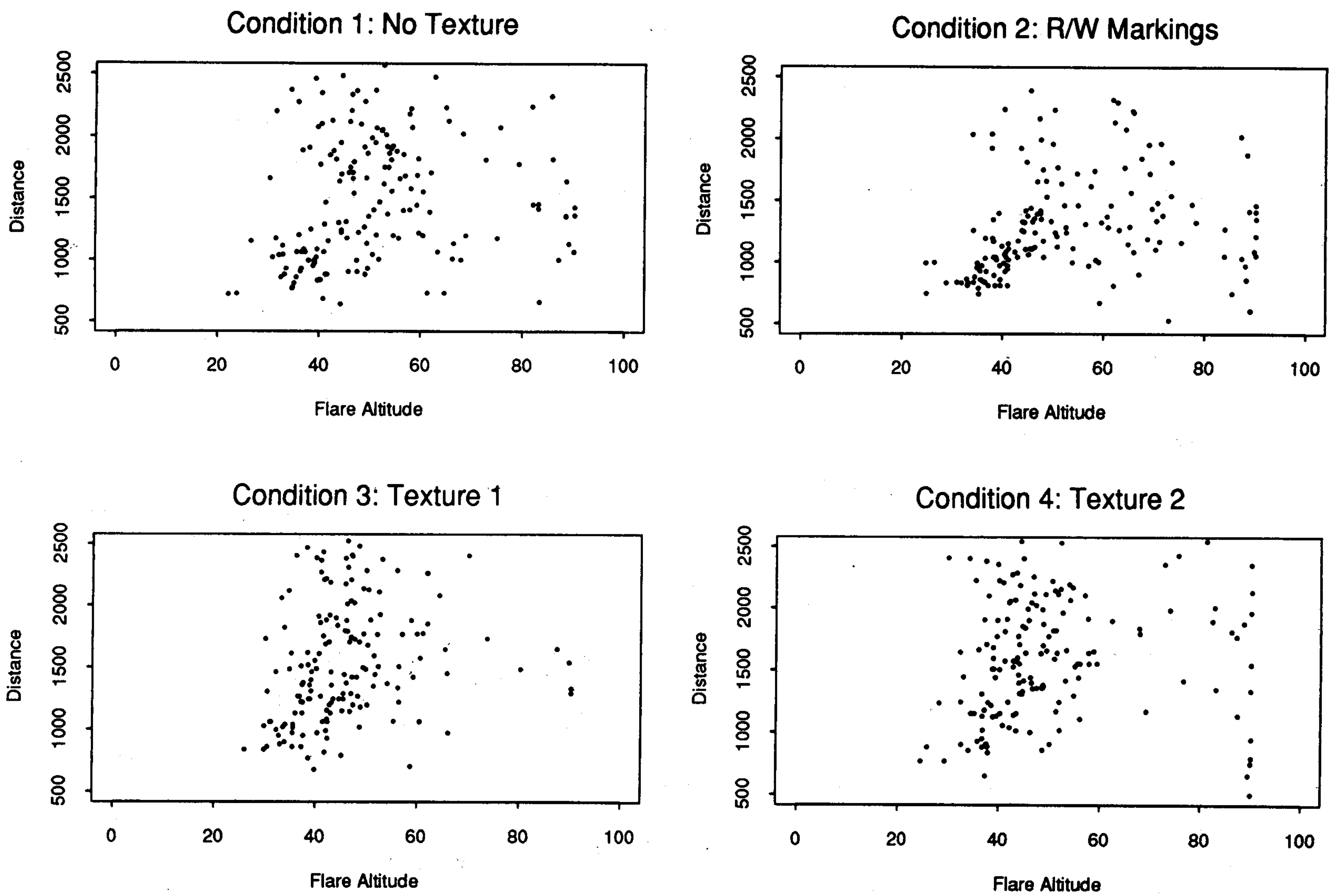

Figure 24. Landing distance at touchdown (ft.) as a function of flare altitude (ft.) at touchdown for each of the four display conditions. Data points represent all landings for all pilots for a total of 160 landing per condition.

