Journal of Experimental Psychology: Applied
2000, Vol. 6, No. 2, 148–157

In the public domain
DOI: 10.1037//1076-898X.6.2.148

Object Identification as a Function of Discriminability and Learning Presentations: The Effect of Stimulus Similarity and Canonical Frame Alignment on Aircraft Identification

Alan R. S. Ashworth III
United States Air Force Research Laboratory
and University of Texas at San Antonio

Itiel E. Dror
University of Southampton

Aircraft that were relatively similar (homogeneous) and relatively dissimilar (heterogeneous) in appearance were studied at orientations either consistent (canonical) or inconsistent (noncanonical) with the environmental frame of reference. At test, participants' identification performance was measured with stimuli rotated to novel orientations within the picture plane. During learning and testing, identification of heterogeneous aircraft was better than that of homogeneous aircraft. At test, only identification of homogenous aircraft revealed a strong linear degradation of performance as angular disparity between the novel test orientations and the original learning orientations increased. During learning and testing, identification was better for aircraft studied at canonical orientations than for those studied at noncanonical orientations. The results are discussed in terms of object identification, aircraft recognition training, categorization, mental representations, and visual mental rotation.

In this study, we examined two factors that bear on object identification theory and its application. First, we considered the overall visual similarity of a set of objects. Second, we considered the presentations of the objects during learning, specifically, the alignment of their internal reference frame with that of the environment. The general context in which we studied these factors was that of shape constancy across orientation manipulations. We measured identification performance as a function of orientation disparity between aircraft images presented during learning and, later, during testing.

Different object identification theories attribute distinct roles to various cognitive factors. Two general approaches govern object identification theories: viewpoint-dependent theories (e.g., Tarr, 1995; Tarr & Bülthoff, 1995; Tarr & Pinker, 1989) and viewpoint-invariant theories (e.g., Biederman, 1987; Biederman & Gerhardstein, 1993). In the former, the object representations maintain the specific viewer-centered properties of the images presented during learning. In the latter, the object representations are object-centered abstractions derived from the images presented during learning. This difference in representational format entails distinct patterns in how learning presentations transfer to novel orientations. Viewpoint-dependent theories entail that the novel orientations used during testing must be normalized to match those acquired during learning. Thus, response time is predicted to be a linear function of the angular disparity between the orientations presented during learning and those used at test. In contrast, viewpoint-invariant theories entail matching input to abstract representations that are not dependent on orientation. Thus, orientation disparities between learning and test should have no effect on response time.

Recent research suggests that both of these identification mechanisms may be used, and that the type of identification task determines which one is more suitable (Hamm & McMullen, 1998). To be specific, identification performance depends on whether participants are required to make superordinate judgments (such as "vehicle" vs. "animal") or subordinate judgments (such as "bee" vs. "fly"). Hamm and McMullen (1998) argued (a) that superordinate category identification judgments are based on a viewpoint-invariant mechanism, as identification performance is not orientation dependent, but (b) that subordinate-level identification requires normalization, as response time increases linearly with departure from the canonical view, and thus this type of judgment is based on a viewpoint-dependent mechanism (for more details, see Hamm & McMullen, 1998).

However, in a recent article, Murray (1998) raised the possibility that the above findings may reflect visual similarity differences

---

Alan R. S. Ashworth III, United States Air Force (USAF) Research Laboratory, Brooks Air Force Base (AFB), San Antonio, Texas, and Department of Psychology, University of Texas at San Antonio; Itiel E. Dror, Department of Psychology, University of Southampton, Highfield, Southampton, England.

Both authors contributed equally to this article—the order of authorship is alphabetical. The research reported here was supported by the USAF Research Laboratory, Brooks AFB, and by the Air Force Office of Scientific Research.

We want to thank Shirley Snooks and Chris Schreiner for their contributions and J. Wesley Regian and John Tangney for his support. We would also like to thank Rob Goldstone, Gregg DiGirolamo, Shimon Edelman, and Larry Parsons for comments on earlier versions of this article.

Correspondence concerning this article should be addressed to either Alan R. S. Ashworth III, USAF Research Laboratory, 2509 Kennedy Circle Building 125, Brooks AFB, Texas 78235, or to Itiel E. Dror, Department of Psychology, University of Southampton, Highfield, Southampton SO17 1BJ, England. Electronic mail may be sent to alan.ashworth@brooks.af.mil or to dror@coglab.psy.soton.ac.uk (http://www.cogsci.soton.ac.uk/~dror/).

between the two types of identification judgments rather than the type of judgments themselves. Subordinate judgments involve discriminations between objects that share many visual features, but superordinate judgments involve discriminations between objects that are relatively dissimilar. Murray suggested that the confound between visual similarity and type of identification judgment (superordinate vs. subordinate) may open the findings of Hamm and McMullen (1998) to different interpretations. Hence, visual similarity is offered as a possible alternative account for what determines whether viewpoint-invariant or viewpoint-dependent mechanisms are used (see Dror, Ashworth, & Stevenage, 1999).

The research reported here was aimed at further investigating the factors that influence which type of object identification mechanism governs the identification process. First, we used an experimental design that is applicable to real-world situations. The previous studies used a name verification task whereby participants had to judge whether a name that initially appeared for 1500 ms matched an object line drawing that followed (essentially, a same–different judgment). Rarely does identification in the real world follow such a sequence of events. Furthermore, the presentation of the name prior to the object may bias the identification process because it may cause the participants to activate and perhaps even generate the typical canonical representation for that object. In our experiment, we not only used highly realistic images rather than line drawings, but we also asked the participants to name the objects rather than just make same–different judgments. We thought it was important to make our experimental stimuli and task closer to real-world situations to allow us both to apply our findings and to use them for theoretical insights into the cognitive mechanisms used in everyday situations.

Second, we chose to examine how visual similarity affects identification performance within a well-defined category. Rather than examining similarities across different categories (superordinate vs. subordinate), we examined performance as a function of visual similarity within the single category of military fighter aircraft (relatively small, highly agile supersonic aircraft that are designed to engage in air-to-air combat; see Figure 1 for examples). By doing so, we were able to focus on the selective contribution of visual similarity to identification isolated from other factors (such as the semantic contributions involved in processing different categorization levels). Third, rather than subjectively manipulating visual similarity, we experimentally quantified similarity on the basis of participants' ratings.

In addition to the theoretical issues of object identification, we were interested in exploring how cognitive processes could be used to optimize identification training for military aircraft. Our experimental design involved training to identify aircraft, with transfer to novel orientations as the measure of learning. This transfer to novel orientations is a crucial issue in aircraft identification training because aircraft are viewed from a wide variety of orientations. An optimal training regimen would incorporate the least number of views necessary to achieve highly accurate identification of aircraft at novel orientations.

This type of research was pioneered by Gibson and Gagne (1947), who used scientific theory to guide the development of aircraft identification training procedures. In turn, their endeavors had a role in testing and forming theory. For example, they suggested that aircraft that look similar should be studied in pairs.

