Journal of Applied Psychology
1993, Vol. 78, No. 5, 763-773

0021-9010/93/$3.00

Visual–Spatial Abilities of Pilots

Itiel E. Dror, Stephen M. Kosslyn, and Wayne L. Waag

U.S. Air Force pilots and control subjects participated in five experiments, each of which assessed a different type of visual–spatial ability. Although pilots judged metric spatial relations better than did nonpilots, they did not judge categorical spatial relations better than did nonpilots. Pilots mentally rotated objects better than did nonpilots, but pilots did not extrapolate motion, scan images, or extract visual features from objects obscured by visual noise better than did nonpilots. The results imply that efficient use of specific processing subsystems is especially important for, and characteristic of, pilots. The possible neuropsychological bases for the enhanced abilities and their susceptibility to change are discussed.

Different professions require different abilities. This is obvious when one considers what distinguishes accountants from interior decorators, but the observation applies to all specialized professions. The special abilities may develop on the job through training and practice or may be brought into the job. In either case, special abilities enable people to excel in occupations that depend critically on specific mental processes. Pilots, as members of a highly specialized profession, rely heavily on a set of specific cognitive abilities; in particular, pilots must respond quickly and accurately in a wide variety of tasks that depend on high-level visual cognition. Visual–spatial skills are critical for both mission accomplishment and safety. In this article, we present evidence that pilots are particularly adept at using some specific high-level visual–spatial abilities, but not others.

Our goal was to understand and characterize the special visual–spatial abilities of pilots from a cognitive neuroscience perspective, which views the brain as a complex information-processing system that is divided into various subsystems (see Kosslyn & Koenig, 1992, chap. 2). Each processing subsystem can be thought of as a "black box" that accepts a specific kind of input and produces a specific kind of output. Thus, the processing labor required to perform a cognitive function is divided among different component subsystems that interact to perform complex—and yet, flexible—information processing. Knowledge of the neuroanatomy of the brain and of how brain damage selectively affects cognitive functioning is crucial for understanding these different subsystems. In a similar way, computational models provide important hints about the nature of the subsystems that underlie specific cognitive abilities. In this article we present an initial step toward examining the efficacy of specific processing subsystems that underlie the high-level visual–spatial abilities of pilots.

High-level vision relies on the use of stored knowledge, whereas low-level vision is driven purely by the stimulus input. Research on low-level vision has laid solid foundations for many tests that allow one to measure such abilities. In contrast, few applications have been developed from research on high-level vision. The research reported here is a step toward remedying this deficiency. We assessed specific abilities that should be drawn upon during piloting, according to the theory of high-level visual processing developed by Kosslyn, Flynn, Amsterdam, and Wang (1990). Specifically, the tasks we administered tap into processes that are used to rotate objects in mental images, to extrapolate motion, to scan imaged objects, to encode spatial relations, and to extract visual features in the presence of visual noise. We explain the role of each process in piloting aircraft when we introduce the relevant experiment.

We assessed the efficacy of each type of process by making use of a variant of "additive factors" methodology (Sternberg, 1969). In each experiment, we manipulated the difficulty of the judgment to force a specific process to work harder in one condition than in another. Thus, increases in response times and error rates with increased difficulty reflect the increased amount of processing needed to perform the more difficult trials. This logic has been used by researchers who have examined changes over age in the ability to rotate objects in visual mental images (e.g., Berg, Hertzog, & Hunt, 1982; Cerella, Poon, & Fozard, 1981; Gaylord & Marsh, 1975; Jacewicz & Hartley, 1979; Puglisi & Morrell, 1986; Sharps & Gollin, 1987). By comparing the increases in time that subjects need to rotate objects greater amounts (the slope), researchers can examine the rotation process itself—independent of the processes involved in encoding the stimulus and in generating the response itself (which are reflected by the intercept of the function).

Itiel E. Dror and Stephen M. Kosslyn, Department of Psychology, Harvard University; Wayne L. Waag, Aircrew Training Research Division, Armstrong Laboratory, Williams Air Force Base, Arizona.

This research was supported by Research Initiation Program Research Grant 92-76, awarded to Itiel E. Dror by the U.S. Air Force Office of Scientific Research (AFOSR) through Contract F49620-90-C-0076 to Research and Development Laboratories; and by AFOSR Grant 91-0100, awarded to Stephen M. Kosslyn.

We thank Gretchen Krueger and the personnel on Williams Air Force Base for their help and time. We also thank James Intriligator, Christopher Chabris, Nachshon Meirn, and Don Lyon for their comments on earlier versions of this article. In addition, we thank two anonymous reviewers for comments and suggestions on an earlier version of this article.

Correspondence concerning this article should be addressed to Itiel E. Dror, Department of Psychology, Harvard University, William James Hall, 33 Kirkland Street, Cambridge, Massachusetts 02138.

This method allowed us to determine whether pilots can perform particular kinds of processing better than can nonpilots. By comparing the measures obtained in a number of tasks, we could begin to characterize pilots' abilities and specify which subsystems are particularly efficient in pilots in comparison with nonpilots. In contrast to our difference-score measures, measures of any differences in overall performance could be caused by other, unrelated factors, such as the speed of pressing the response keys.

General Method

The detailed method and procedure are presented with each individual task. To eliminate redundancy, we present here the common aspects of the procedures and designs.

General Procedure

The subjects were always tested individually, and in all tasks we began by asking them to read instructions on a computer screen. The subjects were then asked to paraphrase the instructions so that any misconceptions that they had could be corrected. Next, subjects were given a set of practice trials. During the practice trials, the computer provided feedback by beeping when a subject made an incorrect response, and the subjects were encouraged to ask questions. During the test trials, no feedback was provided and no talking was allowed.

In all tasks, responses to half of the trials should have been yes and responses to the other half should have been no. The trials were always ordered randomly, except for the constraint that the same level of difficulty or response could not appear more than 3 times in succession. All tasks required the subjects to respond by pressing keys marked *yes* (the *b* key) and *no* (the *n* key) on the computer's keyboard. Before performing the tasks, the experimenter explained the yes and no keys to subjects, and the subjects practiced using them. The practice sessions included 32 trials. In each practice trial the word *yes* or *no* was presented on the computer screen, and the subjects were required to press the corresponding key on the computer keyboard as quickly and as accurately as possible. If the subject pressed the wrong key, the computer beeped. The subjects used two fingers of their dominant hand to press the keys. Tasks were administered on a Macintosh Plus computer with a built-in 9-in. (22.9-cm) monitor. A Polaroid CP-50 filter was placed over the screen to prevent glare. All subjects sat so that their heads were approximately 45 cm from the computer screen. The subjects were asked to respond as quickly as possible while remaining as accurate as possible. The tasks were administered using the MacLab software program (Costin, 1988), which is designed to enable the use of a Macintosh computer for testing subjects on cognitive tasks.

All subjects were tested during one session that lasted, on average, 1 hr and 45 min. Subjects were allowed to take a 5-min break after 1 hr, but no subject took that option. All subjects were tested on the same tasks in the same order: extrapolating motion, extracting visual features from objects in the presence of visual noise, scanning images, judging spatial relations, and rotating images. This order was designed to minimize possible transfer or interference between tasks. Thus, tasks that are potentially related were presented far apart during testing. We report the results in an order that makes sense conceptually.