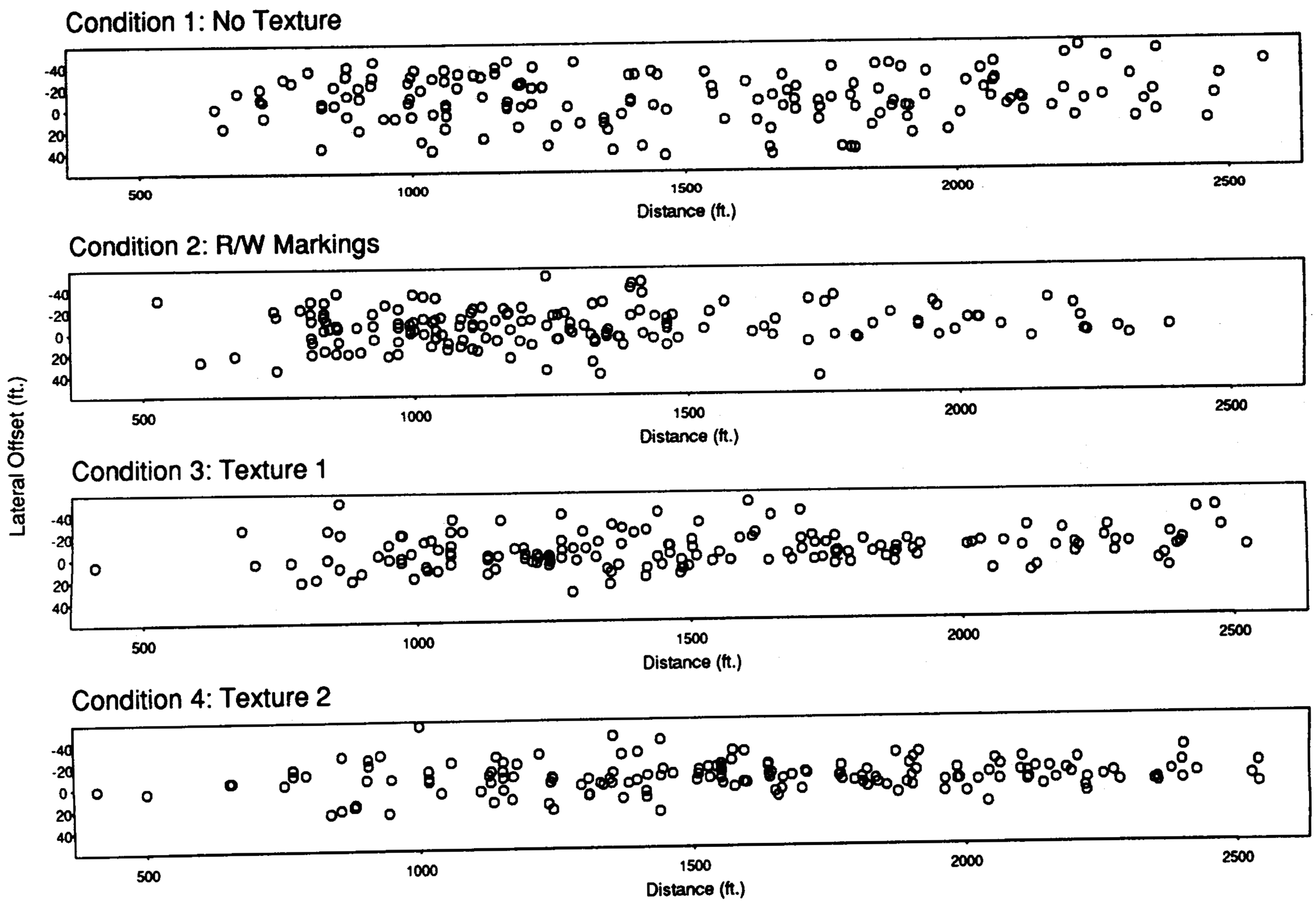

Figure 25. Landing position as a function of display condition. Data points represent all landings for all pilots for a total of 160 landing per condition.

SUBJECTS	CONDITION 1 (plane r/w)	CONDITION 2 (r/w markings)	CONDITION 3 (texture 1)	CONDITION 4 (texture 2)
1	4	3	2	1
2	4	3	1	1
3	4	3	2	1
4	4	1	3	2
5	4	3	2	1
6	4	2	2	1
7	3	4	2	1
8	4	3	1	1

Table 11. Subjective preferences of each pilot for the four display conditions. Rank ordering foes from 1 to 4 (most preferred to least preferred).

Experiment IIB - Texture Level Comparison.

Empirical Results.

Two different sequences of increasing levels of texture were considered. In the first, two display conditions from Experiment I were re-run under the high workload conditions. These include the no texture condition (condition 1, Fig. 12) and a full texture condition (condition 3, Fig. 14). The second sequence included the same no texture condition and a second full texture condition (condition 4, Fig. 15).

Vertical speed at touchdown. The average vertical speed at touchdown for each of the eight test subjects as a function of texture level is shown in Table 12. A graphical representation of this data is provided as part of Figure 16. An analysis of variance indicated a main effect for vertical speed at touchdown as a function of texture level for the condition 1 and 4 comparison ($F(1, 304) = 5.5$, $p < 0.02$), with a decrease in vertical speed found for the display containing a greater amount of texture (condition 4). Analysis of the condition 1 and 3 comparison did not show a main effect for vertical speed ($F(1, 304) = 1.4$, $p < 0.24$). Differences in performance levels between pilots was also found for the condition 1 and 4 comparison ($F(7, 304) = 22.6$, $p < 0.001$), and for the condition 1 and 3 comparison ($F(7, 304) = 12.7$, $p < 0.001$). An analysis of the performance of individual pilots (using a multiple t-test comparison set at the 0.05 level) indicated only one significant difference. That was for pilot 6 in the condition 1 and 3 comparison.

Discussion. The results of this experiment differ from those of Experiment I in which an increase in vertical velocity at touchdown was found to correspond to an increase in texture density. The present results indicating a performance benefit from an increase in texture (for the condition 1 and 4 comparison) are more in keeping with our original expectations based on related experimental findings in the literature. These results tend to indicate that workload level can effect the interaction between perceptual cues and pilot flying performance. With the higher and more realistic workload of Experiment II the pilot's actual performance more closely matched their subjective display preferences.

SUBJECT	CONDITION 1 (no texture)	CONDITION 3 (texture 1)	CONDITION 4 (texture 2)
1	2.5 (1.4)	3.1 (1.6)	2.0 (1.6)
2	4.8 (1.9)	4.7 (2.2)	4.1 (1.9)
3	3.3 (1.3)	4.1 (1.5)	3.4 (1.7)
4	5.3 (2.2)	4.9 (2.1)	4.6 (1.5)
5	7.4 (3.0)	5.4 (1.4)	6.5 (2.4)
6	6.0 (2.2)	3.8 (2.6)	4.6 (2.0)
7	3.1 (1.7)	3.7 (1.8)	2.8 (0.9)
8	4.3 (2.9)	3.8 (2.2)	4.0 (1.7)
Average	4.59	4.19	4.00

Table 12. Average vertical speed at touchdown (ft./sec.) for test subjects as a function of display texture level. The numbers in parentheses indicate standard deviations. Displays for condition 1, 3, and 4 are depicted in Figs. 12, 14, and 15 respectively.

Landing distance. The average landing distance at touchdown for each of the eight test subjects as a function of texture level is shown in Table 13. A graphical representation of this data is provided as part of Figure 17. An analysis of variance found no main effect for landing distance at touchdown as a function of texture level for either the condition 1 and 3 comparison ($F(1, 304) = 0.5$, $p < 0.48$) or the condition 1 and 4 comparison ($F(1, 304) = 3.8$, $p < 0.052$). An analysis of the performance of individual pilots (using a multiple t-test comparison set at the 0.05 level) indicated two significant differences. These were for pilots 2 and 5 in the condition 1 and 3 comparison.

Discussion. Again, the present results differ from those of Experiment I. The decrease in landing distance with increased texture found in the earlier experiment did not appear in the higher workload situation. While not reaching significance, there is actually a trend in the opposite direction.

Flare initiation altitude. The average flare initiation altitude at touchdown for each of the eight test subjects as a function of texture level is shown in Table 14. A graphical representation of this data is provided as part of Figure 16. An analysis of variance indicated a main effect for flare initiation altitude as a function of texture level for the condition 1 and 3 comparison ($F(1, 304) = 13.5$, $p < 0.001$), with a decrease in flare initiation altitude found for the display containing a greater amount of texture (condition 3). Analysis of the condition 1 and 4 comparison did not show this main effect ($F(1, 304) = 0.2$, $p < 0.65$). Differences in performance levels between pilots was also found ($F(7, 304) = 4.7$, $p < 0.001$ for the condition 1 and 3 comparison; $F(7, 304) = 2.3$, $p < 0.04$ for the condition 1 and 4 comparison). An analysis of the performance of individual pilots (using a multiple t-test comparison set at the 0.05 level) indicated only one significant difference. That was for pilot 2 in the condition 1 and 3 comparison.