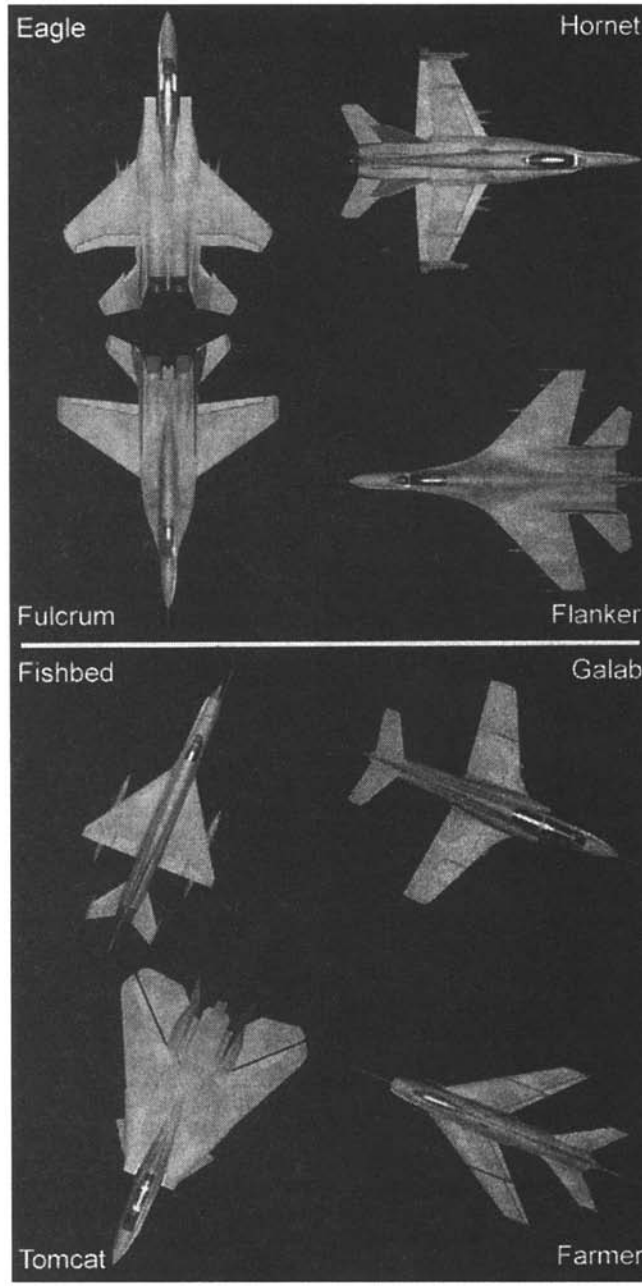

Figure 1. The eight aircraft used for identification training. The four aircraft in the top panel are relatively homogeneous in appearance and are shown from canonical viewpoints. The four aircraft in the bottom panel are relatively heterogeneous in appearance and are shown from noncanonical viewpoints.

They felt that learning to distinguish between similar aircraft would be facilitated by a direct comparison of the differences between critical, but subtle, metric features. However, from a training standpoint, such pairwise comparisons can have the unfavorable effect of substantially increasing the number of learning presentations. Thus, they also suggested that distinctive aircraft, which are less likely to be confused, could be studied individually.

Using identification accuracy as the dependent measure, Whitmore, Rankin, Baldwin, and Garcia (1972) suggested that as few as nine well-chosen views could facilitate accurate identification of aircraft at 45 novel views. With sufficient time to study a target image, accurate identification can be achieved based on only a very few learning views. However, the nature of the aviation environment requires accurate aircraft identification across many viewpoints while precluding lengthy analysis of the target image. The potentially severe consequences of inaccurate aircraft identification combined with the limited examination time of the target image mandate the collection and examination of response time data on the transfer of learning from the learning examples to novel orientations (see Dror, Busemeyer, & Basola, 1999, for the effects of time pressure on making such decisions). Furthermore, response time data are critical for assessing and understanding the underlying cognitive processes of object identification. Thus, we focused on response time for correct identification as the primary dependent measure.

In contrast to the above mentioned past research that focused primarily on the number of views needed for learning, our research was aimed at examining which types of presentations were best for learning. Thus, rather than examining quantitative issues per se, such as number of views used in training, we wanted to examine the qualitative issue of the informational content conveyed by different viewpoints and their contribution to identification. To be specific, we wanted to know whether canonical presentations (those in which the internal reference frame of the aircraft is aligned with the environmental reference frame) or noncanonical presentations (those that present aircraft in views where the reference frame of the aircraft is not aligned with the environment) facilitated learning and later performance at novel orientations. Our definition of canonicality follows Palmer, Rosch, and Chase (1981) and Dror and Kosslyn (1998) to reflect a typical view of an object. In contrast to the above mentioned usage, we limited ourselves to images within the two-dimensional picture plane, designating canonicality to objects that have clearly defined intrinsic frames of reference and face up, down, left, or right.

To summarize, for theoretical and applied reasons, we trained participants to identify military fighter aircraft. Participants learned to identify aircraft that were relatively similar to each other and aircraft that were relatively dissimilar to one another. Some participants learned to identify the aircraft with examples that presented the aircraft from a canonical viewpoint and other participants with examples from noncanonical viewpoints. Our measures included performance improvement during learning on the training examples and transfer identification performance to novel orientations used at test.

Method

Participants

Eighty-two participants were recruited through temporary agencies. There were 10 women and 72 men, with ages ranging from 18–49 years and a mean age of 22.75 years. Two participants were in their 40s, and three were in their 30s. All participants had either a high school diploma or a graduate equivalency diploma. By self-report, no participant had previous experience with aircraft identification. All participants were screened for visual acuity using a standard Schnelling eye chart, and only those whose vision was 20/30 or better took part in the study. Participants were tested at the United States Air Force TRAIN Laboratory at Brooks Air Force Base in San Antonio, Texas, and were paid for their time.

Materials

Stimuli were made from 12 high-resolution gray-scale images of military aircraft. The point of view of the aircraft was from directly above. The images were rendered from virtual three-dimensional models (made by ViewPoint International Inc.). The models were accurate representations of the military fighter aircraft and included visual information about texture and shading. Each aircraft image was scaled to be approximately 10 cm in length along the longest axis, from nose to tail. Participants sat approximately 60 cm from the computer monitor, making the stimulus images subtend a visual angle of approximately 9.5° (see Figure 1 for examples).

An upright 0° orientation was designated for each aircraft such that the aircraft appeared to be standing on its tail with its nose pointing straight up. Sixty-four images of each of the 12 aircraft were made by rotating the 0° image in the picture plane at increments of 5.625°, resulting in 768 total stimulus images (12 aircraft × 64 orientations).

The mask was a visual stimulus composed of a circular image with a gradient fill from black in the center, to gray in the middle, to white at the circumference. Colored in this manner, the mask looked like a dark-centered starburst with white edges. This mask appeared before and after every presentation of an aircraft stimulus.