Subjects

Sixteen pilots and 16 control subjects were tested in all five experiments. The pilots and 4 of the control subjects were recruited and tested at the Aircrew Training Research Division of the Armstrong Laboratory at Williams Air Force Base, Arizona. All 16 pilots were male; their mean age was 30 years (with a range from 23 years to 46 years). Fourteen pilots were right-handed and 2 were left-handed, and all of the pilots had at least a college education. The pilots had an average of 1,773 flight hr (with a range from 218 hr to 4,170 hr; only 3 pilots had less than 1,000 flight hr). The other 12 control subjects were members of the Harvard University community and were tested at the university. The sex, age, handedness, and education of the complete control group were matched to the pilots'; all members of the control group were male, and their mean age was 29 years (with a range from 21 years to 44 years). Fourteen of the control subjects were right-handed and 2 were left-handed, and each control subject had at least a college education.

Experiment 1: Mental Rotation

Visual mental images can be transformed in many ways. One transformation that has received much attention is image rotation, which requires subjects to imagine an object rotating. Previous research has shown that subjects require more time to visualize an object as it rotates greater amounts (e.g., Shepard & Cooper, 1982). Indeed, response times typically increase linearly with increased amounts of mental rotation (e.g., see Shepard & Cooper, 1982). The time needed to alter the orientation of an object is reflected by the slope of the increase in response times with greater amounts of rotation, and many researchers have compared the slopes of subjects' response times in different subject populations (e.g., see Dror & Kosslyn, in press).

In aviation, pilots frequently find themselves in orientations that require them to use imagery to rotate objects back to their upright orientations. For instance, when a pilot is flying straight and level and then rolls the aircraft to 90° and pulls hard to initiate a turn, the pilot must "rotate" his or her internal view to accurately assess the relative position of the aircraft.

Our image-rotation task required that subjects determine whether two sequentially presented shapes were identical or mirror reversed, regardless of each shape's orientation. The shapes were rotated only in the picture plane; the first was upright, and the second was at one of four angular disparities relative to the first. This task is a modified version of the task devised by Cooper (1975).

Method

Materials. The stimuli were composed of two or three connected bars, each of which was made by juxtaposing 0.6 cm × 0.6 cm squares. The shapes never exceeded 2.6 cm × 3.2 cm, which corresponded to approximately 2.9° × 4.1° of visual angle (an example is presented in Figure 1). The first shape was always presented upright, and its top was colored black. The same segment was black in the second shape, which helped subjects to locate the corresponding parts of the two objects. Sixty-four trials were constructed: On 16 trials the stimuli were rotated 0° from the orientation of the first stimulus (i.e., were upright), on 16 trials the stimuli were rotated 90° clockwise, on 16 trials the stimuli were rotated 135° clockwise, and on the remaining 16 trials the stimuli were rotated 180°. Half of the stimuli in each orientation condition were identical to the first member of the pair and half were mirror reversed. For every 8 trials within each cell, 4 included shapes composed of two bars and 4 included shapes composed of three bars. Two additional shapes were used to prepare 16 practice stimuli; the practice trials included 4 trials in each orientation condition, half of which contained mirror-reversed patterns.

Procedure. A trial began when an exclamation mark appeared on the computer screen for 500 ms, and continued with a blank screen for an additional 500 ms. Following this, an upright shape was presented. The

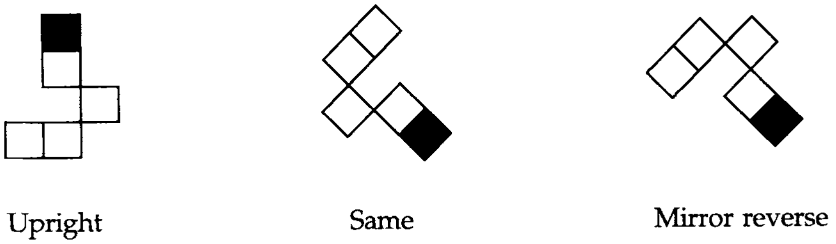

Figure 1. Example of the stimuli used in the image-rotation task.

subject studied this shape, pressed the space bar after memorizing it, and was then ready to view the second shape. The second shape appeared 500 ms thereafter, and the subject was to mentally reorient the shape and then decide whether it was identical to the first member of the pair or was a mirror-reversed version of it. If the shape was identical, the subject was to press the yes key; if it was mirror reversed, the subject was to press the no key. Immediately after the subject responded, another exclamation mark appeared on the screen, and a new trial began.

Results

The data were analyzed using analyses of variance (ANOVAs). Prior to analysis, response times greater than 2.5 times the mean of the remaining scores in each cell for each subject (i.e., outliers) were replaced by the mean of that cell; 1.8% of the data were thus replaced. We did not include incorrect responses when calculating mean response times. We performed separate analyses for response times and error rates, which included subject group and orientation difference as independent variables.

We were primarily interested in whether there was an interaction between subject group (pilots and nonpilots) and orientation difference. However, before we could interpret such a finding we needed to know whether our task was measuring what we thought it was. Thus, it is of interest that response times and error rates varied with angular rotation; for response times, $F(3, 90) = 30.18, p < .01$ (*M*s = 1,063 ms, 1,317 ms, 1,303 ms, and 1,541 ms for 0°, 90°, 135°, and 180° of rotation, respectively), for error rates, $F(3, 90) = 15.05, p < .01$ (*M*s = 5.2%, 13.4%, 14.2%, and 20.3% errors for 0°, 90°, 135°, and 180° of rotation, respectively). Linear contrasts revealed that response times increased linearly with orientation difference, $F(1, 90) = 79.84, p < .01$, as did error rates, $F(1, 90) = 41.54, p < .01$. The slope of the increase in response time was equal to 1,060.9 + [2.42 (angle of rotation)], which accounted for most of the variance ($R^2 = .90$); and the slope of the increase in errors was equal to 5.27 + [.079 (angle of rotation)], accounting for almost all of the variance ($R^2 = .96$). Because we replicated previous findings, we were confident that our task did, in fact, tap mental-rotation ability.

To compare rotation ability between the two groups, we needed to observe how variations in the amount of processing performed by a subsystem affected response times and error rates for each group. As illustrated in Figure 2, we found an interaction in response times between subject group (pilots vs. nonpilots) and orientation difference, $F(3, 90) = 8.13, p < .01$. A linear-by-linear contrast revealed that the pilots did not require as much extra time as the nonpilots when orientation difference increased, $F(1, 90) = 22.08, p < .01$. Pilots and nonpilots had comparable error rates at the different angular disparities, as reflected by the absence of an interaction between subject group (pilots vs. nonpilots) and orientation differences ($F < 1$). Finally, we found that the pilots were faster overall than the nonpilots, $F(1, 30) = 6.75, p = .01$ (*M*s = 1,121 ms vs. 1,491 ms), but that they were no more accurate than the nonpilots ($F < 1$; *M*s = 12.2% error vs. 14.4% error).