SUBJECT	CONDITION 1 (no texture)	CONDITION 3 (texture 1)	CONDITION 4 (texture 2)
1	1807 (363)	1616 (451)	1843 (269)
2	1765 (410)	1414 (382)	1781 (492)
3	1724 (478)	1859 (523)	1704 (464)
4	1283 (419)	1212 (383)	1349 (543)
5	1055 (317)	1564 (317)	1351 (519)
6	1275 (526)	1718 (589)	1687 (511)
7	1317 (379)	1498 (454)	1593 (415)
8	1654 (502)	1314 (390)	1357 (340)
Average	1485	1524	1583

Table 13. Average landing distance (ft.) at touchdown for each of the test subjects as a function of texture condition. Numbers in parentheses are standard deviations. Displays for condition 1, 3, and 4 are depicted in Figs. 12, 14, and 15 respectively.

Discussion. The results for the condition 1 and 3 comparison are the same as those found in Experiment I - a lower average flare initiation altitude for displays with a greater texture content. The failure to find this same effect in the condition 1 and 4 comparison indicates that flare initiation altitude may depend on the type of texture as well as the amount.

SUBJECT	CONDITION 1 (no texture)	CONDITION 3 (texture 1)	CONDITION 4 (texture 2)
1	48.0 (7.8)	46.8 (8.0)	51.4 (11.7)
2	49.3 (11.2)	42.3 (12.6)	46.5 (16.3)
3	54.4 (9.0)	46.6 (9.3)	53.4 (17.2)
4	61.4 (14.8)	54.3 (16.7)	51.5 (17.3)
5	44.2 (17.4)	44.6 (12.2)	55.0 (22.2)
6	47.6 (12.5)	42.3 (7.9)	45.6 (6.6)
7	50.0 (17.9)	46.0 (9.6)	46.6 (12.5)
8	54.0 (20.7)	45.4 (15.4)	54.5 (19.8)
Average	51.1	46.0	50.6

Table 14. Average flare initiation altitude (ft.) for each of the test subjects as a function of texture condition. Numbers in parentheses are standard deviations. Displays for condition 1, 3, and 4 are depicted in Figs. 12, 14, and 15 respectively.

Lateral offset. The average RMS lateral offset at touchdown for each of the eight test subjects as a function of texture level is shown in Table 15. A graphical representation of this data is provided as part of Figure 19. An analysis of variance indicated a main effect for lateral offset as a function of texture level for the condition 1 and 3 comparison ($F(1, 304) = 27.7$, $p < 0.001$) and the condition 1 and 4 comparison ($F(1, 304) = 43.4$, $p < 0.001$), with a decrease in offset found for the display containing a greater amount of texture. Differences in performance levels between pilots was also found ($F(7, 304) = 7.5$, $p < 0.001$ for the condition 1 and 3 comparison; $F(7, 304) = 6.0$, $p < 0.001$ for the condition 1 and 4 comparison). A pilot by condition interaction was found for the condition 1 and 4 comparison ($F(7, 301) = 3.1$, $p < 0.003$). An analysis of the performance of individual pilots (using a multiple t-test comparison set at the 0.05 level) indicated three significant differences. These were for pilots 3, 5, and 8 in the condition 1 and 4 comparison.

Discussion. The strong finding of decreased lateral offset for display conditions 3 and 4 as compared to condition 1 is most likely due to the existence of a runway centerline in these displays and not the existence of texture.

SUBJECT	**CONDITION 1** (no texture)	**CONDITION 3** (texture 1)	**CONDITION 4** (texture 2)
1	8.9 (-3.0)	7.2 (+2.7)	7.3 (-3.6)
2	23.7 (-5.4)	19.2 (-2.9)	13.9 (-6.7)
3	24.8 (-13.4)	16.9 (-6.0)	16.9 (-6.2)
4	23.0 (-5.6)	18.0 (+2.4)	17.1 (+0.9)
5	26.7 (+2.2)	17.5 (-11.7)	8.1 (-2.8)
6	25.6 (-17.8)	22.0 (-14.1)	21.0 (-15.5)
7	20.3 (-7.5)	18.5 (-2.9)	12.8 (-3.9)
8	28.1 (-8.6)	18.4 (-4.4)	10.5 (-1.1)
Average	22.6	15.0	13.5

Table 15. Average RMS lateral offset (ft.) for each of the test subjects as a function of texture condition. Numbers in parentheses are mean offset values. Displays for condition 1, 3, and 4 are depicted in Figs. 12, 14, and 15 respectively.

Track error at touchdown. The average RMS track error at touchdown for each of the eight test subjects as a function of texture level is shown in Table 16. A graphical representation of this data is provided as part of Figure 20. An analysis of variance indicated a main effect for track error as a function of texture level for the condition 1 and 3 comparison ($F(1, 304) = 11.0$, $p < 0.002$) and the condition 1 and 4 comparison ($F(1, 304) = 6.9$, $p < 0.01$), with a decrease in track error found for the display containing a greater amount of texture. Differences in performance levels between pilots was also found ($F(7, 304) = 5.3$, $p < 0.001$ for the condition 1 and 3 comparison; $F(7, 304) = 2.6$, $p < 0.02$ for the condition 1 and 4 comparison).

Discussion. The decrease in track error found for the high texture level conditions is probably due to the existence of runway markings as opposed to a greater level of texture. The runway centerline may provide cues that allow pilots to achieve better alignment during the final approach phase.

SUBJECT	CONDITION 1 (no texture)	CONDITION 3 (texture 1)	CONDITION 4 (texture 2)
1	0.5 (-0.3)	0.3 (0.0)	0.5 (0.0)
2	0.9 (-0.3)	0.9 (+0.1)	0.8 (-0.2)
3	0.7 (-0.1)	0.6 (+0.2)	0.7 (-0.1)
4	0.8 (-0.1)	0.7 (+0.4)	0.8 (+0.2)
5	0.9 (0.0)	0.7 (+0.3)	0.6 (+0.3)
6	1.1 (-0.7)	0.9 (-0.3)	0.7 (-0.3)
7	0.9 (-0.6)	0.5 (-0.1)	0.5 (+0.1)
8	1.4 (+0.3)	0.7 (+0.5)	1.0 (+0.6)
Average	0.90	0.66	0.70

Table 16. Average RMS track error at touchdown (deg.) for each of the test subjects as a function of texture condition. Numbers in parentheses are mean error values. Displays for condition 1, 3, and 4 are depicted in Figs. 12, 14, and 15 respectively.