The 12 aircraft were divided into two sets of stimuli according to their visual similarity. In one set, all the aircraft were relatively different from one another and relatively discriminable (the heterogeneous stimuli set). The other set of stimuli included relatively similar aircraft (the homogeneous stimuli set). The division of the 12 aircraft into the two stimuli sets was based on cluster analyses of similarity ratings. Participants rated the similarity between all possible pairs of 19 aircraft presented at identical upright positions (for more details, see Ashworth & Robbins, 1996). Following cluster analysis of the similarity ratings, we selected aircraft from different clusters for the heterogeneous set of stimuli and from within a single cluster for the homogeneous set. Each set of stimuli included 64 images of each aircraft, as described above.

From the six aircraft in each set of stimuli, two were designated as distractors and were used only during the final testing phase. We used the four remaining aircraft from each set for learning; these are presented in Figure 1. The heterogeneous stimuli set included the Fishbed (Mig-21), Galab (G-4), Tomcat (F-14), Farmer (Mig-19), Mirage (F-1), and Fresco (Mig-17; the latter two were used as distractors). The homogeneous stimuli set included the Eagle (F-15), Hornet (F-18), Flanker (SU-27), Fulcrum (Mig-29), Foxbat (Mig-25), and Falcon (F-16; the latter two were used as distractors).

From both the homogeneous and heterogeneous stimuli sets, two subsets of orientations were chosen for presentation during the learning phase. For half of the participants, we presented the eight aircraft at canonical orientations—that is, the intrinsic reference frame of the aircraft was aligned with the environmental reference frame. The canonical orientations were 0°, 90°, 180°, and 270° in the picture plane (Figure 1, top panel). For the other half of the participants, we presented the eight aircraft at noncanonical orientations—that is, the intrinsic reference frame of the aircraft did not align with the environmental reference frame. The noncanonical orientations were 22.5°, 112.5°, 202.5°, and 292.5° in the picture plane (Figure 1, bottom panel).

Procedure

Participants were tested in one session that took approximately 6 hr and included a 1-hr lunch break and a number of shorter breaks, as specified below. Participants were tested in groups that varied in size from 9 to 20. Each participant sat in a three-sided cubicle. Before the experiment, we gave verbal instructions to the participants and they asked questions and received clarifications. Participants were instructed not to interact with each other.

The experiment was administered on IBM-compatible microcomputers with high-resolution 15-in. (38-cm) monitors. Responses were made on the keyboards. Accuracy and response time were recorded by the computers using a timing technique by Haussmann (1992) that provides millisecond precision.

The experiment had two factors, each with two levels, resulting in four conditions. The first factor, Presentation, manipulated the orientations (canonical or noncanonical) at which the aircraft stimuli were learned and was a between-participant factor. Participants were randomly assigned to either the canonical presentation condition (0°, 90°, 180°, and 270°) or to the noncanonical presentation condition (22.5°, 112.5°, 202.5°, and 292.5°). Those assigned to the canonical condition saw only canonically oriented presentations of aircraft during learning. Those assigned to the noncanonical condition saw only noncanonically oriented presentations of aircraft during learning. The testing phase was identical for all participants, regardless of their learning presentation condition.

The second factor, Discriminability, varied the degree of visual similarity (homogeneous or heterogeneous) among aircraft to be learned and was a within-participant factor. Each participant learned, and then was tested on, both the homogeneous and the heterogeneous stimuli. The administration order of the homogeneous and heterogeneous stimuli was counterbalanced across participants. Thus in each presentation condition (canonical or noncanonical), half the participants first performed the experiment with the homogeneous set and then with the heterogeneous set, and the other half of the participants performed the experiment in the reverse order. Participants were given a lunch break after completing the first stimuli set of the Discriminability factor and before moving on to the next set.

To summarize, the design of the experiment was a 2 (presentation) × 2 (discriminability) full factorial. The Presentation factor (canonical and noncanonical) was a between-participant factor, and the Discriminability factor (homogeneous and heterogeneous) was a within-participant factor. Thus the four experimental conditions were as follows: (a) canonical and homogeneous, (b) canonical and heterogeneous, (c) noncanonical and homogeneous, and (d) noncanonical and heterogeneous.

The procedure was identical for each of the four experimental conditions. There was a learning phase followed by a testing phase. Each phase was composed of blocks of trials. Trials consisted of the presentation of an aircraft stimulus for 5 s or until a response was made, whichever came first. A stimulus mask appeared for 100 ms before and after each aircraft presentation. Participants responded by pressing one of five keys on the computer keyboard. Four of the keys were labeled with the aircraft names. The fifth key was labeled "NA" for "none of the above."

Learning phase. The learning phase contained 10 blocks of trials. Eight of the 10 learning blocks contained 160 trials each; the remaining 2 blocks contained 16 trials each, giving a total of 1,312 learning trials. The trials within each block were randomized. The order of blocks was the same for each experimental condition.

The purpose of Block 1 was to teach the participants to associate the name of the aircraft with its respective images. The four aircraft were presented 10 times at each of the four learning orientations, resulting in 160 trials (4 aircraft × 4 orientations × 10 repetitions). On each trial, the name of the aircraft was presented along with the aircraft image. Participants were instructed to study the image for the full 5 s and to associate the name with the aircraft. After the aircraft image disappeared from the monitor, the participants pressed the key labeled with the correct aircraft name. Participants were instructed to concentrate and strive for perfect accuracy because their identification of the four aircraft would be tested later. Accuracy feedback was provided by a beep if an incorrect key was pressed.

Blocks 2, 3, and 4 were identical to each other. Similar to Block 1, the four aircraft were presented 10 times at each of the four learning orientations, resulting in 160 trials. However, unlike Block 1, the aircraft names were not presented along with the aircraft images. The purpose was to ensure that the participants had learned the names of the aircraft. Participants were instructed to respond as quickly as possible while maintaining high accuracy. Accuracy feedback was provided by a beep if an incorrect key was pressed. After Block 4, participants were given a 10-min rest break.

Because participants had just returned from a 10-min rest break, Block 5 served as a reminder of the aircraft names and offered the participants another opportunity to associate the aircraft names with their respective images without any time pressure. Similar to Block 1, the aircraft names were presented along with the aircraft images, and participants were instructed to study each image for the full 5 s. Unlike Block 1, the four aircraft were presented only once at each of the four learning orientations, resulting in 16 trials (4 aircraft × 4 orientations). Accuracy feedback was provided by a beep if an incorrect key was pressed.

Blocks 6, 7, 8, and 9 were identical to Blocks 2, 3, and 4. The aircraft were presented without names, and participants were instructed to respond as quickly as possible while maintaining high accuracy. Accuracy feedback was provided by a beep if an incorrect key was pressed. After Block 9, participants were given a 30-min rest break.

Block 10 was the last block of the learning phase and was identical to Block 5. The aircraft names were presented along with the aircraft images, and participants were instructed to study each image for the full 5 s. The four aircraft were presented once at each of the four learning orientations, resulting in 16 trials. Accuracy feedback was provided by a beep if an incorrect key was pressed.

Testing phase. The testing phase consisted of 320 randomized trials. The four aircraft were presented once at each of the 64 test orientations for a total of 256 presentations (4 aircraft × 64 orientations). In addition, there were 64 distractor trials, using two distractor aircraft that were evenly distributed across the 64 testing orientations.