Discussion

We found that pilots can mentally rotate objects faster than can nonpilots. Although the task itself probably recruits many processing subsystems, appreciably fewer of these subsystems are used to rotate the image than are used to encode the shapes, evaluate them, and so on. Thus, the difference in rotation slopes allowed us to narrow down the pilots' advantage to those processes that change the represented orientation itself. Indeed, Kosslyn (1987) claimed that only three subsystems are used to rotate imaged patterns: one that shifts the represented positions of segments, one that monitors the spatial relations among segments, and one that looks up stored information (which is used to direct the shift subsystem to realign the segments properly). The results of additional tasks, described below, indicated that pilots are not better than nonpilots at encoding the kinds of spatial relations that would be used to note the relative positions of the segments (i.e., categorical relations, indicating that the segments are preserving right angles). However, it is possible that pilots are better at accessing spatial information in their memory or at shifting the locations of representations.

We also found that pilots were generally faster overall than nonpilots, but, as explained in the General Method section, this could reflect pilots' superior ability in encoding the stimuli, making the responses, or some other factor.

Experiment 2: Motion Extrapolation

Another basic visual-spatial ability allows one to track a moving object and to anticipate its position when it is no longer visible. Because movement—across all orientations and attitudes—is an inherent characteristic of flight, such an ability is vital for piloting. Pilots need to extrapolate motion and spatial positions of objects, especially when they attend to cockpit displays (and are thus disconnected from direct visual contact with their surroundings) and when they need to keep track of many objects that cannot all be attended to at once. We administered

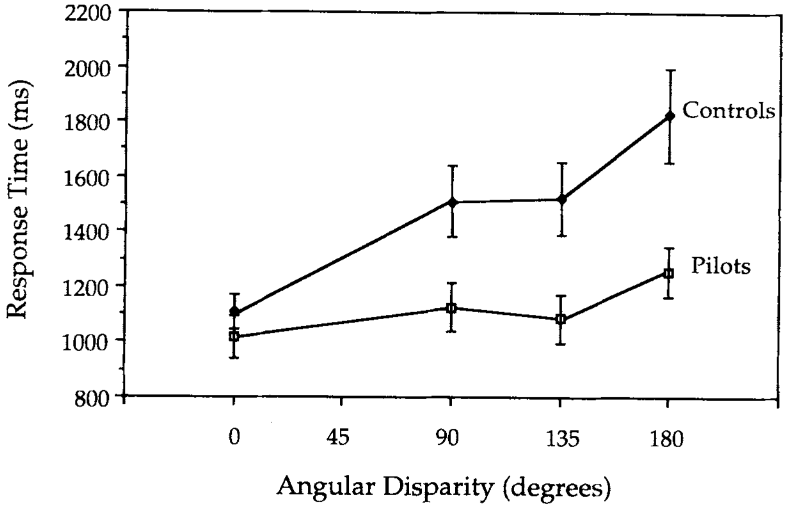

Figure 2. Results from the image-rotation task.

a task that required subjects to track a ball on the computer screen and then extrapolate its future position. The ball traveled at a constant speed for about two full circles and then disappeared from the screen. After a delay, an X was briefly presented, and the subjects decided whether it would have covered the ball if the ball had continued on its path at the same speed.

Method

Materials. The stimulus was a round disk (the ball), 0.5 cm in radius—corresponding to 0.6° of visual angle—that moved in a circular trajectory around the center of the screen. The circular trajectory was divided into 30 positions that were staggered 12° apart. Motion was created by displaying the ball in each successive position for 183 ms, with no interstimulus interval. This method produced reasonably smooth motion on the screen. The ball required 5.49 s to complete one full circle. The radius of the trajectory was 5.40 cm (measured from the center of the circle to the center of the moving ball).

On half of the trials the ball disappeared after completing 12° past two full circles, and on the other half of the trials it disappeared 12° before completing the two full circles. After a delay of 549 ms, 915 ms, or 1,281 ms, an X appeared on the screen for 183 ms, and then the screen went blank. The X was either where the ball would have been, had it continued to move along its trajectory at a constant speed, or 12° or 24° in front of or behind that position. Forty-eight trials were prepared, half of which had Xs that would have fallen "on" the ball and half of which had Xs that would have fallen "off" it; in half of the "off" trials the Xs were ahead of the proper location, and in half they were behind it. In this task, we assumed that the subsystems used to extrapolate motion had to work harder when there was a longer delay and so the trajectory had to be projected a greater distance.

In addition, we could assess the precision of these processes by considering those off trials that had a small disparity between the location of the X and the actual position of the ball (difficult trials) in comparison with those that had a great disparity (easy trials). Half of the off trials were easy and half were difficult. At each level of difficulty, the X appeared before the projected position of the ball on six trials and appeared after that position on six trials. The three delays appeared equally often for each type of trial. Eight practice trials were prepared, half on and half off; half of each of these types of trials had the shortest delay, and half had the longest delay.

Procedure. A trial began with an exclamation mark. When the subject was ready, he pressed the space bar, and then a moving ball appeared. The subject was told to track the motion of the ball and to keep tracking it in his imagination after it disappeared. When the X appeared, the subject was to respond by choosing yes if the X was where the ball should have been and no otherwise. Immediately after the response, another exclamation mark appeared, and a new trial began.

Results

The data were analyzed as in Experiment 1; 1.2% of the data were considered to be outliers. We began by considering only the trials in which the X was presented off the position where the ball should have been (off trials). As before, we first checked the results to ensure that we succeeded in varying the difficulty of the trials. In this case, the difficulty of the trials was manipulated by the distance between the X and the position where the ball should have been. The subjects did require more time for the more difficult trials, $F(1, 28) = 7.60, p = .01$ (Ms = 686 ms and 778 ms for easy and difficult trials, respectively), and they made more errors for these trials, $F(1, 28) = 38.71, p < .01$ (Ms = 29% and 49% errors for easy and difficult trials, respectively); the effect-size measure (r) was .46 for response time and .76 for error rate. However, there were no differences between the pilots and the nonpilots in either overall response times or error rates (701 ms vs. 768 ms and 38% vs. 40% errors for pilots vs. nonpilots, respectively; $F < 1$, in both cases). As illustrated in Figure 3 (top), there was no interaction between subject group and difficulty ($F < 1$ for both response times and error rates).