Bank angle at touchdown. The average RMS bank angle at touchdown for each of the eight test subjects as a function of texture level is shown in Table 17. A graphical representation of this data is provided as part of Figure 21. An analysis of variance indicated a main effect for bank angle at touchdown as a function of texture level for the condition 1 and 3 comparison ($F(1, 301) = 9.1$, $p < 0.003$) and the condition 1 and 4 comparison ($F(1, 301) = 6.0$, $p < 0.02$), with an increase in bank angle found for the display containing a greater amount of texture. An interaction between pilots and conditions was found for the condition 1 and 3 comparison ($F(7, 301) = 2.4$, $p < 0.03$), but not the condition 1 and 4 comparison. An analysis of the performance of individual pilots (using a multiple t-test comparison set at the 0.05 level) indicated only one significant differences. This was for pilot 5 in the condition 1 and 3 comparison.

Discussion. The tendency for pilots to land with a greater bank angle in the high texture level conditions (conditions 3 and 4) is probably due to the existence of a runway centerline in this display which allows them to recognize small lateral errors shortly before touchdown. We speculate that last minute attempts to correct these errors may result in a residual bank angle at touchdown.

Interactions between variables. We looked at a number of different interactions between variables: vertical speed vs. landing distance (Fig. 22), vertical speed vs. flare initiation altitude (Fig. 23), landing distance vs. flare altitude (Fig. 24), and lateral offset vs. distance (Fig. 25).

Vertical speed vs. landing distance. Looking at Figure 22 we see a decreasing trend for higher vertical speeds at touchdown with increased landing distance between conditions 1, 3 and 4

SUBJECT	CONDITION 1 (no texture)	CONDITION 3 (texture 1)	CONDITION 4 (texture 2)
1	0.9 (-0.3)	1.1 (-0.2)	1.1 (-0.1)
2	1.0 (-0.5)	2.3 (-0.8)	2.2 (0.5)
3	2.1 (-0.5)	1.2 (+0.3)	1.2 (0.0)
4	2.1 (-0.4)	2.0 (+0.2)	1.4 (0.0)
5	1.0 (0.1)	2.9 (-0.3)	1.5 (+0.2)
6	2.3 (-0.4)	3.3 (-0.5)	3.9 (-0.7)
7	1.0 (-0.5)	2.0 (-0.1)	2.5 (+0.3)
8	3.0 (-0.8)	2.9 (+1.4)	2.4 (-0.4)
Average	1.68	2.21	2.03

Table 17. Average RMS bank angle at touchdown (deg.) for each of the test subjects as a function of texture condition. Numbers in parentheses are mean bank angle values. Displays for condition 1, 3, and 4 are depicted in Figs. 12, 14, and 15 respectively.

respectively. An analysis of covariance produced the following correlation coefficients: condition 1 = -0.41, condition 3 = -0.30, and condition 4 = -0.20. A comparison of these correlations indicated significant differences between those for conditions 1 and 4, and conditions 3 and 4 (z-statistic for differences in correlations).

Vertical speed vs. flare initiation altitude. These interactions are shown in Figure 23. As mentioned previously, because of the way we were measuring flare initiation altitude in this part of the experiment it is likely that data points with a flare altitude above about 70-80 ft. are representative of earlier pilot maneuvers and do not represent true initiation of flare. Looking at the scatter plots for conditions 1, 3, and 4, and considering only points below 70 ft. flare altitude, indicates no substantial interaction between variables for any of the conditions. The tighter grouping of points in the condition 3 and 4 plots indicate the lower variability of flare altitude and vertical speed for these conditions but no interaction between them.

Landing distance vs. flare initiation altitude. Considering conditions 1, 3, and 4 in Figure 24, and ignoring data points above 70 ft., we see an increase in landing distance with increased flare initiation altitude, although this correlation appears weaker for condition 1.

Lateral offset vs. landing distance. Figure 25 clearly shows a tighter, more controlled pattern of lateral offset performance in conditions 3 and 4 as compared to condition 1. Considering lateral offset as a function of landing distance indicates that there may be a slight decrease in the variability of lateral offset as landing distance increases.

Subjective Results - Experiment IIB.

All pilots expressed a subjective preference for the displays with texture (conditions 3 and 4) over the display with no texture (condition 1). These results are depicted in Table 18. In all cases the preference was a strong one.

SUBJECT S	CONDITION 1 (plane r/w)	CONDITION 3 (texture 1)	CONDITION 4 (texture 2)
1	3	2	1
2	3	1	1
3	3	2	1
4	3	2	1
5	3	2	1
6	3	2	1
7	3	2	1
8	3	1	1

Table 18. Pilot's subjective preferences for displays as a function of level of texture. Preferences range from 1 (most preferred) to 3 (least preferred).

Experiment IIC - Texture Type Comparison.

Empirical results.

In this experiment we compared pilot performance when flying two displays that differed in the type of texture pattern laid down on the runway. One was the same as the full texture scene used in Experiment I (condition 3, Fig. 14). The other used a new texture pattern on the runway (condition 4, Fig. 15). Both displays contained standard runway markings, and had a similar texture pattern on the area surrounding the runway.

Vertical speed at touchdown. The average vertical speed at touchdown for each of the eight test subjects as a function of texture type is shown in Table 19. A graphical representation of this data is provided as part of Figure 16. An analysis of variance indicated no difference in vertical speed at touchdown between the two texture types ($F(1, 304) = 1.4$, $p < 0.23$). An analysis of the performance of individual pilots (using a multiple t-test comparison set at the 0.05 level) also indicated no significant differences between the two conditions. A difference in performance levels between pilots was found ($F(7, 304) = 13.7$, $p < 0.001$).

Discussion. While the results of Experiment IIB indicated better performance for condition 4 than condition 3 when compared to condition 1, this difference does not show up as significant when conditions 3 and 4 are compared directly, as done here.