Participants were instructed to respond as quickly as possible while maintaining high accuracy. Participants were informed that distractors had been added and that they were to press the key labeled "NA" for "none of the above" at the presentation of a distractor. Because there were 64 presentations of each of the aircraft and 64 presentations of the distractors, there were an equal number of responses assigned to each of the five response keys. In contrast to the learning phase, no accuracy feedback was provided during the testing phase.

The testing phase was preceded by 24 practice trials to acquaint the participants with trials that included distractors. The six aircraft (four from the learning phase and two distractors) were presented once at each of the learning orientations for a total of 24 trials (6 aircraft × 4 orientations).

Results

Accuracy rates were very high, most probably due to the explicit instructions that stressed the importance of accurate responses. Accuracy was above 94% at all orientations in all blocks of both the learning phase and the testing phase. Thus, the high accuracy that we wanted, and obtained, resulted in a ceiling effect that prevented meaningful analysis of the error rates. All reported analyses were performed on the response time data for correct responses.

Medians were calculated for each stimuli set at each orientation, for each block, for each participant. Medians were used because they provided an accurate representation of central tendency of individual participants, eliminating the need for various methods for trimming outliers. It was on these medians that all subsequent data analyses were performed. The learning phase data were analyzed separately from the test phase data.

Learning Phase

The analysis included only the learning data from blocks in which the aircraft names were not presented along with the images and the participants were instructed to respond as quickly as possible while maintaining high accuracy (Blocks 2, 3, 4, 6, 7, 8, and 9). Blocks 1, 5, and 10 do not represent meaningful data because these blocks only measured participants' ability to read a name on the monitor, wait 5 s, and then press a key with the same name. For simplicity, the seven blocks from the learning phase that were used in the analysis are henceforth be referred to as Blocks 1–7.

The analysis examined performance across the seven learning blocks. Presentation (canonical vs. noncanonical) was a between-participant factor. Within-participant factors were Discriminability (homogeneous vs. heterogeneous), Learning Blocks (Blocks 1–7), and Learning Orientations (the four orientations used during learning). By itself, the Learning Orientations factor is not meaningful because it collapses across the two presentation conditions (0° and 22.5°, 90° and 112.5°, 180° and 202.5°, 270° and 292.5°), and thus it is not reported. However, all terms that include the Presentation × Learning Orientations interaction are meaningful because this interaction separates the canonical learning orientations from the noncanonical learning orientations. The overall design of the analysis of variance (ANOVA) for the learning phase was 2 (presentation) × 2 (discriminability) × 7 (blocks) × 4 (learning orientations).

The ANOVA revealed significant main effects for presentation, $F(1, 80) = 4.54$, $p < .05$, $MSE = 545{,}751$ (mean response times of 784 ms and 830 ms for canonical and noncanonical, respectively); for discriminability, $F(1, 80) = 86.83$, $p < .001$, $MSE = 110{,}663$ (mean response times of 852 ms and 761 ms for homogeneous and heterogeneous, respectively); and for block, $F(6, 480) = 120.80$, $p < .001$, $MSE = 14{,}128$ (mean response times of 905 ms, 840 ms, 814 ms, 762 ms, 770 ms, 781 ms, and 777 ms for Blocks 1–7, respectively). Learning orientations as a main effect is not reported; as explained previously, it is not meaningful because it collapses across canonical and noncanonical presentations.

As illustrated in the bottom panel of Figure 2, the ANOVA revealed an interaction for Discriminability × Block, $F(6, 480) = 27.89$, $p < .001$, $MSE = 9736.86$, reflecting that during the first four blocks, participants' performance with the homogeneous stimuli set improved more rapidly than with the heterogeneous stimuli set. There was also an interaction for Discriminability × Block × Orientation, $F(18, 1440) = 1.98$, $p < .01$, $MSE = 1566.85$. However, this interaction could not be interpreted because of the manner in which the Orientation factor is collapsed. There were no other significant interactions (all $ps > .20$).

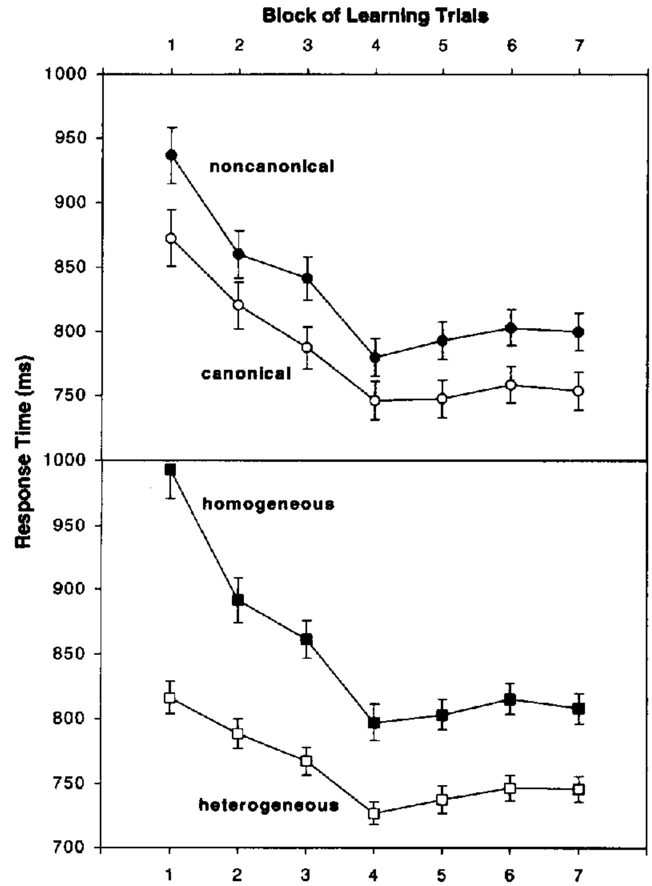

Figure 2. Response time (in milliseconds) for identification of aircraft during learning. The graphs plot performance across learning blocks. Error bars indicate the normalized standard error. The top panel depicts correct identification at the canonical and noncanonical orientations used during learning. The bottom panel depicts correct identification of the relatively dissimilar set of aircraft (heterogeneous set) and of the relatively similar set of aircraft (homogeneous set).

Testing Phase

Participants' identification performance was tested at 64 orientations that were equally spaced every 5.625° in the picture plane. Of these 64 orientations, 4 were the original training orientations, whereas the remaining 60 were novel orientations. To assess identification performance as a function of the learning orientations, we collapsed the data for the novel orientations according to their disparity from the nearest learning orientation. Thus, all data cells that were 5.625° from one of the four learning orientations were collapsed, all cells that were 11.250° from a learning orientation were collapsed, all cells that were 16.875° from a learning orientation were collapsed, and so forth, up to 45°. Because the learning orientations were separated by 90°, the collapsed data could be no more than 45° from a learning orientation (any angular disparity larger than 45° from one learning orientation would be closer to another learning orientation). After collapsing, the data represented identification performance at 0.000°, 5.625°, 11.250°, 16.875°, 22.500°, 28.125°, 33.750°, 39.375°, and 45.000° from the nearest learning orientation.