We conducted a second analysis to consider the trials on which the X was presented where the ball should have been ("on" trials), assuming that longer delays were more taxing. However, the subjects did not require more time to evaluate Xs presented after a longer delay ($F < 1$; Ms = 714 ms, 663 ms, and 695 ms for the 549, 915, and 1,281 ms delays, respectively); in contrast, the subjects did make more errors for the longer delays, $F(2, 54) = 5.16, p < .01$ (Ms = 27.9%, 27.7%, and 42.8% errors for 549 ms, 915 ms and 1,281 ms delays, respectively). The pi-

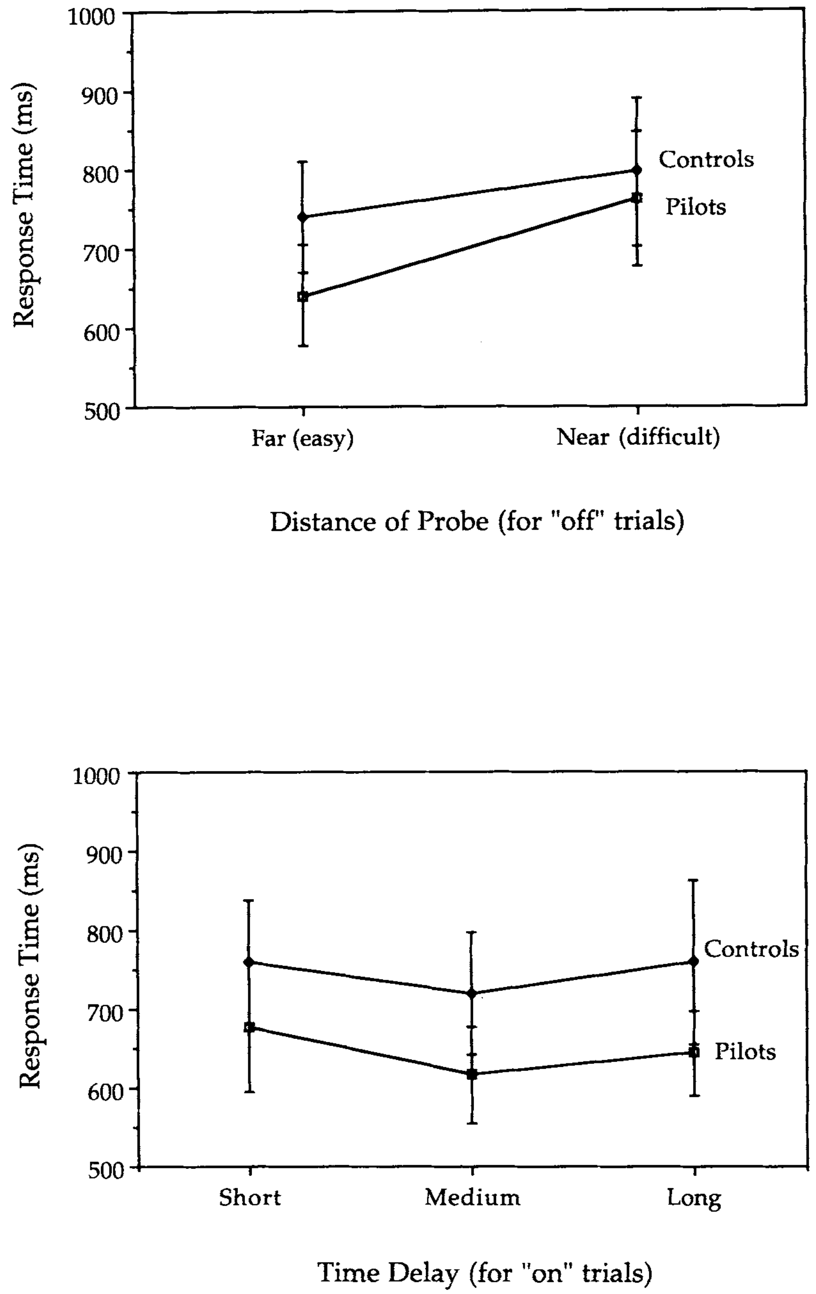

Figure 3. Results from the motion-extrapolation task. Distances between the X probe and position (top) and between delay times for the X probe (bottom).

lots were not significantly faster than the nonpilots ($F < 1$; 646 ms vs. 746 ms) and were not more accurate, $F(1, 27) = 3.00$, $p > .05$ (Ms = 28.6% error vs. 38.0% error). As illustrated in Figure 3 (bottom), delay time had the same effect on the two groups; there was no hint of an interaction between subject group and delay ($F < 1$ for response times and error rates).

Discussion

The motion-extrapolation task did not reveal any difference between pilots and nonpilots. Although increasing the distance of the X and increasing the delay both affected performance, they did so in the same way for subjects in both groups. However, this was a difficult task, as witnessed by the high error rates even for easy trials. Performance may reflect a floor effect, and thus it is possible that we would have found a difference if we had not taxed processing so much. (Even if the pilots are better, they too will perform poorly if the task is difficult enough.) The varying levels of difficulty between the trials might not have been large enough to allow us to detect differences between the groups (there was only a 732-ms difference between the shortest and longest delay and were only 12° or 24° of disparity between the X and the position of the ball). As always, one must be cautious before affirming the null hypothesis.

Experiment 3: Scanning Images

In the previous experiment, subjects were asked to visualize a ball moving along its path after it was no longer present on the screen. This extrapolation process may be distinct from the process of scanning over an imaged object. Scanning an image involves systematically shifting one's attention over an object or scene. This sort of process may typically be used when one visualizes possible scenarios and sees what would happen, inspecting the imaged patterns. In aviation, pilots are often required to visualize such scenarios; for instance, during a landing they have to fly the aircraft so that it will land on the runway in the right angle and orientation. In addition, "situation awareness" often consists of visualizing one's surroundings, including the locations of other aircraft. Both activities would appear to involve scanning over imaged scenes. Thus, we decided to assess image-scanning ability in the two groups.

Our image-scanning task was a variant of one developed by Finke and Pinker (1982) and used by Kosslyn, Margolis, Barrett, Goldknopf, and Daly (1990) in their study of imagery abilities in children and adults. The hallmark of such mental-image scanning is that subjects require more time to shift their attention greater distances over an imaged object or scene, and subjects who performed these tasks have displayed this pattern of performance (replicating the original finding of Kosslyn, 1973). In our version of the task, the subjects saw a square-ring shape that was composed of white and black squares. An arrow appeared briefly in the center of the ring, and then the entire display disappeared, and subjects decided whether the arrow had been pointing to a black square. Because the arrow was presented very briefly, subjects could not make the judgment based on the percept, but rather had to use a mental image of the display. We varied the distance from the arrow to the square ring, expecting more scanning when greater distances had to be traversed.

Method

Materials. A ring was constructed from twenty 0.7 cm × 0.7 cm squares, with 6 squares on each side, giving the shape an overall size of 4.2 cm × 4.2 cm, which corresponded to 5.3° × 5.3° of visual angle. As illustrated in Figure 4, on each trial 3 different squares were filled with black, each on a different side of the ring. A 0.4-cm-long arrow, corresponding to 0.5° of visual angle, appeared within the center region of the ring after the subject pressed the space bar. The arrow pointed either north, northeast, east, southeast, south, southwest, west, or northwest. In addition, the tip of the arrow was positioned 2.1 cm from, 1.2 cm from, or adjacent to the target square, which corresponded to 2.7°, 1.5°, and 0.0° of visual angle, respectively. We constructed 60 test trials, 20 for each distance. On half of the trials at each distance, the arrow pointed to a black square, and on half it pointed to a white square. We also prepared 12 additional, practice trials, 4 for each distance; for each distance, on 2 trials arrows pointed to black squares and on 2 trials arrows pointed to white squares.

Procedure. A trial began when an exclamation mark appeared on the screen for 500 ms; this was then replaced by a blank screen for an additional 500 ms. Following this, a square-ring stimulus appeared. The subject was to study the ring until he could remember the locations of the black squares, at which point he pressed the space bar; 250 ms later, an arrow appeared for 50 ms, and then all stimuli were removed from the screen. The subject pressed the yes key if the arrow pointed to one of the black squares and the no key if it pointed to a white square. Immediately after the response, an exclamation mark appeared, and a new trial began.