Landing distance. The average landing distance at touchdown for each of the eight test subjects as a function of texture type is shown in Table 20. A graphical representation of this data is provided as part of Figure 17. An analysis of variance indicated no difference in landing distance at touchdown between the two texture types ($F(1, 304) = 1.5$, $p < 0.23$). In addition, an analysis of the performance of individual pilots (using a multiple t-test comparison set at the 0.05 level) indicated no significant differences between the two conditions for any of the pilots. A difference in performance levels between pilots was found ($F(7, 304) = 6.78$, $p < 0.001$).

SUBJECT	CONDITION 3 (texture 1)	CONDITION 4 (texture 2)
1	3.1 (1.6)	2.0 (1.6)
2	4.7 (2.2)	4.1 (1.9)
3	4.1 (1.5)	3.4 (1.7)
4	4.9 (2.1)	4.6 (1.5)
5	5.4 (1.4)	6.5 (2.4)
6	3.8 (2.6)	4.6 (2.0)
7	3.7 (1.8)	2.8 (0.9)
8	3.8 (2.2)	4.0 (1.7)
Average	4.19	4.00

Table 19. Average vertical speed (ft./sec.) at touchdown for each of the test subjects as a function of texture type. Numbers in parentheses are standard deviations. Displays for condition 3 and 4 are depicted in Figs. 14 and 15 respectively.

SUBJECT	CONDITION 3 (texture 1)	CONDITION 4 (texture 2)
1	1616 (451)	1843 (269)
2	1414 (382)	1781 (492)
3	1859 (523)	1704 (464)
4	1212 (383)	1349 (543)
5	1564 (317)	1351 (519)
6	1718 (589)	1687 (511)
7	1498 (454)	1593 (415)
8	1314 (390)	1357 (340)
Average	1524	1583

Table 20. Average landing distance (ft.) at touchdown for each of the test subjects as a function of texture type. Numbers in parentheses are standard deviations. Displays for condition 3 and 4 are depicted in Figs. 14 and 15 respectively.

<u>Flare initiation altitude</u>. The average flare initiation altitude at touchdown for each of the eight test subjects as a function of texture level is shown in Table 21. A graphical representation

of this data is provided as part of Figure 18. An analysis of variance indicated a main effect for flare initiation altitude as a function of texture type ($F(1, 304) = 8.7$, $p < 0.004$), with a higher altitude found for condition 4. An analysis of the performance of individual pilots (using a multiple t-test comparison set at the 0.05 level) indicated no significant differences between the two display conditions for any of the subjects. A difference in performance levels between pilots was also found ($F(7, 304) = 2.5$, $p < 0.02$).

Discussion. Flare initiation altitude is the only test measure in which a direct comparison of texture type (conditions 3 and 4) indicates a significant difference. It is interesting to note that the significantly greater flare initiation altitude for condition 4 is not accompanied by a significantly longer touchdown distance (see Table 20). This result weakens any argument for a link between flare altitude and length of landing.

SUBJECT	CONDITION 3 (texture 1)	CONDITION 4 (texture 2)
1	46.8 (8.0)	51.4 (11.7)
2	42.3 (12.6)	46.5 (16.3)
3	46.6 (9.3)	53.4 (17.2)
4	54.3 (16.7)	51.5 (17.3)
5	44.6 (12.2)	55.0 (22.2)
6	42.3 (7.9)	45.6 (6.6)
7	46.0 (9.6)	46.6 (12.5)
8	45.4 (15.4)	54.5 (19.8)
Average	46.0	50.6

Table 21. Average flare initiation altitude (ft.) for each of the test subjects as a function of texture type. Numbers in parentheses are standard deviations. Displays for condition 3 and 4 are depicted in Figs. 14 and 15 respectively.

Lateral offset. The average RMS lateral offset at touchdown for each of the eight test subjects as a function of texture type is shown in Table 22. A graphical representation of this data is provided as part of Figure 19. An analysis of variance indicated no difference in lateral offset at touchdown between the two texture types ($F(1, 256) = 1.38$, $p < 0.24$). An analysis of the performance of individual pilots (using a multiple t-test comparison set at the 0.05 level) indicated no significant differences between the two conditions for any of the pilots. A difference in performance levels between pilots was found ($F(7, 256) = 6.85$, $p < 0.001$).

Track error at touchdown. The average RMS track error (in degrees) at touchdown for each of the eight test subjects as a function of texture type is shown in Table 23. A graphical representation of this data is provided as part of Figure 20. An analysis of variance indicated no difference in track error at touchdown between the two texture types ($F(1, 301) = 0.57$, $p <$

SUBJECT	CONDITION 3 (texture 1)	CONDITION 4 (texture 2)
1	5.3 (-0.9)	7.3 (-3.6)
2	17.9 (-4.3)	13.9 (-6.7)
3	12.8 (-6.9)	16.9 (-6.2)
4	19.1 (-3.7)	17.1 (+0.9)
5	14.6 (-3.3)	8.1 (-2.8)
6	24.6 (-17.6)	21.0 (-15.5)
7	11.7 (2.5)	12.8 (-3.9)
8	13.9 (-2.5)	10.5 (-1.1)
Average	14.99	13.45

Table 22. Average RMS lateral offset (ft.) at touchdown for each of the test subjects as a function of texture type. Numbers in parentheses indicate mean offset for each subject. Displays for condition 3 and 4 are depicted in Figs. 14 and 15 respectively.

0.45). An analysis of the performance of individual pilots (using a multiple t-test comparison set at the 0.05 level) indicated no significant differences between the two conditions for any of the pilots. A difference in performance levels between pilots was found ($F(7, 301) = 4.30$, $p < 0.001$).

SUBJECT	CONDITION 3 (texture 1)	CONDITION 4 (texture 2)
1	0.3 (0.0)	0.5 (0.0)
2	0.9 (+0.1)	0.8 (-0.2)
3	0.6 (+0.2)	0.7 (-0.1)
4	0.7 (+0.4)	0.8 (+0.2)
5	0.7 (+0.3)	0.6 (+0.3)
6	0.9 (-0.3)	0.7 (-0.3)
7	0.5 (-0.1)	0.5 (+0.1)
8	0.7 (+0.5)	1.0 (+0.6)
Average	0.66	0.70

Table 23. Average RMS track error (deg.) at touchdown for each of the test subjects as a function of texture type. Numbers in parentheses are mean track error values. Displays for condition 3 and 4 are depicted in Figs. 14 and 15 respectively.