Presentation (canonical vs. noncanonical) was a between-participant factor. Within-participant factors were Discriminability (homogeneous vs. heterogeneous) and Angular Disparity (the angular disparity of novel orientations from the nearest learning orientation). Thus, the overall design of the ANOVA for the test phase was 2 (presentation) × 2 (discriminability) × 9 (angular disparity).

The ANOVA revealed significant main effects for the Presentation factor, $F(1, 80) = 5.65$, $p < .05$, $MSE = 322{,}875$ (with mean response times of 837 ms and 907 ms for canonical and noncanonical, respectively); for the Discriminability factor, $F(1, 80) = 57.87, p < .001, MSE = 60{,}913$ (with mean response times of 921 ms and 823 ms for homogeneous and heterogeneous, respectively); and for the Angular Disparity factor, $F(8, 640) = 9.88, p < .001, MSE = 4{,}639$ (with mean response times of 850 ms, 855 ms, 857 ms, 866 ms, 879 ms, 877 ms, 875 ms, 896 ms, and 895 ms for increasing angular disparity from the nearest learning orientation).

As illustrated in Figure 3, the ANOVA revealed interactions for Presentation × Angular Disparity, $F(8, 640) = 2.47, p < .05, MSE = 4{,}639$ (top panel), and Discriminability × Angular Disparity, $F(8, 640) = 3.06, p < .01, MSE = 5{,}026$ (bottom panel). There were no other significant interactions (all $ps > .40$).

A linear contrast on the Angular Disparity factor revealed that there was a significant linear trend, $F(1, 80) = 43.27, p < .001, MSE = 7812.31$, with response time increasing as a function of angular disparity from the nearest learning orientation. We per-

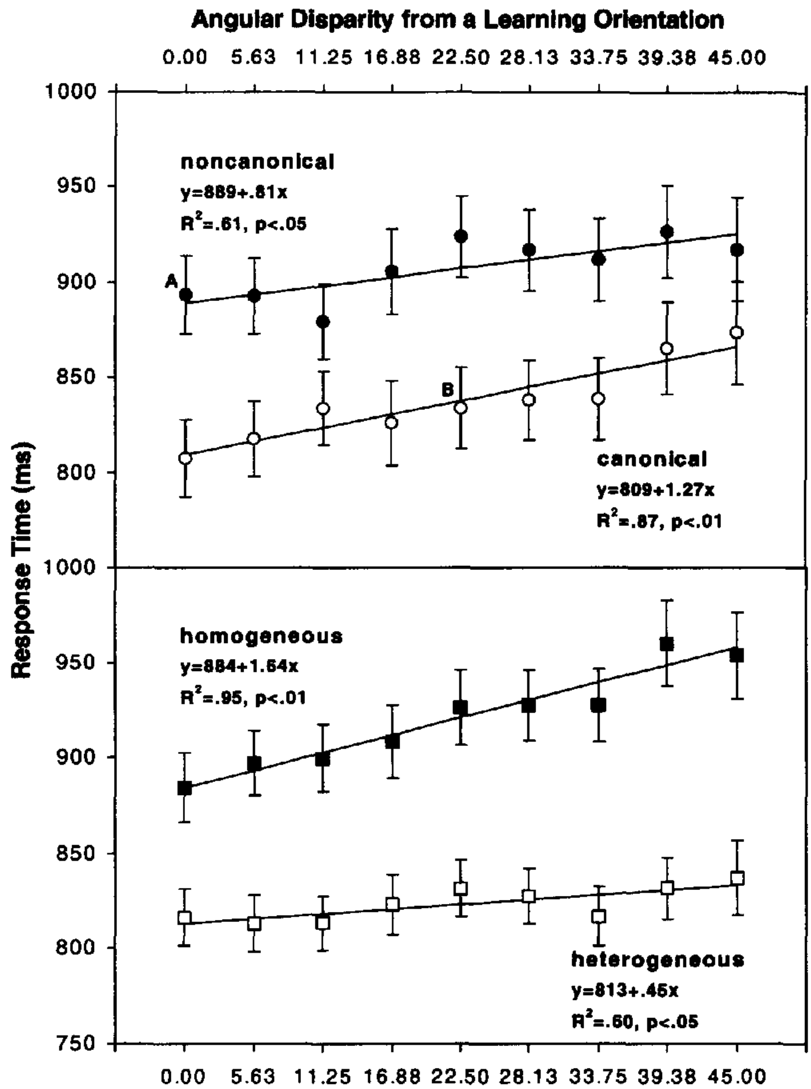

Figure 3. Response time (in milliseconds) for identification of aircraft during testing. Response time is plotted as a function of angular disparity between novel orientations presented during test and the nearest learning orientation. Error bars indicate the normalized standard error. The top panel depicts correct identification for aircraft that were presented during learning at either canonical or noncanonical orientations. Points A and B indicate identification performance at the four noncanonical orientations (22.5°, 112.5°, 202.5°, and 292.5°) at test. Point A represents aircraft that were learned at those same four noncanonical orientations. Point B represents aircraft that were learned at different orientations (the four canonical orientations: 0°, 90°, 180°, and 270°). Nevertheless, identification performance at Point B is better than Point A ($p < .05$). The bottom panel depicts correct identification of the relatively dissimilar set of aircraft (heterogeneous set) and of the relatively similar set of aircraft (homogeneous set). The high linear slope describing the homogeneous set suggests a normalization mechanism used in an orientation-dependent identification process, whereas the lack of such a slope in the heterogeneous set suggests an orientation-invariant process.

formed regression analyses on both levels of the Discriminability factor (homogeneous vs. heterogeneous) as well as both levels of the Presentation factor (canonical vs. noncanonical). In all four instances, a linear relationship was found between angular disparity from the closest learning orientation and response time: homogeneous, slope = 1.64 milliseconds per angle of rotation (ms/°), $R^2 = .95$, $p < .001$; heterogeneous, slope = .45 ms/°, $R^2 = .60$, $p < .05$; canonical, slope = 1.27 ms/°, $R^2 = .87$, $p < .001$; noncanonical, slope = .81 ms/°, $R^2 = .61$, $p < .05$. Figure 3 shows the regression lines fitted to their respective functions.

To examine whether the slopes for the homogeneous and heterogeneous levels of the Discriminability factor were equivalent, we computed a Linear × Linear contrast. The slope relating response time to angular disparity for homogeneous aircraft was steeper than that for the heterogeneous stimuli, $F(1, 80) = 13.26$, $p < .001$, $MSE = 8{,}280$. A similar Linear × Linear contrast computed on the slopes for the canonical and noncanonical levels of the Presentation factor failed to reach significance, $p = .15$, suggesting that they were comparable. Within both the canonical and noncanonical learning conditions, participants' performance was best at the specific orientations used during training.

We were specifically interested in comparing identification performance at 0° angular disparity for the noncanonical stimuli with 22.5° angular disparity for the canonical stimuli (Points A and B, Figure 3, top panel). Both Points A and B represent the same four orientations at test: 22.5°, 112.5°, 202.5°, and 292.5°. However, they represent different orientations during learning. Point A represents learning orientations of 22.5°, 112.5°, 202.5°, and 292.5°, which are identical to those seen at test. Point B represents learning orientations of 0°, 90°, 180°, and 270°, and thus the test orientations at point B are all rotated 22.5° clockwise from the learning orientations. Therefore, Point A represents identification of the same images studied during the learning phase, whereas Point B represents identification of the images at novel orientations. Nevertheless, studying the aircraft at canonical orientations promoted better identification at the noncanonical orientations than did both studying and testing at the noncanonical orientations, $F(1, 80) = 4.05$, $p < .05$, $MSE = 35{,}868$ (with mean response times of 893 ms and 834 ms for Points A and B, respectively).