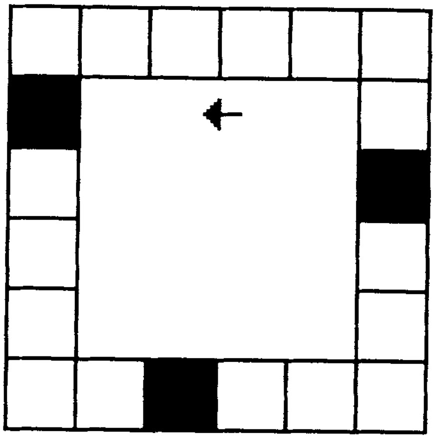

Figure 4. Example of the stimuli used in the image-scanning task.

Results

The data were analyzed as in Experiment 1; 2.1% of the data were considered to be outliers. As before, we first considered whether the results showed that the task was measuring what we assumed it was measuring. Replicating previous results, our results showed that subjects required more time to scan greater distances, $F(2, 60) = 73.57, p < .01$ (Ms = 604 ms, 643 ms, and 706 ms for adjacent to, 1.2 cm from, or 2.1 cm from the target square, respectively), and made more errors with greater distances, $F(2, 60) = 34.90, p < .01$ (Ms = 4.4%, 12.8%, and 16.4% error for the three increasing distances, respectively); the effect-size measure (r) was .74 for response time and .61 for error rate. As illustrated in Figure 5, the subjects in both groups required about the same amount of additional time to scan each additional increment of distance, $F(2, 60) = 1.25, p > .05$, for the interaction in response times. Similarly, subjects made comparable increases in errors for each additional increment of distance ($F < 1$, for the interaction in error rates). We also found that pilots were generally faster than nonpilots, $F(1, 30) = 6.32$, $p = .01$ (Ms = 607 ms vs. 695 ms), but that pilots were no more accurate than nonpilots ($F < 1$; Ms = 11.8% vs. 10.8% errors).

Discussion

Pilots and nonpilots scanned images at comparable rates. This result is similar to our findings in the motion-extrapolation task, in which increasing delays affected performance in the same way for subjects in both groups. Unlike that task, however, this one was not so difficult as to suggest possible floor effects. Thus, we are led to suspect that pilots and nonpilots can mentally scan images equally well.

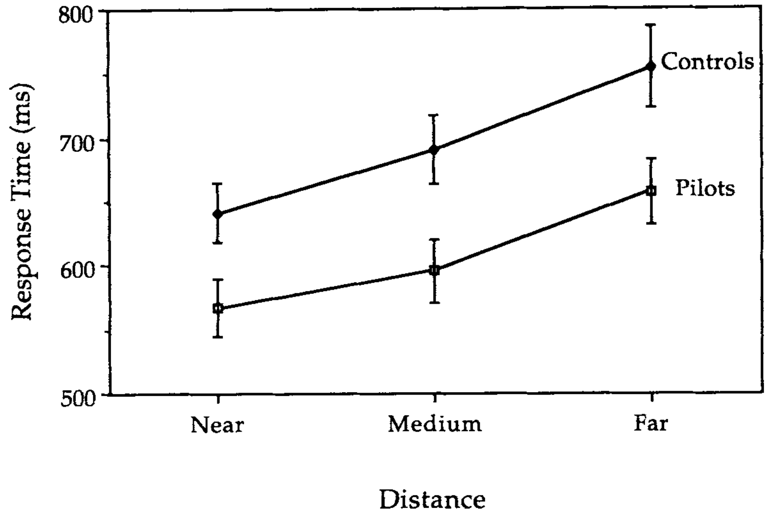

Figure 5. Results from the image-scanning task.

Experiment 4: Spatial Relations Encoding

There are many ways in which spatial relations between objects or parts of objects can be internally represented. Kosslyn et al. (1989) showed that humans can encode spatial relations in at least two different ways. *Categorical spatial relations* collapse over variations in metric distance and position, assigning the relation between two objects to a spatial category. For example, above versus below, on versus off, and left versus right are categorical spatial relations. In contrast, *coordinate* representations specify information in a way that is useful for guiding action (see Kosslyn, 1991; Kosslyn & Koenig, 1992, chap. 3); these representations specify metric distance among objects. Kosslyn et al. (1989), Hellige and Michimata (1989), and others (for a review, see Kosslyn, Chabris, Marsolek, & Koenig, 1992) have shown that the left cerebral hemisphere encodes categorical spatial relations better than does the right hemisphere, whereas the right hemisphere encodes coordinate spatial relations better than does the left. This finding indicates that at least two subsystems must be used to encode spatial relations; if only one had been used, one of the hemispheres would have been better at both types or both hemispheres would have been equally effective.

In aviation there are many tasks that require the rapid encoding of such spatial relations. An example would be an approach to landing, in which the pilot is required to maintain the proper glide slope. At a categorical level, a pilot must know whether he or she is above or below the glide slope. However, although such information may be useful in the initial phases of an approach, by itself it is insufficient for helping a pilot maintain proper glide-slope control. The pilot requires cueing, either in the external visual field or from his or her instruments, that gives information pertaining to the amount of deviation. A pilot must encode the amount of deviation in order to apply the proper amount of control pressure to correct for such deviation. Thus, we wanted to assess how well pilots and nonpilots could encode the two kinds of spatial relations.

We administered two spatial-relations tasks, one that required the subjects to make categorical spatial judgments and one that required them to make metric spatial judgments. We used variants of the tasks developed by Hellige and Michimata (1989). The categorical task required the subjects to decide whether an X was above or below a bar; the distance of the X from the bar was varied to manipulate difficulty. The same stimuli were also used in the metric task, but in that task the subjects were required to decide whether the X was within half an inch (1.27 cm) of the bar; the distance of the X from this criterion distance was varied to manipulate difficulty.

Method

Materials. The stimuli consisted of a horizontal bar and an X. The bar was 0.6 cm wide and 2.4 cm long, corresponding to 0.8° and 3.0° of visual angle, and the X was 0.4 cm × 0.4 cm, corresponding to 0.5° × 0.5° of visual angle. The location of the X was varied to manipulate the difficulty of the trials. For the categorical spatial judgment, the X was placed so that it just touched the bar for the difficult trials and was placed farther than 2 cm from the bar (2.5° of visual angle) for the easy trials. For the metric spatial judgment, the X fell within 2.54 mm (0.3°) of the half-inch border from the bar for the difficult trials and farther than 5.08 mm from the bar (0.6°) for the easy trials. In both tasks, the bar was randomly moved 6.00 mm up and down to ensure that subjects were not memorizing locations on the computer screen; we also randomly moved the X left and right but always kept it within the boundaries of the left and right edges of the bar. For each task, we prepared 64 test trials and 12 practice trials. Half of each type of the trials were easy, and half were difficult. Furthermore, for each level of difficulty, subjects should have responded yes to half of the trials and no to half of the trials. The X appeared equally often above and below the bar in both types of trials.