Bank angle at touchdown. The average RMS bank angle at touchdown for each of the eight test subjects as a function of texture type is shown in Table 24. A graphical representation of this data is provided as part of Figure 21. An analysis of variance indicated no difference in bank angle at touchdown between the two texture types ($F(1, 280) = 0.29$, $p < 0.59$). An analysis of the performance of individual pilots (using a multiple t-test comparison set at the 0.05 level) indicated no significant differences between the two conditions for any of the pilots. A difference in performance levels between pilots was found ($F(7, 280) = 3.13$, $p < 0.004$).

SUBJECT	CONDITION 3 (texture 1)	CONDITION 4 (texture 2)
1	1.1 (-0.2)	1.1 (-0.1)
2	2.3 (-0.8)	2.2 (0.5)
3	1.2 (+0.3)	1.2 (0.0)
4	2.0 (+0.2)	1.4 (0.0)
5	2.9 (-0.3)	1.5 (+0.2)
6	3.3 (-0.5)	3.9 (-0.7)
7	2.0 (-0.1)	2.5 (+0.3)
8	2.9 (+1.4)	2.4 (-0.4)
Average	2.21	2.03

Table 24. Average RMS bank angle (deg.) at touchdown for each of the test subjects as a function of texture type. Numbers in parentheses are mean bank angle values. Displays for condition 3 and 4 are depicted in Figs. 14 and 15 respectively.

Interactions between variables. We looked at a number of different interactions between variables: vertical speed vs. landing distance (Fig. 22), vertical speed vs. flare initiation altitude (Fig. 23), landing distance vs. flare altitude (Fig. 24), and lateral offset vs. distance (Fig. 25). A comparison of the results for the display conditions used in Experiment IIC (condition 3 vs. 4) indicated fairly similar performance patterns between the conditions for each of the interactions.

Subjective Results - Experiment IIC.

Pilots were asked to indicate a preferences for the two texture types. The results are indicated in Table 25. Six of the eight subjects preferred the texture in display condition 4. The other two found the two conditions equally effective and indicated no preference. When a preference was reported it was not a strong one.

SUBJECTS	CONDITION 3 (texture 1)	CONDITION 4 (texture 2)
1	2	1
2	1	1
3	2	1
4	2	1
5	2	1
6	2	1
7	2	1
8	1	1

Table 25. Subjective preferences for texture type. Preferences range from 1 (most preferred) to 2 (least preferred).

GENERAL DISCUSSION AND CONCLUSIONS

In this section we discuss a number of points relevant to our experimental results and indicate conclusions that may be drawn from these results. We will not, in general, repeat the various points discussed, and conclusions drawn, in the Results sections of the different experiments. The reader needs to refer back to these sections to understand the full scope of our findings.

Experiment II was conducted in large part to explore the unexpected results found in Experiment I and a previous experiment by Regal and Whittington (1993). Experiment I results were ambiguous as to the advantages of increasing the amount of texture in displays flown by pilots. The Regal and Whittington (1993) experiment also produced inconclusive results regarding the advantages of increased levels of familiar size cues. In both these experiments the empirical results were not what would have been predicted from the relevant literature, or indicative of the pilot's strong subjective preferences. In Experiment II we increased the difficulty, and thus the workload, of the flying task presented to the pilots to determine if this might affect the results.

Experiment IIA examined the effects of familiar size cues on pilot performance. It was a re-examination of the Regal and Whittington (1993) experiment under higher workload conditions. The results differed in a number of ways. Vertical speeds at touchdown revealed significantly harder landings with increased cues, while having shown only a trend in this direction in the Regal and Whittington study. Flare initiation altitude, which decreased with an increase in cue strength in the Regal and Whittington experiment showed no difference in Experiment IIA. Finally, landing distances got shorter with increased cues, similar to the Regal and Whittington study. Thus, Experiment IIA failed to show an advantage to adding familiar size cues to displays, and provided no evidence that increased workload conditions enhance the benefits of displays containing familiar size cues. There is actually the possibility of a decrement under high workload conditions.

In Experiment IIB the effects of increased display texture were investigated. This is a re-examination of Experiment I (of the present report) under higher workload conditions. Differences in the two experiments include the finding that sink rate at touchdown increased with increasing texture in Experiment IIB - the opposite of findings from Experiment I. Another difference is that the shorter landing distances found with increased texture in Expedient I was not found in Experiment IIB. These results are interpreted as indicating a performance advantage when using displays containing increased texture levels, and an indication that the effects of display texture may be influenced by the workload level.

Experiment IIC compared two patterns of high resolution terrain texture mapping. There was a small, but significant, subjective pilot preference for one of the texture types, and a higher flare initiation altitude for this same display. Otherwise, there was no significant differences in performance between the two textures. While the empirical results did not indicate a clear difference between textures types, it is our feeling, based in part on the subjective reports, and in part on the positive results of texture over no texture (see Experiment IIB), that texture type can make a difference on pilot spatial situational awareness. We feel that further work on texture optimization and the exploration of concepts such as emerging texture should be pursued.

While we were able to measure a number of different pilot performance variables in the present experiments, it would have been desirable to look at the actual flare to landing performance for individual experimental trials. This would involve looking at plots of vertical speed as a function of altitude. Unfortunately, we were not able to carry out this analysis.

Some of the individual differences found among pilots were not expected. We were not surprised to find qualitative differences in the performance levels among pilots, although they were larger than expected in some cases. What was surprising was the degree of qualitative differences found among the subjects. Differential performance trends (e.g., an increase vs. a decrease in sink rate as a function of condition) were not uncommon, and often produced a statistical interaction between pilots and conditions that complicated the interpretation of results. Our initial assumption was that we were dealing with fairly basic perceptual functioning and that there would not be strong qualitative differences between subjects, just as these qualitative differences do not generally show up in basic perceptual or psychophysical experiments. Given our initial assumptions, we ran substantial numbers of trials on relatively few subjects. In further work we would increase the number of subjects to increase our ability to make conclusions concerning population and sub-population trends. The possibility of substantial individual differences between pilots regarding the effectiveness of perceptual cues can significantly affect display design since these designs must accommodate the large majority of pilots.