Discussion

Stimulus Similarity

During the learning phase, the aircraft from the heterogeneous and homogeneous stimuli sets showed a comparable pattern of results. Identification performance improved as learning progressed, and around the fourth block improvement reached asymptote. As expected, aircraft from the heterogeneous stimuli set were consistently easier to identify than aircraft from the homogeneous stimuli set. After performance had leveled off, the heterogeneous aircraft were identified approximately 65–70 ms faster than the homogeneous aircraft.

The test phase continued to reflect that aircraft from the homogeneous set were more difficult to identify. However, in contrast to the learning phase, a different pattern of results emerged from the two sets of aircraft stimuli. Degradation of identification performance as a function of the angular disparity between the learning orientations and the novel orientations was much higher for the homogeneous aircraft than for the heterogeneous aircraft. In fact, the heterogeneous set of aircraft had almost no slope. The slope of 0.45 ms/° for this condition is below what is considered to reflect a normalization process, whereas the slope of 1.64 ms/° for the homogeneous aircraft is typical of normalization (such as mental rotation; see Cohen & Kubovy, 1993, and Tarr & Pinker, 1989, for reviews of mental rotation slopes).

The increase in response time found for the homogeneous set of aircraft is consistent with viewpoint-dependent theories (e.g., Tarr, 1995; Tarr & Bülthoff, 1995; Tarr & Pinker, 1989), which suggest that the incoming visual images are normalized and compared to viewpoint-specific representations stored in memory. Such theories assign a very specific role for the training examples, namely, that they are used as basic representational exemplars for identification of future stimuli. Given the importance of initial presentations in establishing long-lasting representations (DiGirolamo & Hintzman, 1997; Dror, Ashworth, Schreiner, Robbins, & Snooks, 1997), careful consideration must be given to selection of the training examples.

However, viewpoint-dependent theories were only supported by the homogeneous set of stimuli. The data from the heterogeneous set of stimuli reflected that novel stimuli were identified without a normalization process, thus supporting a viewpoint-invariant mechanism (e.g., Biederman, 1987; Biederman & Gerhardstein, 1993). Such theories attribute a different role for the training examples, namely, that they are the means for establishing abstract, orientation-invariant representations. For such theories, the examples used in training may not be as critical as in viewpoint-dependent theories. However, they are still important tools and have to be well chosen in order to facilitate formation of good abstract representations.

Hence, our data demonstrate that neither a viewpoint-dependent nor a viewpoint-invariant mechanism underlie all identification processes. The question remains, What is the underlying cognitive reason for the use of one mechanism or the other? The answer has theoretical significance because it highlights the distinctiveness and relative advantage of each mechanism. It also has applied significance because it can help construct training procedures that best facilitate the mechanisms used.

Hamm and McMullen (1998) proposed that the type of category identification judgment (subordinate or superordinate) plays a key role in determining if viewpoint-dependent or viewpoint-invariant mechanisms are used. Murray (1998) suggested that the similarity of the stimuli sets could equally well account for the data reported by Hamm and McMullen (1998).

Consistent with Murray (1998), our findings suggest that visual similarity is a crucial factor in determining which mechanism will govern the identification process. However, we interpret the similarity effect as reflecting different computational demands needed for identification. Orientation-invariant representations are abstractions and, hence, do not encode all of the fine metric details. Thus, when the identification task was relatively easy (when the aircraft were dissimilar), then fine metric details were not needed for successful identification and an orientation-invariant mechanism was used. Its use enables identification without resorting to computationally expensive cognitive processes, such as normalization. In contrast, orientation-dependent presentations more fully preserve fine metric details. Thus, when the identification task was relatively difficult (when the aircraft were similar), then fine metric details were necessary to enable correct identification and an orientation-dependent mechanism was used. In this case, similarity is the factor that influences the computational demands that drive the identification process.

Furthermore, because the stimuli used in our study were members of a single highly constrained category (military fighter aircraft), it is difficult to argue that differences in category membership (e.g., subordinate, superordinate) are responsible for the performance differences between the homogeneous and heterogeneous aircraft. Using similar and dissimilar military fighter aircraft enabled us to isolate the similarity effects from the category membership effects. The use of a highly specialized domain (military fighter aircraft) with which naive participants had no prior experience enabled us to clearly address our research questions, avoiding issues related to the use of common objects (cars, dogs, and so forth) for which participants already had prior knowledge. And finally, cluster analyses of visual similarity ratings of the aircraft were used to determine the homogeneous and heterogeneous aircraft sets, thus empirically validating that similarity was indeed the distinguishing factor between the stimuli sets. This finding establishes similarity as an experimentally quantifiable factor that can determine which mechanism will govern the identification process.

Our data support Edelman's (1995) use of abstract three-dimensional representations of animals to demonstrate that stimuli more similar in appearance are governed more by an orientation-dependent mechanism. Edelman (1991a, 1991b, 1995) proposed a model based on so-called features of recognition in which recognition is based on two-dimensional, image-based representations rather than structural three-dimensional images. Processing based on fine-grain metric details that are well-localized within the image-based representations will lead to performance that is viewpoint-dependent. In contrast, viewpoint-invariant performance results from processing gross qualitative features that are defined over larger regions of the representation. As objects become more similar in appearance, the necessity to rely on finer details for discrimination increases, and thus viewpoint-dependency also increases.

The above conclusions can be derived with relative confidence because we were able to form sets of stimuli with consistent levels of similarity across orientations. This consistency was achieved by using picture plane orientations in which visual information is held constant for each aircraft regardless of orientation. Introducing out-of-picture-plane orientations changes the information that is presented by the image and, hence, varies the similarity of that aircraft to other aircraft in the set (e.g., when distinctive features are obscured, two dissimilar aircraft may appear more similar).

The data reported in this article have considerable utility in that they demonstrate identification dissociations between different sets of aircraft. In contrast to so-called pure laboratory research that has been conducted with oversimplified or artificially contrived stimuli and unrealistic procedures (such as same–different judgments), our data are based on stimuli and procedures that are closer to real-world circumstances. Moreover, because of the stimuli used, the results have direct bearing on identification training of military aircraft.

In terms of training military aircraft identification, our data demonstrate that the more distinctive the individual aircraft are within a set, the more transfer of learning will occur from the initial learning orientation to novel orientations. This in turn suggests that for distinctive aircraft, fewer total views would be needed during training to ensure accurate identification at novel orientations. In comparison, aircraft that are more similar in appearance will require incorporating more views to ensure accurate identification across novel orientations. Thus, the relationship between homogeneity of the stimuli set and the views required to optimize identification performance across all novel orientations is highly dependent on the specific stimuli set (i.e., possible confusers). Furthermore, our data show the need to construct the difficulty of training examples in accordance with the demands of real-world situations, thus ensuring the development of representations that are cognitively suited to the kind of identification tasks that observers will perform. Our use of picture plane orientations is appropriate for such real-world fighter aircraft identification applications because fighter aircraft are extremely maneuverable and, thus, top views are readily available to a wide variety of viewpoints.