Procedure. A trial began with an exclamation mark, at which point the subject pressed the space bar; 1 s thereafter, a bar-and-X stimulus appeared. For the categorical task, the subject was asked to decide whether the X was above the bar. For the metric task, he was asked whether the X was within half an inch from the bar. Immediately after the subject responded, another exclamation mark appeared on the screen, and a new trial began.

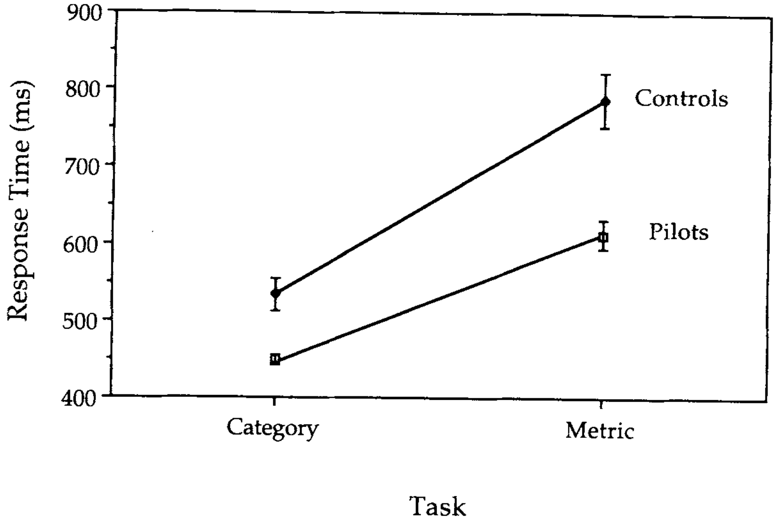

Figure 6. Results from the spatial judgment tasks.

The categorical task was administered before the metric task. Before the metric task, subjects were shown what a half-inch distance looked like on the computer screen and were asked to memorize it. Before the practice trials began, they were shown a picture of the bar with a half-inch border drawn above and below the bar. They then were shown four stimuli that included the half-inch border: two examples in which the X was within a half-inch of the bar and two in which it was not. In each type of example, one had the X above the bar and one had the X below the bar. After the practice trials, subjects were once again shown the picture of the bar with a half-inch border drawn above and below it and were asked to study this interval. Before the actual testing began, subjects were asked whether they had memorized the half-inch distance and whether they wanted to see it again. All subjects memorized the half-inch distance and did not ask to see it again.

Results

The data were analyzed as in Experiment 1; 1.9% of the data were considered to be outliers. In this analysis we included task (categorical vs. metric) as an independent variable in the ANOVA in addition to subject group and difficulty of the trials. We found that subjects needed more time and made more errors with the more difficult trials; for response time, $F(1, 30) = 45.86$, $p < .01$ (*M*s = 567 ms and 624 ms for easy and difficult trials, respectively), and for error rates, $F(1, 30) = 16.51$, $p < .01$ (*M*s = 3.6% and 6.8% errors for easy and difficult trials, respectively). The effect-size measure (r) was .77 for response time and .60 for error rate. We also found an interaction in response time between difficulty and subject group (pilots vs. nonpilots), $F(1, 30) = 4.90$, $p < .05$. In addition, this difference varied for the two tasks, as shown by an interaction between difficulty, subject group, and task (categorical vs. metric) in the response times, $F(1, 30) = 8.58$, $p < .01$, and by a trend toward an interaction in the error rates, $F(1, 30) = 3.40$, $p > .05$. As illustrated in Figure 6, we also found an interaction in response time between subject group (pilots vs. nonpilots) and task (categorical vs. metric), $F(1, 30) = 8.67$, $p < .01$. A linear-by-linear contrast performed separately on the results for each task revealed that the pilots were less affected by difficulty than were the nonpilots in the metric task: for response time, $F(1, 30) = 6.80$, $p = .01$ (*M*s = 581 ms and 644 ms vs. 713 ms and 862 ms for easy and difficult conditions for pilots vs. nonpilots, respectively), and for error rates, $F(1, 30) = 3.63$, $p > .05$, (*M*s = 5.7% and 9.4% errors vs. 3.7% and 12.3% errors for easy and difficult conditions for pilots vs. nonpilots, respectively). But there was no such interaction in the categorical task: for response time, $F(1, 30) = 1.06$, $p > .05$ (*M*s = 443 ms and 455 ms vs. 533 ms and 534 ms for easy and difficult conditions for pilots vs. nonpilots, respectively), and for error rates, $F < 1$ (*M*s = 3.1% and 3.9% errors vs. 1.9% and 1.8% errors for easy and difficult conditions for pilots vs. nonpilots, respectively). Finally, we found that the pilots were faster overall than the nonpilots, $F(1, 30) = 11.36$, $p < .01$ (*M*s = 531 ms and 660 ms), but were no more accurate ($F < 1$; *M*s = 5.5% and 5.0% errors for pilots and nonpilots, respectively).

Discussion

The pilots judged metric spatial relations better than did the nonpilots, but did not judge categorical spatial relations better than did nonpilots. Given the findings reported in the neuropsychological literature and noted earlier, our results suggest that pilots may be better than nonpilots at specific kinds of right-hemisphere processing. The performance on the categorical-judgment task may reflect a ceiling effect, and thus, it is possible that we would have found a difference if we had taxed processing more than we did.

Experiment 5: Recovering Visual Features

One often sees objects when they are partially occluded or obscured by various kinds of visual noise. For example, an object might be partially behind a bush, off in the distance on a foggy day, and so on. The visual system copes with such problems by extracting critical features that characterize the object.

These features distinguish one object from another and thus enable object recognition even when the object is partly obscured or distorted. Biederman (1987), Lowe (1987), and others have proposed that the visual system's remarkably robust ability to recognize objects in different orientations arises because it extracts relatively invariant features, such as parallel edges, points where edges intersect, and symmetries. In aviation, visibility is often limited, and pilots are required to extract the important visual features that they need for flying the aircraft. Thus, we thought it was important to assess this ability in pilots.

We examined the ability to extract visual features in the presence of visual noise by comparing performance in two conditions. In the "easy" condition, the subjects saw a shape composed of curved contours and an X and were asked to decide whether the X was on the shape; in the "difficult" condition, the subjects performed the same task while noise lines were randomly superimposed over the stimulus. The more difficult version of this task has been shown to be especially hard for a brain-damaged patient who had prosopagnosia, an inability to recognize faces (Kosslyn, Hamilton, & Holmes-Bernstein, 1993).

Method

Materials. The shapes were constructed by connecting one, two, or three rounded bars. As illustrated in Figure 7, the resulting stimuli were segmented, bloblike shapes. Each bar had an average width of 0.6 cm (corresponding to 0.8° of visual angle), and the lengths of the bars varied from 0.6 cm to 2.0 cm (corresponding to 0.8° to 2.5° of visual angle). The X was also created out of curved lines and was 0.6 cm × 0.6 cm (corresponding to 0.8° × 0.8° of visual angle). These bars were gray, allowing one to see a black X if it fell on the pattern. For the difficult condition, eight curved lines were superimposed on each stimulus; the lines were placed randomly over the figure. Forty-eight trials were prepared for each condition, 16 trials for each size of stimuli. For each type of stimulus in each condition, half of the Xs fell on the shapes and half fell near the shape but not on it. We prepared 12 additional trials for practice in each condition, 4 with each number of segments; half of each type of stimuli had Xs on the shape.