We would like to conclude with a word of caution regarding interpretation of the present results. The findings of the present set of experiments show only limited benefits from adding perceptual complexity to displays. We feel it would be very premature to assume from these results alone that enhanced perceptual cues of some sort will not add significantly to pilot performance using a synthetic vision system. We looked only at a very limited number of possible visual cues. We do not know if more optimal cues for familiar size, texture mapping, and texture type would have resulted in superior performance. In addition, there are many other types of cues and combinations of cues that were not considered. Another important point is that we were able to look only at certain aspects of pilot performance. We did not, for example examine the pilot's ability to maintain a specified flight path, and we did not measure whether pilots had overall higher levels of situational awareness, a very important, but hard to measure, concept.

We believe the present results do indicate that determining which perceptual cues are most important for perspective displays, and discovering how these cues should be rendered is not an easy problem. Human perception is very complex and the applied research necessary to tell us which cues, or combination of cues, are needed to optimize pilot performance is still in its early stages.

REFERENCES

Abbott, T. S. (1993). Function categories for future flight deck designs. NASA Technical Memorandum 109005.

Andre, A. D. and Johnson, W. W. (1992). Stereo effectiveness evaluation for precision hover tasks in a helmet-mounted display simulator. *Proceedings of the 1992 Conference on Systems, Man and Cybernetics.*

Arditi, A. (1986). Binocular vision. in K. R. Boff, L. Kaufman, and J. P. Thomas (Eds.), *Handbook of Perception and Human Performance*, Wiley, New York.

Awe, C. A., Johnson, W. W. and Schmitz, F. (1989). Inflexibility in selecting the optical basis for perceiving speed. *Proceedings of the Human Factors Society 33rd Annual Meeting*, pp. 1440-1444. (h)

Bainbridge, L. (1983). Ironies of automation. *Automatica, 19*, 775-779.

Barfield, W., Rosenberg, C. and Kraft, C. (1990). Relationship between scene complexity and perceptual performance for computer graphics simulations. *Displays*, October, 1990, pp. 179-185.

Boff, K. R., Kaufman, L., and Thomas, J. P., (Eds.). (1986). *Handbook of Perception and Human Performance: Vol. 1. Sensory Processes and Perception*. New York, Wiley.

Buckland, G. H. (1980). Flight simulator runway visual textural cues for landing. *Proceedings of the Human Factors Society - 24th Annual Meeting*, 285-287.

Carterette, E. C. and Friedman, M. P. (Eds.)(1975). *Handbook of Perception*, Academic Press, New York

Denton, G. G. (1980). The influence of visual pattern on perceived speed. *Perception, 9*, 393-402.

Flach, J. M., Hagen, B. A., and Larish, J. F. (1989). Sources of information in optic flow for the regulation of altitude. EPRL-89-11.

Foyle, D. C., Kaiser, M. K., and Johnson, W. W. (1992). Visual cues in low-level flight: Implications for pilotage, training, simulation, and enhanced/synthetic vision systems. *American Helicopter Society 48th Annual Forum, Vol. 1*, 253-260.

Gibson, J. J. (1950). *The Perception of the Visual World*. Houghton Mifflin, Boston.

Graham (Ed.) (1966). *Vision and Visual Perception*, John Wiley and Sons, Inc. New York.

Grantham, W. D. (1989). Comparison of flying qualities derived from in-flight and ground based simulators for a jet-transport airplane for the approach and landing pilot task. NASA Technical Paper 2962.

Hochberg, J. (1971). Space and movement. In J. W. Kling and L. A. Riggs (Eds.), *Woodworth and Schlossberg's Experimental Psychology, (3rd ed.)* Holt, Rinehart and Winston, New York.

Jewel, J. W. and Stickle, J. W. (1962). Landing-contact conditions for turbine-powered aircraft. In Operation Experiences of Turbine-Powered Commercial Transport Airplanes. NASA TN D-1392, p 10-70.

Johnson, W. W., Bennett C. T., O'Donnell, K, and Phatak A. V. (1988). Optical variables useful in the active control of altitude. Presented at the 23rd Annual Conference on Manual Control, Cambridge, MA.

Johnson, W. W., Tsang, P. S., Bennett, C. T., and Phatak, A. V. (1989). The visually guided control of simulated altitude. *Aviation Space and Environmental Medicine, 60,* pp.152-156.

Kaiser, M. K., Perrone, J. A., Andersen, G. J., Lappin, J. S. and Proffitt D. R. (1990). The effect of motion on surface slant perception. Presented at the Annual Meeting of the Association for Research in Vision and Ophthalmology, Sarasota, FL.

Kantowitz, B. and Sorkin, R. (1987). Allocation of functions. In *Handbook of Human Factors,* Ed. G. Salvendy. John Wiley and Sons.

Kaufman, L. (1974). *Sight and Mind,* Oxford University Press, New York.

Kelly, L., Flach, J. M., Garness, S., and Warren, R. (1993). Altitude control: Effects of texture and global optical flow. *Proceedings of the Seventh International Symposium on Aviation Psychology.*

Kleiss, J. A., Curry, D. G., and Hubbard D. C. (1988). Effect of three-dimensional object type and density in simulated low-level flight. In *Proceedings of the Human Factors Society 32nd Annual Meeting,* Santa Monica , CA, pp.1299-1303.

Kleiss, J. A. and Hubbard, D. C. (1993). Effects of three types of flight simulator visual scene detail on detection of altitude change. *Human Factors, 35,* 653-671.

Langewiesche, W. (1944). *Stick and Rudder,* McGraw-Hill Book Co, Inc., New York

Lintern, G., Sheppard, D. J., Parker, D. L., Yates, K. E. and Nolan, M. D. (1989). Simulator design and instructional features for air-to-ground attack: A transfer study. *Human Factors, 31,* 87-99.

Lintern, G. D. C., Thomley-Yates, K. E., Nelson, B. E., and Roscoe, S. N. (1987). Content, variety, and augmentation of simulated visual scenes for teaching air-to-ground attack, *Human Factors, 29*, pp. 45-59.

Lintern, G. and Walker M. B. (1991). Scene content and runway breadth effects on simulated landing approaches. *The International Journal of Aviation Psychology, 1*, 117-132.

Mann, T. L. (1987). Autonomous landing guidance concept - The effects of video and symbology dynamics on pilot performance. Proceedings of the Sixth Aerospace Behavioral Engineering Technology Conference, 131-138. SAE Technical Paper Series 872390.