Canonical Frame Alignment

During the learning phase, aircraft that were presented at noncanonical orientations (22.5°, 112.5°, 202.5°, and 292.5°) consistently took more time to be identified than those presented at canonical orientations (0°, 90°, 180°, and 270°). Even after performance reached asymptote, the aircraft presented at canonical orientations were identified approximately 40–45 ms faster.

During the test phase, aircraft identification was tested at 64 different orientations, which included both the canonical and noncanonical orientations used during learning. The initial learning presentation (canonical vs. noncanonical) had a strong impact on identification performance during testing. Participants who studied the aircraft at canonical orientations were much better at identifying aircraft during testing than participants who studied aircraft at noncanonical orientations. As seen in the top panel of Figure 3, this performance advantage for canonically presented aircraft was evident at the learning orientations (0° for canonical is faster than 0° for noncanonical) and held across all novel orientations.

Furthermore, identification performance at 22.5° from a canonical learning orientation was better than that at 0° from a noncanonical learning orientation (Figure 3, top panel, Points A and B). This finding means that training at the canonical orientations and then testing at the noncanonical orientations produces better identification than training at the noncanonical orientations and then testing at the same noncanonical orientations! These effects for the canonical versus noncanonical presentations appear robust because performance was asymptotic after four training blocks (Figure 2). We attribute these data to the effect of canonicality per se because information is consistent across all picture plane orientations. We chose not to include images out of the picture plane because they alter the information available across orientations and hence introduce alternative explanations for data.

Our results are consistent with the reference frame literature. Summarizing the literature, Palmer (1989) noted that object identification is mediated primarily by the intrinsic reference frame for objects whose structure clearly defines one and by an environmental reference frame for objects whose structure does not clearly define an intrinsic reference frame. The best identification occurs when intrinsic reference frames align with the environmental reference frame (see also Appelle, 1972; Corballis, Nagourney, Shetzer, & Stefanatos, 1978; Friedman & Hall, 1996; Hinton & Parsons, 1988; Palmer et al., 1981; Pani, 1993, 1994; Pani & Dupree, 1994; Pani, Jeffres, Shippey, & Schwartz, 1996; Pani, William, & Shippey, 1995; Parsons, 1995; Shiffrar & Shepard, 1991). Our study suggests that identification performance will be enhanced in any pattern recognition task in which the intrinsic reference frames align with the environmental reference frame.

Furthermore, our results are consistent with the literature that shows that the initial examples used during training have long-lasting effects (e.g., DiGirolamo & Hintzman, 1997; Dror et al., 1997). This finding is not limited to military fighter aircraft identification training but, rather, applies to any knowledge acquisition process. It highlights the importance of considering factors that govern the formation of initial mental representations.

The existence of cognitive mechanisms that perform best when aircraft (and other objects) are studied at canonical orientations is probably attributable to our experience with objects and the environment. As Shepard (1982, 1984) and Shiffrar and Shepard (1991) suggested, this effect is likely attributable to the invariant nature of the terrestrial gravitational field and its predictable influence on objects. Objects with primary internal axes are most stable if their axes are either in line with gravity or orthogonal to gravity, and our perceptual apparatus exploits these regularities in the world. Thus, our study suggests that when attempting to understand cognition, one must take into account the nature of mental representations in addition to "pure" computational considerations. Although our cognitive system is computational, it is not necessarily the most efficient (Dror & Gallogly, 1999; Dror, Zagaeski, & Moss, 1995). Our findings demonstrate that two computationally identical learning presentations (canonical and noncanonical orientations) containing the same information do not lead to the same behavioral results. The different results derive not from a computational difference in the processing of the two learning presentations but, rather, from the nature of the underlying mental representations.

In terms of teaching identification of objects with well-defined intrinsic reference frames (e.g., aircraft), our results suggest that it would be efficient for the training examples to focus on views in which the intrinsic and environmental reference frames align. This conclusion holds true at least for novice observers, because there is evidence that skilled observers, such as experienced pilots, may develop unique information-processing mechanisms that incorporate different representational formats and reference frames (Dror, 1992; Dror, Kosslyn, & Waag, 1993). The study of mental representations and how they interact with expertise is important in developing real-world applications. Such representations may also depend on the type of stimulus, its meaningfulness, and our familiarity with it (Dror, Ivey, & Rogus, 1997).

Summary and Conclusions

Explication of the cognitive factors underlying aircraft identification provides insight into the nature of, and processes associated with, the representations used to identify real-world objects. This explication in turn serves as a touchstone for theories of object identification, whose significance and validation should ultimately derive from the extent to which their predictions hold in real-world circumstances. Furthermore, the results reported here have direct application in the development of effective aircraft identification training.

Aircraft identification within the military aviation environment must be accurate and rapid. Aircraft identification training that meets these requirements is a challenge because observers are required to learn a large set of complex and variant stimuli in a relatively short time. Furthermore, there are severe consequences associated with inaccurate or slow identification of military aircraft.

In the research reported here, we examined the effect of stimulus similarity and canonical frame alignment on aircraft identification. We found that both the level of homogeneity of the set of aircraft and the initial orientation at which they are learned have a powerful effect on identification performance during learning and later during testing at novel orientations. That is, homogeneous stimuli sets are harder to learn than heterogeneous stimuli sets and are governed by an orientation-dependent mechanism. This finding, in turn, suggests that normalization processes for aligning incoming percepts with stored representations are more readily engaged by homogeneous stimuli sets than by heterogeneous stimuli sets. In contrast, heterogeneous stimuli sets are easier to learn and are governed by an orientation-invariant mechanism.

Furthermore, even if performance in the real world requires identification at noncanonical orientations (e.g., a military aviation environment), training should nevertheless be based on canonical presentations that align the intrinsic and environmental reference frames. Our findings also question the validity of the prevalent "pure" viewpoint-dependent and viewpoint-invariant theories and support the notion that the underlying representations of objects and their processing is highly dependent on the task and computational demands during learning and testing.

We contend that this research has implications both for theories of object identification and for the development of training procedures. We showed how such a relationship can be reciprocal—one in which theories are applied to real-world settings but also one in which real-world settings provide the test for theories and can constrain and guide their development.

References

Appelle, S. (1972). Perception and discrimination as a function of stimulus orientation: The oblique effect in man and animals. *Psychological Bulletin, 78,* 266–278.

Ashworth, A. R. S., III, & Robbins, R. D. (1996). Visual homogeneity affects orientation sensitivity. *Psychonomic Abstracts, 1,* 63.

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

Biederman, I., & Gerhardstein, P. C. (1993). Recognizing depth-rotated objects: Evidence and conditions for three-dimensional viewpoint invariance. *Journal of Experimental Psychology: Human Perception and Performance, 19,* 1162–1182.

Cohen, D., & Kubovy, M. (1993). Mental rotation, mental representation, and flat slopes. *Cognitive Psychology, 25,* 351–382.

Corballis, M. C., Nagourney, B. A., Shetzer, L. I., & Stefanatos, G. (1978). Mental rotation under head tilt: Factors influencing the location of the subjective reference frame. *Perception & Psychophysics, 24,* 263–273.

DiGirolamo, G. J., & Hintzman, D. L. (1997). First impressions are lasting impressions: A primacy effect in memory for repetitions. *Psychonomic Bulletin and Review, 4,* 121–124.

Dror, I. E. (1992). Visual mental rotation: Different processes used by pilots. *Proceedings of the Human Factors Society 36th Annual Meeting, 2,* 1368–1372.

Dror, I. E., Ashworth, A. R. S., III, Schreiner, C. S., Robbins, R. D., & Snooks, S. F. (1997). The primacy effect on identification: Initial presentations during training establish long lasting representations. *Psychonomic Abstracts, 2,* 64.

Dror, I. E., Ashworth, A. R. S., III, & Stevenage, S. V. (1999). Mediating aircraft identification by manipulating distinctiveness, stimulus similarity, and learning presentations. *Psychonomic Abstracts, 4,* 23–24.

Dror, I. E., Busemeyer, J. R., & Basola, B. (1999). Decision making under time pressure: An independent test of sequential sampling models. *Memory & Cognition, 27,* 713–725.

Dror, I. E., & Gallogly, D. (1999). The role of computational analyses in cognitive neuroscience: In defense of biological implausibility. *Psychonomic Bulletin & Review, 6,* 173–182.

Dror, I. E., Ivey, C., & Rogus, C. (1997). Visual mental rotation of possible and impossible objects. *Psychonomic Bulletin & Review, 4,* 242–247.

Dror, I. E., & Kosslyn, S. M. (1998). Age degradation in top-down processing: Identifying objects from canonical and noncanonical viewpoints. *Experimental Aging Research, 24,* 203–216.

Dror, I. E., Kosslyn, S. M., & Waag, W. L. (1993). Visual-spatial abilities of pilots. *Journal of Applied Psychology, 78,* 763–773.

Dror, I. E., Zagaeski, M., & Moss, C. F. (1995). Three-dimensional target recognition via sonar: A neural network model. *Neural Networks, 8,* 143–154.

Edelman, S. (1991a). A self-organizing multiple-view representation of 3D objects. *Biological Cybernetics, 64,* 209–219.

Edelman, S. (1991b). *Features of recognition. CS-TR 91-10.* Rechovot, Israel: Weizmann Institute of Science.

Edelman, S. (1995). Class similarity and viewpoint invariance in the recognition of 3D objects. *Biological Cybernetics, 72,* 207–220.

Friedman, A., & Hall, D. L. (1996). The importance of being upright: Use of environmental and viewer-centered reference frames in shape discriminations of novel three-dimensional objects. *Memory & Cognition, 24,* 285–295.

Gibson, J. J., & Gagne, R. M. (1947). Research on the recognition of aircraft. In J. J. Gibson (Ed.), *Motion picture testing and research* (Army Air Forces Aviation Psychology Program Research Report No. 7, pp. 113–168). Washington, DC: U.S. Government Printing Office.

Hamm, J. P., & McMullen, P. A. (1998). Effects of orientation on the identification of rotated objects depend on the level of identity. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 413–426.

Haussmann, R. E. (1992). Tachistoscopic presentation and millisecond timing on the IBM PC/XT/AT and PS/2: A turbo Pascal unit to provide general-purpose routines for CGA, Hercules, EGA, and VGA monitors. *Behavior Research Methods, Instruments, and Computers, 24,* 303–310.

Hinton, G. E., & Parsons, L. M. (1988). Scene-based and viewer-centered representations for comparing shapes. *Cognition, 30,* 1–35.

Murray, J. E. (1998). Is entry level recognition viewpoint invariant or viewpoint dependent? *Psychonomic Bulletin & Review, 5,* 300–304.

Palmer, S. E. (1989). Reference frames in the perception of shape and orientation. In B. E. Shepp & S. Ballesteros (Eds.), *Object perception: Structure and process* (pp. 121–163). Hillsdale, NJ: Erlbaum.

Palmer, S. E., Rosch, E., & Chase, P. (1981). Canonical perspective and the perception of objects. In J. Long & A. Baddeley (Eds.), *Attention & performance IX* (pp. 135–151). Hillsdale, NJ: Erlbaum.

Pani, J. R. (1993). Limits on the comprehension of rotational motion: Mental imagery of rotations with oblique components. *Perception, 22,* 785–808.

Pani, J. R. (1994). The generalized cone in human spatial organization. *Spatial Vision, 8,* 491–501.

Pani, J. R., & Dupree, D. (1994). Spatial reference systems in the comprehension of rotational motion. *Perception, 23,* 929–946.

Pani, J. R., Jeffres, J. A., Shippey, G. T., & Schwartz, K. J. (1996). Imagining projective transformations: Aligned orientations in spatial organization. *Cognitive Psychology, 31,* 125–167.

Pani, J. R., William, C. T., & Shippey, G. T. (1995). Determinants of the perception of rotational motion: Orientation of the motion to the object and to the environment. *Journal of Experimental Psychology: Human Perception and Performance, 21,* 1441–1456.

Parsons, L. M. (1995). Inability to reason about an object's orientation using an axis and angle of rotation. *Journal of Experimental Psychology: Human Perception and Performance, 21,* 1259–1277.

Shepard, R. N. (1982). Perceptual and analogical bases of cognition. In J. Mehler, E. C. T. Walker, & M. Garrett (Eds.), *Perspectives on mental representation* (pp. 49–67). Hillsdale, NJ: Erlbaum.

Shepard, R. N. (1984). Ecological constraints on internal representation: Resonant kinematics of perceiving, imagining, thinking, and dreaming. *Psychological Review, 91,* 417–447.

Shiffrar, M. M., & Shepard, R. N. (1991). Comparison of cube rotations around axes inclined relative to the environment or to the cube. *Journal of Experimental Psychology: Human Perception and Performance, 17,* 44–54.

Tarr, M. J. (1995). Rotating objects to recognize them: A case study on the role of viewpoint dependency in the recognition of three-dimensional objects. *Psychonomic Bulletin & Review, 2,* 55–82.

Tarr, M. J., & Bülthoff, H. H. (1995). Is human object recognition better described by geon-structural-descriptions or by multiple-views? Comment on Biederman and Gerhardstein (1993). *Journal of Experimental Psychology: Human Perception and Performance, 21,* 1494–1505.

Tarr, M. J., & Pinker, S. (1989). Mental rotation and orientation dependency in shape recognition. *Cognitive Psychology, 21,* 233–282.

Whitmore, P. G., Rankin, W. C., Baldwin, R. O., & Garcia, S. (1972). *Studies of aircraft recognition training* (Human Resources Research Organization Technical Report 72-5). Fort Bliss, TX: Human Resources Research Organization Division 5 (Air Defense).

Received July 13, 1998
Revision received November 29, 1999
Accepted December 3, 1999 ■