Procedure. The subjects were tested in two blocks of trials: one containing only stimuli from the easy condition (without added noise) and one containing only stimuli from the difficult condition (with the added noise). The easy stimuli were always presented first. The appropriate practice trials preceded each block. In both conditions, a trial began when an exclamation mark appeared, at which point the subject pressed the space bar. A blank screen was presented for 1 s, and then a stimulus appeared at the center of the screen. The subject decided whether the X fell on or off the stimulus. Immediately after the subject responded, an exclamation mark appeared, and a new trial began.

Figure 7. Example of the stimuli used in the visual-feature-encoding tasks.

Results

The data were analyzed as in Experiment 1; 1.1% of the data were considered to be outliers. We found that our manipulation did indeed vary the difficulty, as witnessed by the fact that subjects required more time to make a decision when visual noise was present, $F(1, 30) = 4.53$, $p < .05$ (Ms = 595 ms and 611 ms for the easy and difficult trials, respectively), and made more errors when visual noise was present (Ms = 3.9% and 5.2% errors for the easy and difficult trials, respectively), $F(1, 30) = 7.39$, $p = .01$. The effect-size measure (r) was .36 for response time and .45 for error rate. In addition, subjects required different amounts of time to evaluate stimuli of different complexity, $F(2, 30) = 5.26$, $p < .01$ (Ms = 595 ms, 603 ms, and 610 ms for one-, two-, and three-bar stimuli, respectively). As illustrated in Figure 8, variations in difficulty had the same effects for both groups of subjects; there was no interaction between subject group (pilots vs. nonpilots) and the presence or absence of visual noise; for response time, $F(1, 30) = 2.44$, $p > .05$, and for error rates, $F(1, 30) = 2.83$, $p > .05$. No other interactions were found. Pilots did, however, respond faster overall than nonpilots, $F(1, 30) = 5.48$, $p < .05$ (Ms = 571 ms vs. 636 ms for pilots vs. nonpilots, respectively), but did not make fewer mistakes ($F < 1$). However, the p values for the interaction effects were approaching a marginal level of significance, and so we decided to calculate the power of the results to ensure that the lack of interactions were not caused by lack of statistical power. The power was approximately .5 for response times and .7 for error rates. Thus, one must be careful about affirming the null hypothesis from the results of this experiment.

Discussion

We did not find that pilots were better than nonpilots at extracting visual features from images through random noise. However, the magnitude of the effect sizes for the main effects and the power estimates was only moderate; thus, we are cautious in interpreting the results because we cannot argue conclusively that such an interaction would not be found with a greater number of subjects.

General Discussion

In most of our experiments, we found that the pilots responded faster overall than the nonpilots. However, this result could indicate that pilots have better sensory or motor abilities and may say nothing about the specific kind of high-level processing examined in a given task. Hence, in each experiment we varied the difficulty in a way that should have selectively taxed a distinct process. We found that pilots have exceptional abilities in at least two respects: They can mentally rotate objects and can judge metric spatial relations exceptionally well. In contrast, we did not find evidence that pilots have unusual abilities to judge categorical spatial relations, extrapolate motion, scan visual mental images, or extract visual features from objects obscured by visual noise. It is always difficult to interpret

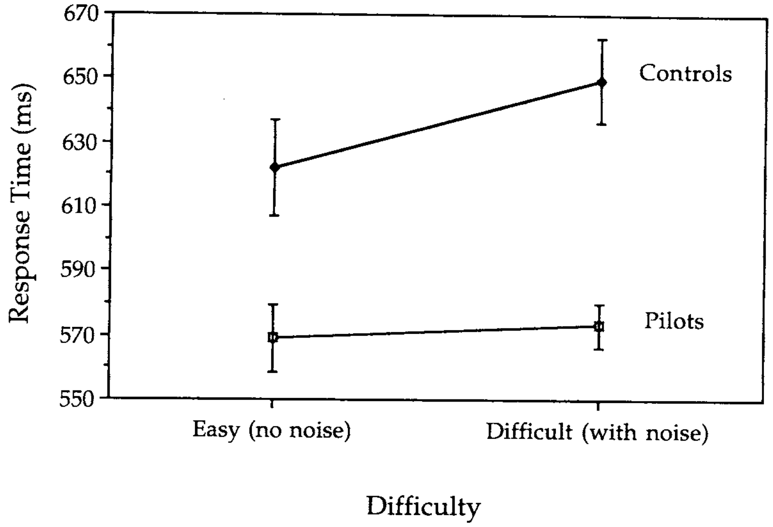

Figure 8. Results from the visual-feature-encoding tasks.

null findings; however, the *p* value for most of these interactions was very high, as was the main-effect size, and when it was not high we validated that the absence of an interaction was not caused by a lack of statistical power. (The one note of caution we sound concerns the final experiment; because we obtained only moderate estimates of power, it is possible that pilots may encode visual features better than do nonpilots but that our sample size was not large enough to detect this difference.) Furthermore, the fact remains that such findings were obtained in the face of positive results on other tasks that had similar sizes of effects. Thus, there is evidence that pilots have selective advantages, not overall superior performance.

If researchers were eventually able to identify certain cognitive processes with certain brain areas, this would open up an entirely new approach to personnel testing and assessment. Thus, the fact that we found evidence that pilots are better than nonpilots only on spatial tasks is intriguing given the division of labor that takes place in high-level visual processing. Ungerleider and Mishkin (1982; for a review, see Kosslyn & Koenig, 1992, chap. 3) documented that information about "what" and "where" is processed by different cortical pathways, with object properties (such as shape) being the province of the inferior temporal lobes, and spatial properties (such as location and size) being the province of the posterior parietal lobes. Further evidence for this division of labor comes from computational models (Rueckl, Cave, & Kosslyn, 1989) that show that the what and where processes are computationally distinct. From yet another perspective, it has been established that human patients suffering from damage to the parietal lobes often have visual-spatial deficits (e.g., Benton, 1985; Kolb & Whishaw, 1990; for a review, see De Renzi, 1982; Kosslyn & Koenig, 1992), whereas humans suffering from damage to the temporal lobes typically have problems with recognizing objects and perceiving their properties (e.g., Benton, 1985; Kolb & Whishaw, 1990; Kosslyn & Koenig, 1992).

The image-rotation task requires one to mentally transform and move an imaged pattern. Such processing relies on a set of complex computations that involve parietal- and frontal-lobe structures (Deutsch, Bourbon, Papanicolaou, & Eisenberg, 1988). The metric spatial-relations encoding task requires one to make precise distance judgments. Such processing relies on accurately making small spatial distinctions, which involves the parietal lobes (particularly right-parietal-lobe structures; cf. Hellige & Michimata, 1989; Kosslyn et al., 1989). For both of these tasks, the performance of pilots in our study was superior to that of nonpilots. We did not find evidence that pilots are especially efficient in the sort of object-recognition processes that are localized in temporal lobe structures (Desimone, Albright, Gross, & Bruce, 1984; Gross, Desimone, Albright, & Schwartz, 1984). Nor did we find evidence that pilots excel at extrapolating motion or image scanning, both of which may involve the middle temporal area (e.g., see Allman, Miezin, & McGuinness, 1985).

The present research dovetails nicely with research on the effect that aging has on mental imagery (Dror & Kosslyn, in press). Dror & Kosslyn found that the ability to rotate images declines with age, whereas the ability to scan images does not. This dissociation is also evident in our present research with pilots. The ability to rotate images was superior in pilots in relation to nonpilots and has previously been shown to be inferior in elderly people in relation to young people, whereas the ability to scan images was no different between nonpilots and pilots in this study or between young people and elderly subjects in Dror and Kosslyn's study. Thus, one is led to suspect that some processes, such as image rotation, are more plastic and thus susceptible to change, whereas other processes, such as image scanning, are less plastic. One possible reason for such differences is that some processes rely more strongly on more primitive, hardwired structures than do others.

It is clear that our research is only the first step toward discovering the ways in which pilots distinctively process information and determining which of these skills are susceptible to change, as well as how they can be changed. This is all a prelude to using measures of the abilities that are not susceptible to change as criteria for screening and selecting pilots and then focusing training on developing those skills that are susceptible to change. Nevertheless, applying studies from a broad range of domains to understand the cognitive components that are important for piloting and the factors that influence them is a fertile and promising approach. Of course, this approach can be applied to a wide range of specialized professions.

References

Allman, J., Miezin, F., & McGuinness, E. (1985). Direction and velocity specific responses from beyond the classical receptive fields in the middle temporal visual area (MT). *Perception, 14,* 105–126.

Benton, A. (1985). Visuoperceptual, visuospatial, and visuoconstructive disorders. In K. M. Heilman & E. Valenstein (Eds.), *Clinical neuropsychology* (pp. 151–186). New York: Oxford University Press.

Berg, C., Hertzog, C., & Hunt, E. (1982). Age differences in the speed of mental rotation. *Developmental Psychology, 18,* 95–107.

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review, 94,* 115–147.

Cerella, J., Poon, L. W., & Fozard, J. L. (1981). Mental rotation and age reconsidered. *Journal of Gerontology, 36,* 620–624.

Cooper, L. A. (1975). Mental rotation of random two-dimensional shapes. *Cognitive Psychology, 7,* 20–43.

Costin, D. (1988). MacLab: A Macintosh system for psychology labs. *Behavior Research, Methods, Instruments, and Computers, 20,* 197–200.

De Renzi, E. (1982). *Disorders of space exploration and cognition.* New York: Wiley.

Desimone, R., Albright, T. D., Gross, C. G., & Bruce, C. J. (1984). Stimulus selective properties in inferior temporal neurons in the macaque. *Journal of Neuroscience, 4,* 2051–2062.

Deutsch, G., Bourbon, W. T., Papanicolaou, A. C., & Eisenberg, H. M. (1988). Visuospatial tasks compared via activation of regional cerebral blood flow. *Neuropsychologia, 4,* 2051–2062.

Dror, I. E., & Kosslyn, S. M. (in press). Mental imagery and aging. *Psychology and Aging.*

Finke, R. A., & Pinker, S. (1982). Spontaneous imagery scanning in mental extrapolation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 8,* 142–147.

Gaylord, S. A., & Marsh, G. R. (1975). Age differences in the speed of a spatial cognitive process. *Journal of Gerontology, 30,* 674–678.

Gross, C. G., Desimone, R., Albright, T. D., & Schwartz, E. L. (1984). Inferior cortex as a visual integration area. In F. Reinoso-Suarez & C. Ajmone-Marsan (Eds.), *Cortical integration* (pp. 291–315). New York: Raven Press.

Hellige, J. B., & Michimata, C. (1989). Categorization versus distance: Hemispheric differences for processing spatial information. *Memory & Cognition, 17,* 770–776.

Jacewicz, M. M., & Hartley, A. A. (1979). Rotation of mental images by young and old college students: The effects of familiarity. *Journal of Gerontology, 34,* 396–403.

Kolb, B., & Whishaw, I. Q. (1990). *Fundamentals of human neuropsychology.* New York: Freeman.

Kosslyn, S. M. (1973). Scanning visual images: Some structural implications. *Perception and Psychophysics, 14,* 90–94.

Kosslyn, S. M. (1987). Seeing and imagining in the cerebral hemispheres: A computational approach. *Psychological Review, 94,* 148–175.

Kosslyn, S. M. (1991). A cognitive neuroscience of visual cognition: Further development. In R. Logie & M. Denis (Eds.), *Mental images in human cognition* (pp. 351–381). Amsterdam: North-Holland.

Kosslyn, S. M., Chabris, C. F., Marsolek, C. M., & Koenig, O. (1992). Categorical versus coordinate spatial representations: Computational analyses and computer simulations. *Journal of Experimental Psychology: Human Perception and Performance, 18,* 562–577.

Kosslyn, S. M., Flynn, R. A., Amsterdam, J. B., & Wang, G. (1990). Components of high-level vision: A cognitive neuroscience analysis and accounts of neurological syndromes. *Cognition, 34,* 203–277.

Kosslyn, S. M., Hamilton, S. E., & Holmes-Bernstein, J. (1993). The perception of curvature can be selectively disrupted in prosopagnosia. Unpublished manuscript, Harvard University, Department of Psychology.

Kosslyn, S. M., & Koenig, O. (1992). *Wet mind: The new cognitive neuroscience.* New York: Free Press.

Kosslyn, S. M., Koenig, O., Barrett, A., Cave, D. B., Tang, J., & Gabrieli, J. D. E. (1989). Evidence for two types of spatial representations: Hemispheric specialization for categorical and coordinate relations. *Journal of Experimental Psychology: Human Perception and Performance, 15,* 723–735.

Kosslyn, S. M., Margolis, J. A., Barrett, A. M., Goldknopf, E. J., & Daly, P. F. (1990). Age differences in imagery abilities. *Child Development, 61,* 995–1010.

Lowe, D. G. (1987). The viewpoint consistency constraint. *International Journal of Computer Vision, 1,* 57–72.

Puglisi, T. J., & Morrell, R. W. (1986). Age-related slowing in mental rotation of three-dimensional objects. *Experimental Aging Research, 12,* 217–220.

Rueckl, J. G., Cave, K. R., & Kosslyn, S. M. (1989). Why are "what" and "where" processed by separate cortical visual systems? A computational investigation. *Journal of Cognitive Neuroscience, 2,* 171–186.

Sharps, M. J., & Gollin, E. S. (1987). Speed and accuracy of mental image rotation in young and elderly adults. *Journal of Gerontology, 42,* 342–343.

Shepard, R. N., & Cooper, L. A. (1982). *Mental images and their transformations.* Cambridge, MA: MIT Press.

Sternberg, S. (1969). The discovery of processing stages: Extensions of Donder's method. *Acta Psychologica, 30,* 276–315.

Ungerleider, L. G., & Mishkin, M. (1982). Two cortical visual systems. In D. J. Ingle, M. A. Goodale, & R. J. W. Mansfield (Eds.), *The analysis of visual behavior* (pp. 549–586). Cambridge, MA: MIT Press.

Received March 13, 1992
Revision received January 27, 1993
Accepted January 28, 1993 ■