McClumpha, A. J., James, M., Green, R. G. and Belyavin, A. J. (1991). Pilots' attitudes to cockpit automation. In *Proceedings of the 35th Annual Meeting of the Human Factors Society*, Santa Monica, CA

Morrison, J. D. and Whiteside, T. C. D. (1984). Binocular cues in the perception of distance of a point source of light. *Perception, 13*, 555-566.

Nataupsky, M., and Crittendon L. (1988).Stereo 3-D and non-stereo presentations of a computer-generated pictorial primary flight display with pathway augmentation. P*roceedings of the AISS/IEEE 8th Digital Avionics Systems Conference*.

Owen, D. H., Wolpert, L., and Warren, R. (1983). Effects of optical flow acceleration, edge acceleration, and viewing time on the perception of egospeed acceleration. NASA Scientific and Technical Information Facility, 79-140.

Parrish, R. V., Kahlbaum, W. M. and Steinmetz, G. G. (1979). Effect of image tilt of a vertical image display on simulated transport touchdown performance. NASA TP-1541, November, 1979.

Reardon, K. A. (1988). The effects of nested texture on a landing-judgment task. In *Proceedings of the Human Factors Society 32nd Annual Meeting*, Santa Monica , CA, pp. 10-14.

Regal, D. M. and Braune, R. J. (1992). Toward a flight deck automation philosophy for the Boeing high speed civil transport. *Proceedings of the 22nd International Conference on Environmental Systems*, Seattle, WA.

Regal, D. M., Rogers, W. H. and Boucek, G. P. (1988). Situational Awareness in the commercial flight deck: Definition, measurement, and enhancement. SAE Technical Paper Series 881508. Presented at Aerospace Technology Conference and Exposition, Anaheim, CA.

Regal, D. M. and Whittington, D. H. (1993). Synthetic vision in commercial aviation - display requirements. *Proceedings of the Society for Information Display International Symposium, Volume XXIV*, 441-444.

Reising, J., Barthelemy, K., and Hartstock, D. (1989). Pathway-in-the-sky evaluation. *Proceedings of the 5th Symposium on Aviation Psychology*. Santa Monica, CA.

Riley, V. (1989). A general model of mixed-initiative human-machine systems. *Proceedings of the Human Factors Society 33rd Annual Meeting*.

Silsby, N. S. (1955). Statistical measurements of contact conditions of 478 transport-airplane landings during routine daytime operations. NACA Report 1214.

Silsby, N. S. and Livingston, S. P. (1959). Statistical measurements of contact conditions of commercial transports landing on airports at an altitude of 5,300 feet and at sea level. NASA TN D-147, November, 1959.

Stickle, J. W. and Silsby, N. S. (1960). An investigation of landing-contact conditions for a large turbojet transport during routine daylight operations. NASA TN D-527, October, 1960.

Warren, R. (1988). Visual perception in high-speed low-altitude flight. *Aviation, Space, and Environmental Medicine*, p A116-A124.

Warren, R. and McMillan, G. R. (1984). Altitude control using action demanding interactive displays: Toward an active psychophysics. Proceedings of the 1984 Image Conference III, 38-51

Way, T. (1989). 3D in pictorial formats for aircraft cockpits. Proceedings of the 33rd Annual Meeting of the Human Factors Society, Santa Monica, CA.

Way, T. C., Hobbs, R. E., Qualy-White J., and Gilmour, J. D.(1989). 3D imagery cockpit display development. Boeing Final Report WRDC-TR-89-April.

Wickens, C. D., Todd, S. and Seidler K. (1989). Three-dimensional displays: perception, implementation, and application. Contract report under contract DLA900-88-D-0393 from AAMR/HEL/SCERIAC, Wright-Patterson Air Force Base, Dayton, OH.

Wolpert, L. (1988). The active control of altitude over differing texture. In *Proceedings of the Human Factors Society - 32nd Annual Meeting*, Santa Monica, CA, pp.15-19.

Wolpert, L., Owen, D. H., and Warren, R. (1983). The isolation of optical information and its metrics for the detection of decent. In D. H. Owen (Ed.), *Optical flow and texture variables useful in simulating self motion (II)*. Final Technical Report for Grant No. AFOSR-81-0078).

Yeh, Y. Y. and Silverstein, L. D.(1992). Spatial judgments with monoscopic and stereoscopic presentation of perspective displays. *Human Factors*, 34(5), 583-600.

Zenyuh, J. P., Reising, J. M., Walchli, S., and Biers, D. (1988). A comparison of a stereographic 3-D display versus a 2-D display using an advanced air-to-air format. Proceedings of the 32nd Annual Meeting of the Human Factors Society. Santa Monica, CA.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)

2. Report Date
August 1994

3. REPORT TYPE AND DATES COVERED
Contractor Report

4. TITLE AND SUBTITLE

Synthetic Vision Display Evaluation Studies

5. FUNDING NUMBERS
NAS1-19360
537-08-20-01

6. AUTHOR(S)
David M. Regal and David H. Whittington

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Boeing Commercial Airplane Group
Flight Deck Research
Seattle, Washington 98124-2207

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(2) AND ADDRESS(ES)
National Aeronautics and Space Administration
Langley Research Center
Hampton, Virginia 23681-0001

10. SPONSORING / MONITORING AGENCY REPORT NUMBER
NASA CR-194963

11. SUPPLEMENTARY NOTES
Langley Technical Monitor: Randall L. Harris, Sr.
Final Report - Task 16

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Unclassified - Unlimited
Subject Category 06

12b DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

The goal of this research was to help us understand the display requirements for a synthetic vision system for the High Speed Civil Transport (HSCT). Four experiments were conducted to examine the effects of different levels of perceptual cue complexity in displays used by pilots in a flare and landing task. Increased levels of texture mapping of terrain and runway produced mixed results, including harder but shorter landings and a lower flare initiation altitude. Under higher workload conditions, increased texture resulted in an improvement in performance. An increase in familiar size cues did not result in improved performance. Only a small difference was found between displays using two patterns of high resolution texture mapping. The effects of increased perceptual cue complexity on performance was not as strong as would be predicted from the pilot's subjective reports or from related literature.

A description of the role of a synthetic vision system in the High Speed Civil Transport is provide along with a literature review covering applied research related to perceptual cue usage in aircraft displays.

14. SUBJECT TERMS
Display design, HSCT, flight deck, spatial perception, situational awareness, synthetic vision, computer generated imagery.

15. NUMBER OF PAGES
67

16. PRICE CODE
A04

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified		