Enhancing Air Traffic Displays via Perceptual Cues

EVAN M. PALMER
University of California, Los Angeles and HRL Laboratories, LLC
TIMOTHY C. CLAUSNER
HRL Laboratories, LLC
and
PHILIP J. KELLMAN
University of California, Los Angeles

We examined graphical representations of aircraft altitude in simulated air traffic control (ATC) displays. In two experiments, size and contrast cues correlated with altitude improved participants' ability to detect future aircraft collisions (conflicts). Experiment 1 demonstrated that, across several set sizes, contrast and size cues to altitude improved accuracy at identifying conflicts. Experiment 2 demonstrated that graphical cues for representing altitude both improved accuracy and reduced search time for finding conflicts in large set size displays. The addition of size and contrast cues to ATC displays may offer specific benefits in aircraft conflict detection.

Categories and Subject Descriptors: H.5.2 [**Information Interfaces and Presentation**]: User Interfaces—*Graphical user interfaces (GUI), screen design, user-centered design*; H.1.2 [**Model and Principles**]: User/Machine Systems—*Human factors, human information processing, preformance*

General Terms: Design, Experimentation, Human Factors

Additional Key Words and Phrases: Human–computer interaction, air traffic control, applied cognitive science, visualization

ACM Reference Format:

Palmer, E. M., Clausner, T. C., and Kellman, P. J. 2008. Enhancing air traffic displays via perceptual cues. ACM Trans. Appl. Percpt. 5, 1, Article 4 (January 2008), 22 pages. DOI = 10.1145/1279640.1279644 http://doi.acm.org/10.1145/1279640.1279644

1. INTRODUCTION

Air traffic controllers maintain safe separation of aircraft flying within their assigned sector by monitoring altitudes and headings to ensure that no two aircraft are on a collision course. This is an attention-demanding task that requires the controller to continuously monitor the airspace, to simultaneously track multiple aircraft, and to continuously perform visual search for potential mid-air collisions [Benel

Authors' addresses: Current address, Evan M. Palmer, Department of Psychology, Wichita State University, Wichita, Kansas 67260; email: evan.palmer@wichita.edu; Current address, Timothy C. Clausner, University of Maryland Center for Advanced Study of Language, College Park, Maryland 20742; Philip J. Kellman, Department of Psychology, University of California, Los Angeles, Los Angeles 90095. Correspondence concerning this article should be addressed to Dr. Evan M. Palmer.

1998; Remington et al. 2000; Roske-Hofstrand and Murphy 1998]. In this paper, we address the third of these attentional tasks—we explore the potential benefits of adding the visual cues of size and contrast[1] to air traffic displays to aid the apprehension of aircraft altitude, and to help avoid traffic conflicts (and potential mid-air collisions).[2] We show that these display changes can improve participants' search for potential mid-air collisions in a simulated air traffic control (ATC) conflict-detection task.

The purpose of the present studies is to evaluate potential enhancements to ATC displays. Currently, controllers monitor air traffic using planview displays that are implicitly viewed from overhead. This display convention depicts aircraft latitude and longitude graphically, but provides no graphical display of aircraft altitude. Instead, air traffic controllers receive aircraft altitude information (along with other information, such as the aircraft's speed and flight number) from alphanumeric data tags associated with each aircraft icon in the display. Thus, in current ATC displays, the three dimensional (3D) position of each aircraft (i.e., latitude, longitude, and altitude) is communicated through two different channels of human visual processing: one two-dimensional (2D) graphical channel (longitude and latitude) and one alphanumeric channel (altitude). Holding other factors constant, this system, in which the altitude information for each aircraft must be read, may impose a limitation in the perceptual processing of pairwise aircraft trajectories and, therefore, the detection of aircraft conflicts. We believe that conflict detection may be improved through the graphical encoding of aircraft altitude.

There are several motivations for this effort, including both the limitations of the current scheme and the potential benefits of graphical encoding of altitude. The former include the fact that textual information can be processed only in a very small region of the visual field, including the 1–2 degrees of the fovea and perhaps a few degrees parafoveally (e.g., Anstis [1974]; Rayner, et al. [1980]). The unavailability of altitude information to peripheral vision may limit the effectiveness of the currently used data tags. Therefore, comparisons of altitude necessarily require serial fixations to specific display regions containing the alphanumeric data tags, creating a potential bottleneck in the controller's task.

Detecting aircraft conflicts is a uniquely difficult form of visual search because the target is not a single item, but rather a spatial relationship between two items. ATC displays, in their current form, often use aircraft icons that are all the same size and shape. In terms of search efficiency, there are no visual cues that can guide a controller's attention to aircraft of a particular altitude. Consequently, the search for aircraft conflicts may involve a serial, exhaustive examination through all pairs of aircraft that might be in conflict. If, on the other hand, observers could easily search through only aircraft of similar altitudes, then the efficiency of conflict detection might be greatly improved.

Previous work with ATC-like displays has shown that the addition of perceptual cues can improve situation awareness. Remington et al [2000] showed that coding aircraft altitudes via color differences improved conflict detection in an ATC task with retired controllers. Johnston et al. [1993] found that color coding of aircraft altitudes improved participants' ability to identify aircraft that were within 1000 feet of a target aircraft. The color-coding schemes in these studies improved performance, but neither study varied the presentation of aircraft icons in a manner consistent with monocular depth cues.

Other efforts have focused on improving situation awareness via cockpit displays of air traffic information. Ellis et al. [1987] employed perspective displays, where the air traffic situation was viewed from above and behind the pilot's aircraft. They evaluated the quantity and quality of evasive maneuvers

[1]We use the term "contrast" for this stimulus variable as a cover term, which might also be described as lightness, a perceptual variable that the observer extracts, or stimulus luminance, or grayscale in the computer's color table. All of these descriptions are highly correlated in our experimental manipulations.

[2]Even with increasing automation in the system, both current air traffic control technologies and those envisioned in the near future require that humans monitor air traffic information and detect potential conflicts.

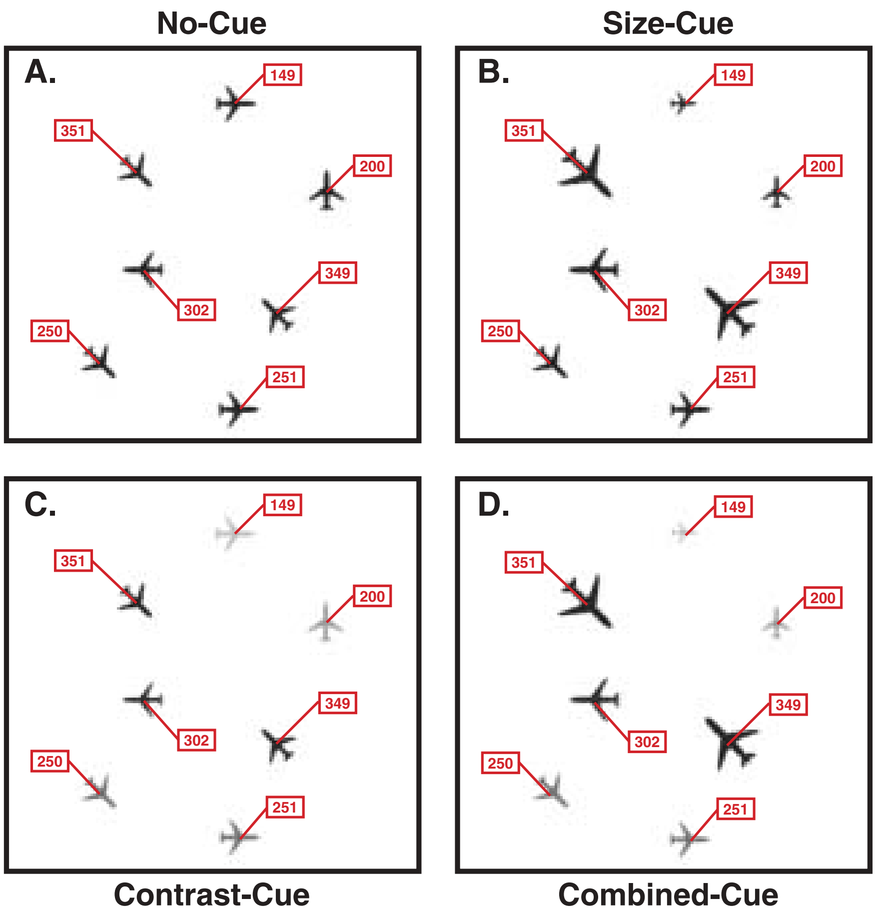

Fig. 1. Display conditions used in Experiments 1 and 2. Search displays represent an air traffic situation with altitude and flight path restrictions removed. (A) The no-cue condition: Altitude is conveyed solely by the data tags associated with each plane (altitude presented in hundreds of feet, see below). (B) The size-cue condition: Aircraft with higher altitudes are portrayed as being larger since they would be closer to the observer, given an overhead viewing position. (C) The contrast-cue condition: Aircraft at lower altitudes are farther away from the observer, so contrast is reduced. (D) The combined-cue condition: Size and contrast cues are used together to indicate altitude.

proposed by pilots in response to the detection of an aircraft conflict. Results showed that pilots tended to propose more ascending and descending evasive maneuvers with perspective displays, indicating that a more robust 3D mental representation of air traffic had been formed. Thus, there is evidence to suggest that the use of perspective cues in ATC displays can lead to more robust mental representations of air traffic, both for pilots and controllers.

In the present work, we consider two information visualization techniques for conveying aircraft altitude graphically (as in Figure 1) and evaluate their effect on the discovery of aircraft conflicts. While maintaining the use of data tags for conveying aircraft altitude, we varied the size and contrast of aircraft icons in a manner consistent with the monocular depth cues of relative size and contrast from an overhead viewing perspective of the air traffic situation. There is considerable evidence that relative size is an effective cue to relative depth, especially when stereoscopic and motion cues are absent (e.g., Hochberg and McAlister [1955]; Holway and Boring [1941]; Stenger et al. [1981]). Likewise, variations in the relative contrast of aircraft icons may also lead to the impression of depth segregation [Beck 1994]. The monocular depth cue of aerial perspective, first described by Leonardo Da Vinci [1492], predicts that if an object appears lower in contrast to the background than another, the lower contrast object will be perceived as farther away. Perceptual depth cues that correlate with aircraft altitude may allow for improved efficiency of conflict detection over the current system.

Variations in the size or contrast of aircraft icons can be easily incorporated into planview ATC displays and may be useful for conveying altitude in a more perceptually usable format. Application of these cues to ATC displays is straightforward. Assuming the controller has an overhead view of the airspace, icons would vary in size or contrast such that those at the highest altitudes are largest and darkest and those closest to the ground are smallest and lightest. In addition, the cues of size and contrast can be combined into a single icon and used simultaneously.

There are several reasons that graphically encoding altitude via aircraft icons might improve conflict detection. First, size or contrast cues that correlate with aircraft altitude might improve the efficiency of visual search for conflicts. By providing perceptual features that support attentional segregation of the displays according to altitude, attention may be more easily guided to the icons of interest Egeth et al. (1984); Wolfe (1994); Wolfe et al. (1989). For instance, Healy et al. [1996] successfully applied a similar technique to real-world quantitative data by using the cues of hue and orientation to aid in the rapid numerical estimation of data from salmon migration simulations. Second, variations in size, and contrast are available to peripheral vision and could lessen the need to foveate particular icons before determining that they are in a distinct altitude band. Third, size differences of similar shapes comprise the pictorial depth cue of relative size and contrast differences also function as pictorial depth cues. Any impression of depth in these displays, however modest, might allow attention to be allocated to different depth planes [He and Nakayama 1992] and possibly speed comparisons of aircraft at similar altitudes. Finally, increases in size and contrast may be easily understood as representing increases in altitude, possibly because conceptual magnitude is known to have conventionalized correspondence with increases in spatial height, such as the metaphor MORE IS UP Clausner [2002]; Clausner and Croft [1997, 1999]; Johnson [1987]; Lakoff [1993]; Lakoff and Johnson [1980].

Altitude information provided by size and contrast cues need not be as precise as that provided by data tags, nor would it replace them. The potential gain is that certain discriminations could be coded graphically (e.g., two aircraft at vastly different altitudes would appear as different sized icons). We hypothesized that the addition of size and contrast cues that had a consistent relationship with aircraft altitude would improve conflict detection and decrease response times (RTs) in an ATC-like task. The representation of aircraft altitude in terms of these visual cues should lessen the need to rely solely on alphanumeric data tags and thus decrease the time required to assess the altitude of each aircraft when searching for potential conflicts.

2. EXPERIMENT 1

We evaluated participants' ability to detect the presence of an aircraft conflict in simulated planview ATC displays under four different cue conditions. In all four conditions, each aircraft icon had an associated data tag indicating its altitude in alphanumeric form. In one condition, only data tags indicated altitude. In three conditions, altitude was indicated by data tags and one of three visual cues: size, contrast, or both size and contrast. In order to assess the merits of different altitude-encoding schemes fairly, we used naive participants rather than controllers, who are experienced with specific styles and details of displays in current use.

Examples of the experimental displays are shown in Figure 1. Often, ATC displays have structured flight paths and altitude bands such that aircraft proceed along defined routes and eastbound aircraft

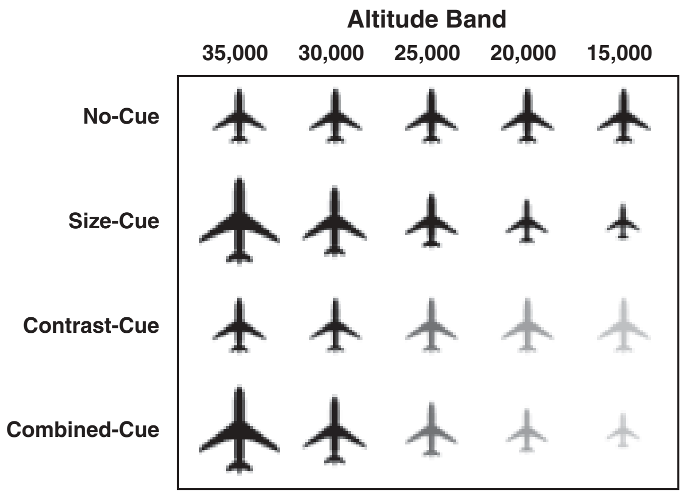

Fig. 2. Altitude band encoding using size and contrast. Altitudes are given in feet above mean sea level and refer to the middle altitude value within the band.

are at different altitudes from westbound aircraft. However, we relax these constraints consistent with current direct routings and possible future free-flight operations, in which aircraft may proceed off of established airways. Removing constraints on aircraft positions allowed us greater power and sensitivity in our experimental task. In addition, we were interested in evaluating observers' ability to detect aircraft conflicts as a function of angle of intersection, a variable that would be quite constrained using established airways. Consequently, the displays had route and altitude restrictions removed so that any two aircraft could potentially be in conflict.

Since alphanumeric data tags indicating altitude were always present in the experimental displays, size or contrast coding of aircraft altitude did not provide new information, but was (potentially) more accessible to the observer. In all conditions, perfect performance could be achieved by relying solely on the alphanumeric data tags for altitude information. Consequently, any improvement in conflict-detection performance for the cue conditions could be attributed to the presence of the cues themselves. We used novice observers rather than experienced controllers to avoid ceiling effects and to allow comparison of several display conditions using observers who were not already highly practiced in one of those conditions.

2.1 Methods

2.1.1 *Participants.* Forty University of California, Los Angeles undergraduates participated in the experiment in partial fulfillment of course requirements for an introductory psychology class. All participants reported normal or corrected-to-normal vision.

2.1.2 *Displays.* Displays designed as stimuli consisted of aircraft icons that appeared in one of five altitude bands, as depicted in Figure 2. The use of aircraft-shaped icons, as opposed to hexagons or some other shape, is similar to the methods of Ellis et al. [1987]. An aircraft always had a displayed altitude within 300 feet of its assigned altitude and, therefore, each altitude band may be considered a relatively stable feature of the aircraft being monitored (assuming it is in the cruise segment of flight).[3] In the

[3]These aspects of the displays, although somewhat simplified, reflected certain realistic aspects of air-traffic management situations. Controllers have responsibility for separating aircraft operating under instrument flight rules. These aircraft are assigned

size-cue, contrast-cue, and combined-cue conditions, the categorical nature of these altitude bands was depicted graphically as aircraft icons having uniform size, contrast, or both within a given altitude band.

In the no-cue condition (Figure 1A), aircraft icons were always presented at the same size and contrast, similar to the ATC displays in current use. In the size-cue condition (Figure 1B), aircraft icons varied in size such that those having the highest altitude were largest and those having the lowest altitude were smallest (consistent with the visual cue of relative size from an overhead vantage point). Aircraft at intermediate altitudes were depicted by icons of intermediate sizes, proportional to their altitudes. The contrast-cue condition (Figure 1C) was arranged similarly, with the highest altitude aircraft being darkest and the lowest altitude being lightest, consistent with the depth cue of aerial perspective from an overhead vantage point. The combined-cue condition (Figure 1D) employed both the size and contrast cues simultaneously.

We created 120 static displays, each depicting a hypothetical air traffic situation. At the observers' viewing distance of 24 inches, the full display area subtended 17.2 $\times$ 17.2 deg of visual angle. Five aircraft icon sizes and contrasts were used. Expressed in terms of the diameter of the smallest circle that could encompass the aircraft icon, the sizes were 0.87, 1.16, 1.45, 1.74, and 2.03 deg of visual angle. Expressed in terms of Michaelson contrast $((L_{MAX} - L_{MIN})/(L_{MAX} + L_{MIN}))$, with L_{MAX} equal to the white background and L_{MIN} equal to the luminance of the aircraft icon, the five contrast values were 0.32, 0.48, 0.65, 0.81, and 0.94. Each aircraft in the displays could be oriented in one of eight possible directions (0, 45, 90, 135, 180, 225, 270, or 315° from vertical) and was accompanied by a data tag that indicated the altitude of the aircraft in hundreds of feet (e.g., 25,000 feet was displayed as "250").

For each participant, the particular altitudes of the aircraft were randomly assigned to altitude bands. Consequently, two aircraft might have been in conflict because they had intersecting paths and occupied the same altitude band, but the particular altitude band that they occupied varied randomly from participant to participant. This method of randomization ensured that any differences observed in participants' abilities to detect aircraft conflicts could not be attributed to any particular altitude band (and consequently, not to any particular icon size or contrast value). In addition, several aircraft might occupy a single altitude band, but have slightly different altitudes. For instance, at the bottom of any one of the displays in Figure 1 there are two aircraft with data tags reading "250" and "251". These two aircraft occupy the same altitude band since they are within 300 vertical feet of each other (their altitudes are 25,000 and 25,100 feet, respectively).

We controlled the angle at which aircraft would intersect in these displays, given their present heading. Path intersections between two aircraft could form one of four angles: 45, 90, 135, or 180°. All angles appeared equally often in the 120 displays. In Figure 1, there are two sets of path intersections. In the middle of the display, the two aircraft labeled "351" and "349" are headed directly toward each other, forming a 180° path intersection. In addition, these two aircraft occupy the same altitude band indicating a conflict. In the upper right corner of the display, the two aircraft labeled "149" and "200" form a 90° path intersection; however these aircraft occupy two different altitude bands, and thus do not represent a conflict.

The displays always depicted either 2, 7, or 12 aircraft, with each set size occurring equally often among the set of 120 displays. One-half of the displays contained an aircraft conflict (conflict-present displays) and one-half did not (conflict-absent displays).

altitudes in some integral number of thousands of feet. Actual aircraft altitude may fluctuate somewhat, but the extent of these fluctuations is limited. A pilot whose aircraft strays 300 feet or more from its assigned altitude may be charged with an altitude deviation, often depending on whether the altitude error results in a loss of legally required separation between aircraft.

2.1.3 *Apparatus.* Stimuli were presented and responses recorded using an Apple Macintosh Power PC with a 17" (diagonal) monitor at a resolution of 1024×768 pixels. Participants were seated 24 inches from the monitor and responses were entered on designated keys on the keyboard. The experiment was programmed in the MacProbe programming language [Hunt 1994].

2.1.4 *Design.* Ten participants were run in each of the four cue conditions. Participants searched simulated ATC displays for aircraft conflicts. We defined an aircraft conflict as any situation in which two aircraft occupied the same altitude band and were on intersecting paths. This definition is simpler than the FAA definition of an aircraft conflict, but it retains the basic premise. With the simplifying assumption that all aircraft were traveling at the same speed, a path intersection was defined as a situation in which two aircraft would occupy the same horizontal and vertical location should they continue on their present course.

All participants were tested on the same set of 120 displays. These stimuli were presented in random order, one display per trial. Sixty conflict-present trials were generated by using all combinations of 5 (altitudes) $\times$ 3 (set sizes) $\times$ 4 (intersection angles). For both conflict-present and conflict-absent displays, the presence of path intersections between two aircraft was counterbalanced across the four angles. The 60 conflict-absent trials were divided into three sets of 20 trials each for the 2, 7, and 12 plane conditions. Within each 20-display set, 15 displays depicted a situation in which two aircraft were on an intersecting path but had different altitudes and five displays depicted a situation in which several aircraft shared the same altitude, but had no path intersection. For the 7 and 12 plane conditions, aircraft not forming a path intersection were assigned to random locations, orientations, and altitudes with the constraint that their placement did not form a new path intersection in the display.

2.1.5 *Procedure.* Participants read a series of instruction screens at the beginning of the experiment to acquaint them with the conflict detection task and the displays, particularly the convention in which data tags report altitude in hundreds of feet. If the participant was in the size-cue condition, they were instructed that "The different-sized aircraft reflect their different altitudes. Larger aircraft are closer to you and smaller aircraft are farther away from you." Participants in the contrast-cue condition were instructed, "The different shades of the aircraft reflect their different altitudes. Darker aircraft are closer to you and lighter aircraft are farther away from you." Participants in the combined-cue condition saw both descriptions. Participants were instructed to press the "C" key on the keyboard if they detected a conflict, or the "N" key on the keyboard if they did not detect a conflict.

The participants completed a practice phase of 60 trials during which they received feedback about their performance. The procedure for the experimental phase was identical to the practice phase except that no feedback was given. This phase consisted of 120 trials and took approximately 30 min. Each participant was given a short break after 60 trials. Following the experimental phase, participants were debriefed about the purposes of the experiment.

2.2 Results

2.2.1 *Accuracy Data.* Overall performance in the four cue conditions as a function of set size is shown in Figure 3. Participants were more accurate at identifying conflicts in displays with added visual cues than in displays lacking them. The presence of size cues in particular enhanced performance, with the combined-cue condition showing the best performance.

These patterns were confirmed by the analyses. Conflict-detection accuracy was analyzed in a $4 \times 2 \times 4 \times 3$ (cue type $\times$ conflict presence/absence $\times$ angle of intersection $\times$ set size) mixed ANOVA, with the first factor treated as a between-subjects variable. The analysis revealed main effects of cue type, $F(3, 36) = 9.27$, $p < .0005$, conflict presence, $F(1, 36) = 48.09$, $p < .0001$, angle of intersection, $F(3, 108) = 22.38$, $p < .0001$, and set size, $F(2, 72) = 98.98$, $p < .0001$. There were also significant interactions

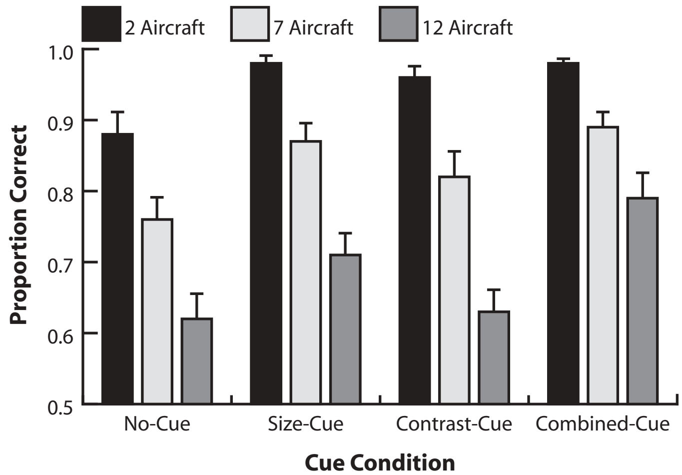

Fig. 3. Conflict detection accuracy as a function of cue condition and number of aircraft in the display.

between set size and conflict presence/absence, $F(2, 72) = 27.96$, $p < .0001$, conflict presence/absence and angle of intersection, $F(3, 108) = 12.01$, $p < .0001$, number of aircraft and angle of intersection, $F(6, 216) = 6.64$, $p < .0001$, and a three-way interaction between conflict presence/absence, set size, and angle of intersection, $F(6, 216) = 5.08$, $p < .0001$.

The addition of size cues to the displays improved performance, while the combination of both size and contrast together appeared to be the most potent. The t-tests on the cue conditions (with a Bonferroni correction against type-I error, $p_{crit} = 0.0083$) showed that participants in the no-cue condition had lower conflict-detection accuracy, overall, than participants in either the size-cue or combined-cue conditions, all $t(18) \geq 3.45$, all $p \leq 0.003$[4]. In addition, participants in the combined cue condition outperformed those in the contrast-cue condition, $t(18) = 3.39$, $p = 0.003$.

Overall, the proportion of correct trials decreased as set size increased. The angle of intersection between aircraft also had a significant impact on participants' accuracy. The t-tests between the angle of intersection conditions revealed that participants were more sensitive to 180° path intersections than to any other angle of intersection, all $t(39) \geq 6.18$, all $p < 0.0001$. There were no other reliable differences between the angle of intersection conditions, relative to the Bonferroni-corrected alpha level of 0.0083 (all $p > 0.05$, n.s.).

Participants were best at identifying conflicts when path intersections formed a 180° angle, but this advantage was obscured in the two-plane condition because of apparent performance at ceiling. The interaction of number of aircraft by angle of intersection was driven by uniformly high performance in the two-plane condition combined with the main effect of angle of intersection noted earlier.

There were no other reliable main effects or interactions in the proportion correct data (all $p > 0.10$).

2.2.2 *Response Time Data.* The addition of visual cues to simulated ATC displays did not have a significant impact on observers' mean RTs (Figure 4). However, more aircraft in the displays yielded longer RTs and participants responded faster to displays containing conflicts than to those that did not. The RT advantage of conflict-present displays increased along with the number of aircraft. Finally, participants were fastest when aircraft formed a 180° angle of path intersection.

[4]We use the term "all" when several planned comparisons were used in an analysis and more than one was significant. The test with the least degree of significance is reported. In addition, all t-tests were two-tailed.

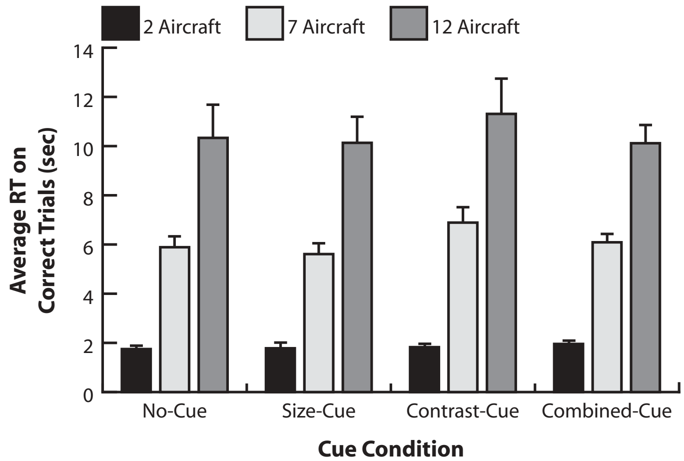

Fig. 4. Average response time on correct trials as a function of cue condition and set size in Experiment 1.

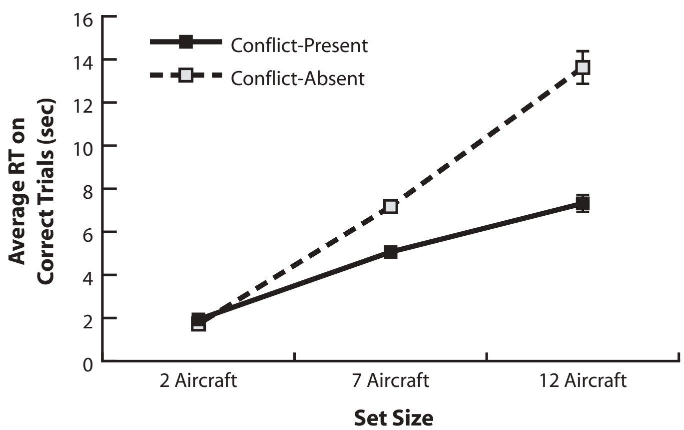

Fig. 5. Average response time on correct trials as a function of set size and conflict presence/absence for Experiment 1.

Response time data (correct trials only) were analyzed with a 4 × 2 × 4 × 3 (cue type × conflict presence/absence × angle of intersection × set size) mixed ANOVA, with the first factor as a between-subjects variable. The ANOVA revealed reliable main effects of conflict presence/absence, $F(1, 27) = 111.77$, $p < 0.0001$, angle of intersection, $F(3, 81) = 4.15$, $p < 0.01$, and set size, $F(2, 54) = 222.20$, $p < 0.0001$, as well as a significant interaction of set size by conflict presence/absence, $F(2, 54) = 69.58$, $p < 0.0001$.

The analysis showed no reliable main effect of cue type, $F(3, 27) = 1.92$, $p = 0.15$, although participants in the no-cue condition had the numerically shortest RTs and those in the combined cue condition had the longest (Figure 4). Participants' enhanced conflict-detection accuracy in the size-cue, contrast-cue, and combined-cue conditions came with no statistically reliable increase in RTs, compared to the no-cue condition.

The presence of a conflict interacted with the number of aircraft in the display, $F(2, 54) = 69.58$, $p < 0.0001$ (Figure 5), reflecting the higher RT × set size slope for conflict-absent displays. The conflict-absent RT slope (1.13 s/aircraft) was roughly twice as large as the conflict-present slope

(0.53 s/aircraft). Planned comparisons between the RT × set size slopes for conflict-present and conflict-absent displays between the cue conditions did not reveal any significant differences between conditions (all $p > 0.10$).

Finally, the main effect of angle of intersection was driven by faster RTs to aircraft forming a 180° angle of intersection than to any other angle. Planned comparisons between the four angle of intersection conditions confirmed the superiority of the 180° condition, all $t(39) \geq 2.20$, all $p < 0.05$.

There were no other reliable main effects or interactions in the RT data (all $p > 0.10$).

2.3 Discussion

The graphical encoding of altitude using size and contrast variations produced significant gains in conflict-detection by our participants. We can roughly quantify the improvement in performance in terms of the equivalent reduction of display complexity. Inspection of Figure 3 reveals that adding both size and contrast cues to displays consistently yielded improvements equivalent to reducing the complexity of the displays by about five aircraft. Performance on 12-aircraft displays in the combined-cue condition was nearly equivalent to performance on seven aircraft displays in the no-cue condition. Likewise, performance on seven-aircraft displays in the combined-cue condition produced a level of performance equivalent to the two-aircraft displays in the no-cue condition.

2.3.1 *Interaction of Cues.* If size and contrast cues each enhance conflict-detection performance, how do they interact when they are both present in the displays? In the combined-cue condition, the relative size and relative contrast of the aircraft icons were correlated to determine whether participants would benefit from having more than one visual cue for altitude. We might expect some improvement in the combined-cue condition over either single-cue condition alone because of simple probability summation. If either cue alone increases the probability of detection, the joint presence of two cues predicts even better detection, depending on the independence of the two cues.

In this experimental scenario, the probability summation equation must be expressed in terms of the proportional reduction in error for the size-cue, contrast-cue and combined-cue conditions over the no-cue condition. For example, the proportional reduction in error for the size-cue condition ($R_{Size\text{-}Cue}$) is expressed in Eq. (1):

$$R_{Size\text{-}Cue} = \frac{(1 - P_{No\text{-}Cue}) - (1 - P_{Size\text{-}Cue})}{(1 - P_{No\text{-}Cue})} \quad (1)$$

Where $P_{No\text{-}Cue}$ and $P_{Size\text{-}Cue}$ are the proportion correct for the no-cue and size-cue conditions, respectively. The term $R_{Size\text{-}Cue}$ represents the effectiveness of the size-cue, that is, the chance that the cue reduced errors compared to the no-cue condition. The proportional reduction in error for the other cue conditions over the no-cue condition is calculated similarly.

Using the results of Eq. (1), the predicted reduction in error for the combined-cue over the no-cue condition can be expressed as:

$$1 - R_{Combined-Cue} = (1 - R_{Size\text{-}Cue})(1 - R_{Contrast-Cue}) \quad (2)$$

This revised probability summation equation predicts a 0.53 proportional reduction in error for the combined-cue condition, if the size and contrast cues were additive in their benefit to the participants. We observed a 0.56 proportional reduction in error in the combined-cue condition which, given the similarity of these two numbers, indicates that the benefits of the size and contrast cues were each providing an independent benefit in Experiment 1.

People perform as well or better in the combined-cue condition than predicted by probability summation. This indicates that the two cues of size and contrast may be operating independently and additively to help the detection of conflicts in the displays.

2.3.2 *Clues About Processing.* Participants were more accurate and faster at identifying conflicts forming a 180° angle of path intersection than any other angle. This advantage may indicate that participants searched for conflicts by scanning along the future trajectory of aircraft icons. This strategy would provide a selective advantage for 180° path intersections because the future trajectory of one aircraft icon would lead directly to another aircraft icon in the display.

Though consistent with some previous work [Ellis et al. 1987], the finding that observers were better able to detect conflicts with larger angles of intersection stands in contrast to other findings [Ellis and Palmer, 1981; Remington, et al. 2000; Smith et al. 1984]. The Remington et al. [2000] study used retired air traffic controllers as participants and attributed their finding to a proximity effect. Pairs of icons with small angles of intersection are closer in absolute space than pairs with large angles of intersection. Remington et al. [2000] notes that controllers are trained to search in a clockwise direction from each icon in order to find possible conflicts. It is possible that the difference in experimental outcomes is because of a difference in conflict detection training and techniques.

For the response time data, the best-fitting linear functions relating RT to number of aircraft showed that slopes on target-absent trials were 2.34 times that of target-present trials (Figure 5). This slope ratio of slightly more than 2:1 is a familiar pattern in visual search data [Wolfe 1998], sometimes interpreted as arising from a serial self-terminating search of the displays in which search is halted upon discovery of a target (here, a conflict), and exhaustive otherwise (e.g., Treisman and Gelade [1980]). However, searching for a conflict in these displays required inspecting pairs of items, so an exhaustive serial search would be expected to lead to RTs that increased as a quadratic function of set size. Given that the RT × set size data did not increase quadratically, it appears that search was slow and inefficient, but apparently not exhaustively serial. Alternatively, search of pairs might not be the correct metric; for example, search might proceed across several areas or strips of a display, with time per area varying with aircraft density. Such a model would also be consistent with the data. The search slopes indicate that participants spent 539 ms/item on conflict-present displays and 1190 ms/item on conflict-absent displays, on average. These numbers suggest that the search task was very difficult, because average search slopes in typical visual search experiments are 14.6 ms/item for target-present displays and 33.0 ms/items in target absent displays [Wolfe, 1998]. It is likely that the aircraft icons were each checked several times during the search for conflicts. Whether more efficient processing could be attained with training is a question for further research.

The graphical cues in this experiment may have allowed participants to set up visual categories (corresponding to the altitude bands) that could have guided their search for a conflict. Prior to the experiment, we hypothesized that the benefit of such coding might be in the rapid exclusion of certain path intersections that were not conflicts because of altitude differences in the converging aircraft. A different mode of improvement appears likely, given that our data showed enhanced accuracy but no differences in RT. Participants' attention may have been drawn to some conflicts that would otherwise have been missed. This would account for improvement in accuracy, but not speed. It remains possible that improvement in rapid exclusion of some path intersections might be seen with extended practice.

The data suggest that encoding altitude as size, contrast, or both offers clear advantages over data tags alone for observers' processing relevant relations in displays. There are several reasons why these cues might have been beneficial. First, the use of size and contrast coding might allow the observer to efficiently guide their attention to the relevant altitude categories in their search for conflicts. Second, size and contrast information, to a greater extent than the numbers in data tags, should be available in peripheral vision. Finally, size and contrast variations may be conceptually compatible with mental representations of magnitude.

Experiment 1's results did not provide much evidence for improved search efficiency as a function of cue type. Both enhanced attentional processing and greater contributions of information from peripheral vision may have contributed to improved detection of conflicts that would otherwise have been missed. The finding of additive contributions of size and contrast is consistent with this possibility. Compatibility with conceptual notions of magnitude also remains a possible contributor to the performance improvements we observed.

3. EXPERIMENT 2

Although Experiment 1 supported the notion that adding perceptual cues for altitude can enhance conflict detection, there was a limitation in its design. Displays in Experiment 1 always had, at most, one path intersection. This feature ensured comparability of the conflict-present and conflict-absent displays. In conflict-present displays, the path intersection involved aircraft in the same altitude band, whereas in conflict-absent displays, it did not. This is not representative of air traffic displays, since several path intersections may be present in air traffic at once, despite the fact that no two aircraft are in conflict. Moreover, the presence, of at most, one path intersection per display could have allowed participants to determine the absence of any conflict using a shortcut strategy. Specifically, in a conflict-absent display, participants could find a path intersection, then determine that the aircraft were at two different altitudes, and stop searching. Despite the availability of this shortcut strategy, the RT data from Experiment 1 suggested a pattern of extensive serial search (i.e., participants on average spent a longer time looking at conflict-absent than conflict-present displays). This result suggests that the shortcut strategy was not used by many participants, if any. Nonetheless, in the absence of a detailed model of participants' processing in our task, we cannot evaluate the importance of the shortcut strategy in Experiment 1.

Consequently, in Experiment 2 we manipulated the number of path intersections in our displays. New displays were developed that contained one, two, or three path intersections. The presence of multiple path intersections in most of the displays eliminated the effectiveness of the shortcut strategy, because a decision about the absence of a conflict in the displays could not be made from evaluating one path intersection alone.

A second goal of Experiment 2 was to obtain more detailed information on the use of peripheral vision in performance on our task. To do so we systematically manipulated the distance to path intersection between conflicting aircraft in the Experiment 2 displays. We were interested in determining the effectiveness of the size and contrast cues as a function of distance between conflicting aircraft. We hypothesized that the size-cue and contrast-cue manipulations would improve participants' ability to detect aircraft conflicts over large distances, compared to the no-cue condition.

3.1 Methods

3.1.1 *Participants.* Fifty-four University of California, Los Angeles undergraduates participated in the study in partial fulfillment of course requirements in an introductory Psychology class. Six HRL Laboratories employees participated in the study on a volunteer basis. All participants reported normal or corrected-to-normal vision.

3.1.2 *Displays.* Displays designed as stimuli for this experiment had the same general appearance as the displays in Experiment 1, with the following exceptions: First, 12 aircraft were depicted in every display. Second, all the displays contained one, two, or three path intersections. In one-half of the displays, one (and only one) path intersection formed an aircraft conflict. Finally, for all sets of path intersections within the displays, the distance between the center of each aircraft icon and the point of intersection was randomly assigned to be one of five specific distances.

3.1.3 *Apparatus and Materials.* Stimuli were presented and responses recorded using the same computer screen and apparatus specified in Experiment 1.

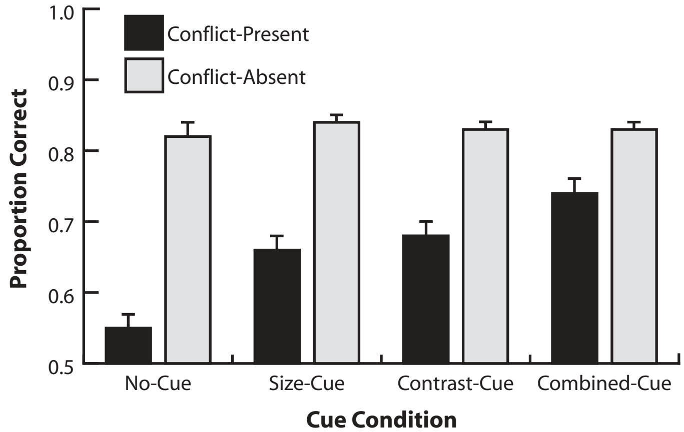

Fig. 6. Conflict-detection accuracy as a function of cue condition and conflict presence/absence in Experiment 2.

3.1.4 *Design.* Fifteen participants were run in each of the four cue conditions. All participants were tested on the same 120 displays, with the icon presentation style being determined by the cue condition. The 120 displays were created by combining the factors: 2 (conflict presence/absence) × 5 (altitudes) × 3 (numbers of path intersections) × 4 (intersection angles). For both conflict-present and conflict-absent trials, path intersections between two aircraft were counterbalanced across four angles: 45, 90, 135, or 180°. For each conflict-present display, the distance between the center of the two conflicting aircraft icons to the point of path intersection was chosen randomly to be either 1.79, 3.43, 5.06, 6.98, or 8.59 deg of visual angle.

3.1.5 *Procedure.* The same instructions used in Experiment 1 were used in this experiment. The practice phase of the experiment consisted of 20 displays, with 12 aircraft in each display. One-half of the training scenarios had conflicts. We reduced the number of training displays from 60 to 20 to allow participants more time to complete the experimental phase, which took longer than Experiment 1, because every display contained 12 aircraft. Participants were given a short break after 60 experimental trials and all were able to complete the experiment in less than 1 hr.

3.2 Results

3.2.1 *Accuracy Data.* Figure 6 shows accuracy by cue type. Conflict detection was enhanced by the perceptual coding of altitude. The size-cue, contrast-cue, and combined-cue conditions all produced conflict detection performance that was superior to performance in the no-cue condition. Number of path intersections had no reliable impact on participants' conflict-detection accuracy or RTs.

We analyzed the accuracy data using a 4 × 2 × 4 × 3 (cue type × conflict presence/absence × angle of intersection × number of path intersections) mixed ANOVA, treating the first factor as a between-subjects variable and the latter three factors as repeated measures. The analyses revealed main effects of cue type, $F(3, 56) = 7.16$, $p < 0.0005$, conflict presence/absence, $F(1, 56) = 57.94$, $p < 0.0001$, and angle of intersection, $F(3, 168) = 25.57$, $p < 0.0001$. The analyses also detected significant interactions of conflict presence/absence by angle of intersection, $F(3, 168) = 20.65$, $p < 0.0001$, and number of paths by angle of intersection, $F(6, 336) = 16.41$, $p < 0.0001$, as well as a three-way interaction of conflict presence/absence by number of path intersections by angle of intersection, $F(6, 336) = 7.80$, $p < .0001$.

Comparisons of the accuracy scores for the four cue conditions via t-tests (with a Bonferroni correction against type-I error, $p_{crit} = 0.0083$) revealed that participants in the no-cue condition were significantly

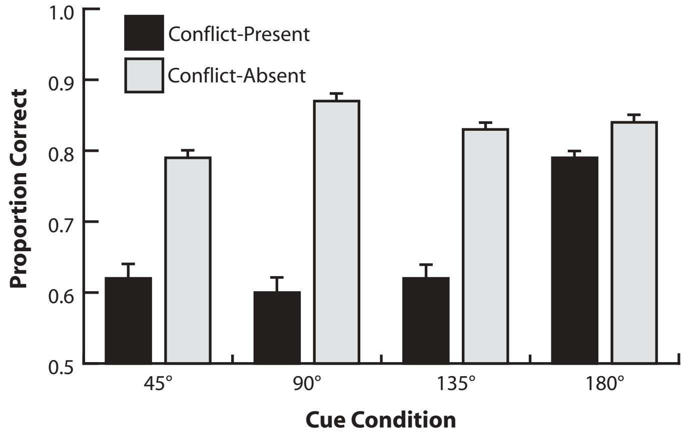

Fig. 7. Accuracy to aircraft conflicts as a function of angle of path intersection and conflict presence/absence in Experiment 2.

less accurate at identifying aircraft conflicts than were participants in the contrast-cue or combined-cue conditions, all $t(28) \geq 3.24$, all $p < 0.0083$. No other comparisons reached significance (all $p > 0.0083$, n.s.).

Inspection of Figure 6 indicates that the main effect of cue condition was driven by improved performance on conflict-present trials for displays with size and/or contrast cues, however the interaction of cue type by conflict presence/absence was only marginally significant, $p = .072$, n.s. Figure 6 also indicates that participants were biased toward responding "no-conflict" more often than "conflict," since performance on conflict-absent displays was consistently higher than performance on conflict-present displays. However, despite this bias, participants were about 20% more likely to detect the presence of a conflict in the combined-cue compared to the no-cue condition.

As in Experiment 1, participants were much better at detecting 180° angle of intersection conflicts than any other angle of intersection. The superiority of the 180° angle of intersection condition compared to other conditions was confirmed by t-tests with a Bonferroni correction, all $t(59) \geq 6.39$, all $p < .0001$ (Figure 7). No other t-test reached significance, all $t(59) \leq 1.98$, all $p \geq 0.05$. The ANOVA also revealed that participants' conflict-detection accuracy as a function of angle of intersection was modulated by the number of path intersections in the displays. This interaction appears to be because of consistently superior performance in the 180° condition across all numbers of path intersections, in contrast to varying levels of performance for the other angles of intersection. Likewise, the significant interaction of conflict presence/absence by angle of intersection was driven by uniformly high performance on conflict-absent displays for all angles of intersection, as opposed to higher performance on conflict present displays for only 180° intersections (Figure 7).

The three-way interaction of conflict presence/absence by number of path intersections by angle of intersection appears to be driven mostly by the main effects of conflict presence/absence and angle of intersection. The number of path intersections manipulation elicited quite variable performance as a function of angle of intersection and do not appear to be contributing anything meaningful to the three-way interaction.

Analyses of the accuracy data revealed no other significant main effects or interactions (all $p > 0.05$). Of particular interest is the fact that the number of path intersections in the display did not have a significant impact on participants' accuracy at identifying aircraft conflicts, indicating that the findings of Experiment 1 were probably not compromised by participants using the shortcut strategy described above.

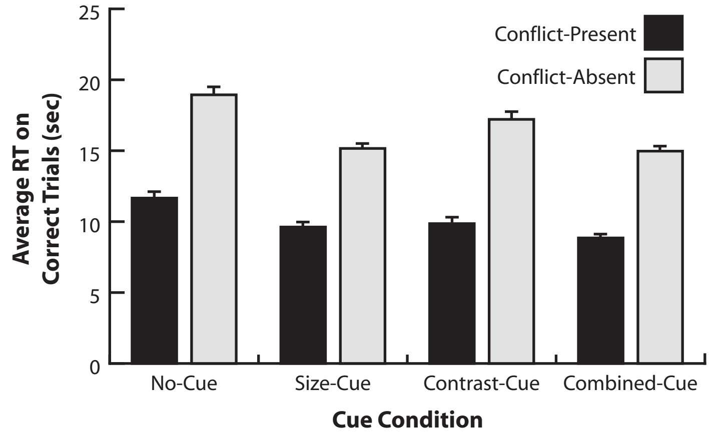

Fig. 8. Average response time on correct trials as a function of cue condition and target presence/absence in Experiment 2.

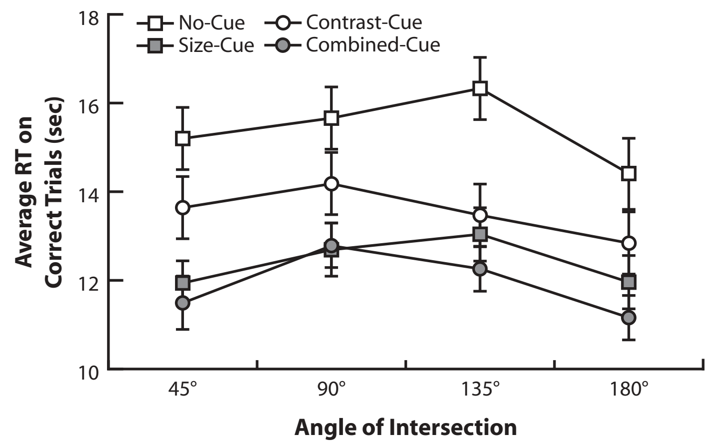

Fig. 9. Average response time on correct trials as a function of cue condition and angle of path intersection in Experiment 2.

3.2.2 *Response Time Data.* The average RT data are depicted in Figures 8 and 9. Response times in the no-cue condition were slower, on average, than RTs in the size-cue or combined-cue conditions. In addition, participants responded fastest to displays with conflicts and to displays with aircraft forming a 180° angle of intersection.

These observations were confirmed by the analyses. Response time data for correct trials were analyzed with a $4 \times 4 \times 3 \times 2$ (cue type $\times$ angle of intersection $\times$ number of path intersections $\times$ conflict presence/absence) mixed ANOVA, with the first factor as a between-subjects variable and the latter three factors as repeated measures. The analysis revealed main effects of cue condition, $F(3, 47) = 4.35$, $p < 0.01$, conflict presence/absence, $F(1, 47) = 255.95$, $p < 0.0001$, and angle of intersection, $F(3, 141) = 7.88$, $p < 0.0001$. The ANOVA also detected significant two-way interactions of conflict presence/absence by cue condition, $F(3, 47) = 3.14$, $p < 0.05$, conflict presence/absence by number of path intersections, $F(2, 94) = 7.38$, $p < 0.005$, and conflict presence/absence by angle of intersection, $F(3, 141) = 16.22$, $p < 0.0001$. In addition, there were significant three-way interactions of conflict presence/absence by angle by cue condition, $F(9, 141) = 2.83$, $p < 0.005$, and conflict presence/absence by number of path intersections by angle, $F(6, 282) = 2.63$, $p < 0.05$.

Overall, the combination of both size and contrast cues in the displays significantly reduced participants' average time to identify the presence or absence of a conflict. Planned comparisons between the four cue conditions revealed that participants in the no-cue condition were slower to respond than participants in the combined-cue condition, $t(28) = 2.93$, $p < 0.01$. Response times in the size-cue condition were faster than those in the no-cue condition, though this difference was only marginally reliable, $t(28) = 2.20$, $p = 0.046$, n.s., relative to the Bonferroni corrected significance level of $p = 0.0083$. Response times in the contrast-cue condition, though numerically faster, were not reliably different from RTs in the no-cue condition ($p > 0.10$, n.s.).

The main effect of conflict presence/absence was because of the fact that participants responded 6.5 s faster, on average, to conflict-present displays than to conflict-absent displays. This substantial difference in RTs is likely because of the large number of aircraft used in this experiment. When participants find a conflict among the 12 aircraft in these displays, on some trials they may find it at the beginning of their search, and on other trials they may not find it until the end of their search. Presuming that participants have to search only one-half of a display, on average, to find a conflict, we can estimate that conflict-absent displays involve inspection of about six more aircraft icons than conflict-present displays. If participants take roughly 1 s to inspect each icon in the display (as suggested by Experiment 1, see Figure 5), then this might yield the 6.5-s difference in RTs that we observed.

The angle formed between two aircraft in a path intersection had a significant impact on participants' RTs (Figure 9). Participants were fastest to respond to displays that involved aircraft heading on a direct collision course (the 180° condition), or involved aircraft that were fairly close to each other (the 45° condition). The 180° condition was reliably faster than performance in both the 90° condition, $t(59) = 4.72$, $p < 0.0001$, and the 135° condition, $t(59) = 4.73$, $p < 0.0001$. The 45° angle of intersection was reliably faster than the 90° condition, $t(59) = 2.89$, $p = 0.0028$, but the difference between the 45° and 135° conditions was only marginally significant, relative to a Bonferroni corrected alpha value of 0.0083, $t(59) = 2.49$, $p = 0.0085$, n.s.

Conflict presence or absence interacted with cue type, number of path intersections, and angle of intersection. The difference between conflict-present and conflict-absent displays was larger for the no-cue and contrast-cue conditions than for the size-cue and combined-cue conditions. Likewise, the presence of a conflict had a larger effect on accuracy for displays with three path intersections than for the other displays. Finally, RTs for 180° conflict-present displays were much shorter than any other condition, causing a significant interaction.

The three-way interaction of conflict presence/absence by angle by cue condition was as a result of shorter RTs for 180° conflict-present displays, particularly for participants in the no-cue condition. The interaction of conflict presence/absence by number of path intersections by angle was because of relatively short RTs in 180° conflict-present displays with three path intersections.

3.2.3 *Distance to Conflict Data.* Finally, we conducted separate statistical analyses of performance as a function of distance to path intersection. For these analyses, we grouped participants' accuracy and RTs on conflict-present trials according to the five distances to path intersection by collapsing across all numbers and angles of path intersections for a given distance. Since many displays had multiple path intersections, only the distances for path intersections forming a conflict were used in this analysis. For the RT data, only correct trials on conflict-present displays for each distance were entered into the analysis.

Regardless of the distance between icons that were path-intersecting, participants in the conditions with visual cues were better at identifying conflicts than were participants in the no-cue condition

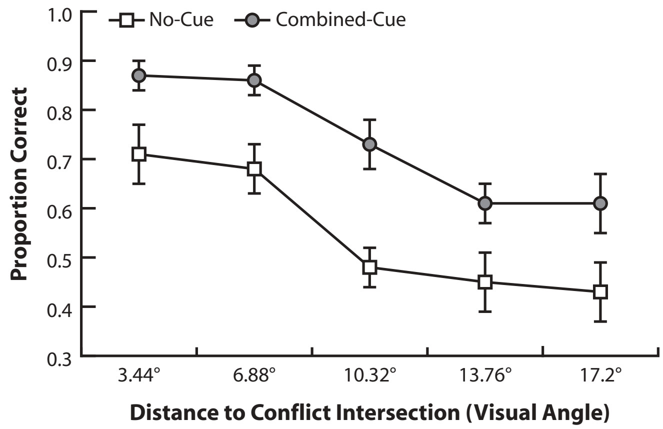

Fig. 10. Conflict detection accuracy as a function of distance to path intersection for the no-cue and combined-cue conditions in Experiment 2. Performance curves for the size-cue and contrast-cue conditions fell between the no-cue and combined-cue conditions and have been omitted to improve readability of the graph.

(Figure 10). The accuracy data were submitted to a 4×5 (cue type $\times$ distance) mixed ANOVA with the first factor as a between-subjects variable; the second factor as a repeated-measures variable. Similar to the previous findings, we again observed a main effect of cue type, $F(3, 56) = 5.87$, $p < 0.005$. We also observed a main effect of distance to conflict $F(4, 224) = 57.65$, $p < 0.0001$, with accuracy at identifying aircraft conflicts steadily declining as the distance between the aircraft and the point of intersection increased (Figure 10). There was no interaction of cue condition $\times$ distance ($p > 0.10$).

The response time data were analyzed with a 4×5 (cue condition $\times$ distance) mixed ANOVA, with the first factor as a between-subjects variable and the second factor as a repeated-measures variable. Unlike the RT analysis reported above, this analysis did not show a reliable main effect of cue condition, $F(3, 56) = 2.10$, $p > 0.10$. The lack of a main effect is likely because of the inclusion of only conflict-present trials in our analyses since conflict-absent RTs varied more as a function of cue type than conflict-present trials. We did observe a main effect of distance to path intersection, $F(4, 224) = 21.04$, $p < 0.0001$, with RTs increasing steadily as a function of distance (Figure 11). The interaction of cue type by distance was not significant ($p > 0.10$). These results suggest that the relative performance benefits from size and contrast cues did not vary as a function of distance to path intersection.

3.3 Discussion

The results of Experiment 2 confirm and extend those of Experiment 1. Both experiments indicate that perceptual encoding of altitude information confers advantages in accuracy at extracting information from ATC-like displays. In addition, Experiment 2, with only high-density displays, showed clear benefits of such encoding for detection speed. The accuracy and RT advantages in the cue conditions held over the full range of distances tested (up to about 17 deg separation between conflicting aircraft) and were not found to interact with distance. These results suggest that enhanced information extraction using peripheral vision may contribute to the performance advantages. Although there was no increase in benefit with increasing separation, it is not clear that one would be expected. Even our smallest separation between conflicting aircraft (1.8 deg, for the 45° intersection condition at the shortest distance) makes it unlikely that both aircraft cannot be entirely in the fovea simultaneously (as the fovea is 1–2 deg in diameter).

Experiment 2 again provided evidence that the combination of size and contrast cues yields performance advantages greater than either cue alone. Relative to the no-cue condition, accuracy in the

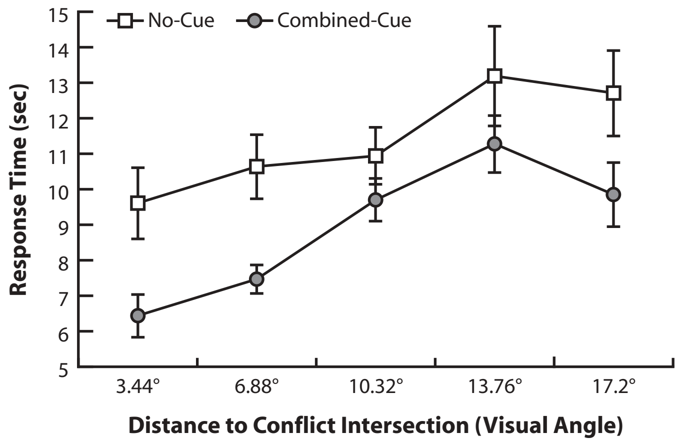

Fig. 11. Average response times on correct trials as a function distance to path intersection for the no-cue and combined-cue conditions in Experiment 2. Performance curves for the size-cue and contrast-cue conditions fell between the no-cue and combined-cue conditions, and have been omitted to improve readability of the graph.

combined condition increased by 10%, from 68.5 to 78.5%. Using the probability summation formulas and rationale presented in the discussion of Experiment 1, the expected proportional reduction in error for the combined-cue condition was 0.37, while the actual proportional reduction in error was 0.28. This may suggest that the size and contrast cues, when combined, did not function completely independently in this experiment. Nonetheless, it is still clear that the two cues together are more beneficial than either cue alone.

Response times in the cue conditions were more than 2 s shorter than RTs in the no-cue condition. We believe the lack of a main effect of cue condition for RTs in Experiment 1 combined with significantly faster RTs in Experiment 2 argues against the notion of any speed-accuracy tradeoff in these displays. The best response times were found in the combined-cue condition. The use of only high-density displays with many possible path intersections in this experiment (i.e., 12 aircraft having up to three path intersections) appeared to particularly favor the use of size and contrast cues for encoding altitude. Burnett and Barfield [1991] surveyed air traffic controllers, who reported that they are more likely to rely on data tag information in high-density displays, because they provide the most accurate altitude information. We believe that the size and contrast coding of altitude in our displays reduced the need for observers to rely solely on data tags for altitude information when searching for conflicts. The net effect of the altitude encoding is that participants had a greater ability to perform a visual search for conflicts. However, the effectiveness of these cues with ascending or descending aircraft has yet to be evaluated.

Improved visual search in the cue conditions may result either from participants quickly and accurately identifying aircraft at the same altitudes for analysis, or from participants quickly rejecting aircraft that are clearly at different altitudes from each other. In either case, the addition of size and contrast cues to displays may be of particular benefit to air traffic controllers monitoring high-density airspaces with many potential path intersections, such as in departure or approach control.

The interaction of conflict presence/absence with cue condition suggests that when participants must exhaustively search a display (as in a conflict-absent display), the addition of size and contrast cues for altitude encoding can significantly reduce the time and increase the accuracy of the search. This may be because the addition of these cues allows more efficient attentional selection of relevant aircraft and/or greater use of peripheral visual information. The results are consistent with the interpretation that participants exclude some path intersections from detailed consideration.

On the other hand, average RTs in this experiment were 11.9 s or more, which is enough time for participants to inspect each aircraft icon for roughly 1 s before making a decision about the presence or absence of a conflict. There is little indication that this type of processing is reduced, even in the cue conditions. The addition of size and contrast cues in these experiments may not change the visual search strategy, but instead may improve its efficiency. A study of the eye movements executed by participants when searching for conflicts in these displays could provide further information on this issue.

4. GENERAL DISCUSSION

Taken together, Experiments 1 and 2 provide evidence that extraction of information from ATC-type displays can be enhanced by size and contrast cues for the encoding of altitude. The beneficial effects of these cues were apparent in accuracy (Experiments 1 and 2) and in response times (Experiment 2). Increases in conflict detection accuracy were observed across all display densities (i.e., 2, 7, or 12 aircraft) for the combined-cue condition, relative to the no-cue baseline. The size-cue and contrast-cue conditions showed an intermediate level of improvement, which was sometimes, but not always, reliably better than the no-cue condition. The advantage in response time to detect conflicts in displays with size and contrast cues was apparent in the high-density displays of Experiment 2. The benefits of using graphical representations for altitude were evidenced by performance in both low- and high-density displays, and across both short and long separation distances between aircraft.

What accounts for the improvement in performance from the new depictions of altitude we have studied? The experiments were motivated by several nonexclusive possibilities. One is that size and contrast cues are stimulus dimensions to which the visual system can efficiently guide attention for making relational comparisons. The probability summation analyses of Experiment 1 (and to a lesser extent, Experiment 2) indicated that participants performed as well in the combined-cue condition as would be predicted from the independent summing of two visual cues. Based on this view, it is important for enhanced displays to choose stimulus dimensions that support attentional guidance (for a list of such dimensions, see Wolfe and Horowitz, [2004].

It is interesting to note that the combined-cue displays were more visually heterogeneous than the no-cue displays, but elicited faster and more accurate search performance. Typically, heterogeneous search displays yield less efficient searches when compared to homogenous displays [Duncan and Humphreys 1989]. We believe the reason for the opposite result was because the visual heterogeneity contained useful information for the search process.

A second possible explanation for the improvement of conflict detection observed in conditions with added cues is that the features of size and contrast allowed for greater contributions from peripheral vision. Conventional displays using only data tags necessarily limit the possibilities of processing an array in the periphery and, consequently, focal attention is required to decipher the numbers indicating altitude. Whereas the precise numbers given by data tags are required for some purposes, many aspects of the operator's task do not require such precision. For example, ruling out a conflict between an aircraft at 25, 000 and one at 5000 feet using our altitude encoding scheme need not involve foveation or the symbolic aspects of data tags.

Another explanation for the superiority of displays with size and contrast cues is the possible engagement of depth perception processes. Relative size and contrast can both function as pictorial depth cues to relative depth, so some modest segregation of the scenes into different depth planes could have contributed to the observed effects. Subjective reports did not suggest vivid 3D percepts of these displays, but any sense of depth stratification could have allowed attention to be allocated to different depth planes [He and Nakayama 1992] possibly aiding efficient exclusion of aircraft pairs that did not share the same cues.

Finally, participants may have benefited from our size and contrast altitude encoding scheme because these dimensions are not only perceptually salient, but also conceptually compatible with mental representations of magnitude. Correspondences between perceptual dimensions in visual displays and abstract representations can be attributed to conventional metaphors [Clausner 2002]. It is natural to interpret graphical symbols that are larger or darker as meaning more and this may have helped our participants apprehend the altitude information from the data tags and speeded their processing of this information. Future work that attempts to distinguish among the relative contributions of attentional guidance, improved peripheral visibility, depth processing, and metaphoric correspondence to the observed benefits of size and contrast cues would be useful.

Another issue for future investigation is the effectiveness of size and contrast cues in situations where aircraft are changing altitudes quickly, such as approach and departure control. In these situations, aircraft at similar altitudes would have similar size and contrast features, making visual search for aircraft at close altitudes quick and effective. However, since the sizes and contrasts of the aircraft icons would be changing, it remains to be shown that observers can effectively search for, say, a conflicting pair of aircraft where one is ascending and the other is descending.

Similarly, the detection of conflicts in situations with aircraft traveling at different velocities warrants investigation. In real-world situations, aircraft conflicts may be complicated by the fact that aircraft going in different directions may have significantly different ground speeds. The benefits we found of perceptual encoding were useful even in the somewhat simplified situation of assuming of constant velocity for all aircraft. We suspect that making the task even more challenging in this way might show even greater benefits of perceptual encoding, due to the selectivity with which controllers can look at relevant aircraft.

The present results showing improvement in aircraft conflict detection from size and contrast cues to altitude may be valuable because adding these features to existing ATC displays would be relatively straightforward, as opposed to some other efforts at redesigning ATC displays. A number of proposals have been made for more radical changes in ATC displays, such as using stereoscopic and motion parallax information to produce vivid impressions of depth. These proposals encounter a variety of problems, ranging from the complexity and physical discomfort associated with some implementations (e.g., “simulator sickness” in virtual reality displays; see Nichols and Patel [2002] for a review) to tradeoffs in which altitude information is made more salient at the expense of lateral position information [St. John et al. 2001]. Moreover, radically new display formats cause serious difficulties in acceptance by experienced users and may attenuate the accumulated expertise of these users. The addition of size and contrast cues to ATC displays has the virtue of building incrementally on existing displays, possibly allowing better acceptance and transfer of expertise, while still offering significant value.

Finally, we believe that more extensive training with this encoding scheme might allow perceptual learning to occur, which would optimize information pickup from the displays [Gibson 1969; Kellman and Kaiser 1994; see Kellman 2002 for a review]. The encoding of critical abstract information by means of visual metaphors and the consequent fluency acquired through training may be helpful for designing optimal interfaces. These issues are ripe for continued investigation, both in air traffic control and other contexts.

ACKNOWLEDGMENTS

The authors gratefully acknowledge partial funding for this research by Raytheon Company to T. C. Clausner, with the supportive sponsorship of David Bloomstran, and support from the U.S. Office of Naval Research MURI Program to P. J. Kellman. We also thank Tom Wickens for his helpful statistical advice.

The altitude encoding schemes of air traffic control displays described herein are the property of Raytheon Company (US patent pending; European patent #1474789, issued 06/2005 to inventors P. Kellman, T. Clausner, and E. Palmer).

REFERENCES

ANSTIS, S. M. 1974. A chart demonstrating variations in acuity with retinal position. *Vision Research 14*, 589–592.

BECK, J. 1994. Interference in the perceived segregation of equal-luminance element-arrangement texture patterns. *Perception & Psychophysics 56*, 4, 424–430.

BENEL, R. A. 1998. Workstation and software interface design in air traffic control. In *Human Factors in Air Traffic Control*, M. W. Smolensky and E. S., Stein, Eds. Academic Press, San Diego, CA. 17–63.

BURNETT, M. S. AND BARFIELD, W. 1991. Perspective vs. plan-view air traffic control displays: Survey and empirical results. In *Proceedings of the Human Factors Society, 35th Meeting*. 87–91.

CLAUSNER, T. C. 2002. How conceptual metaphors are productive of spatial-graphical expressions. In *Proceedings of the 24th Annual Conference of the Cognitive Science Society.* Erlbaum, Mahwah, NJ. 208–213.

CLAUSNER, T. C. AND CROFT, W. 1997. Productivity and schematicity in metaphors. *Cognitive Science 21*, 247–282.

CLAUSNER, T. C. AND CROFT, W. 1999. Domains and Image Schemas. *Cognitive Linguistics 10*, 1–31.

DA VINCI, L. c. 1492. Della prospettiva aerea [Of aerial perspective]. In *The Notebooks of Leonardo da Vinci*, J. P., Richter (Transl., 1883/1970). Dover, London, UK. 159–166.

DUNCAN, J. AND HUMPHREYS, G. W. 1989. Visual search and stimulus similarity. *Psychological Review 96*, 433–458.

EGETH, H. E., VIRZI, R. A., AND GARBARD, H. 1984. Searching for conjunctively-defined targets. *Journal of Experimental Psychology: Human Perception & Performance 10*, 32–39.

ELLIS, S. R. AND PALMER, E. A. 1981. Threat perception while viewing single intruder conflicts on a cockpit display of traffic information (NASA TM 81341), NASA Ames Research Center, Moffett Field, CA.

ELLIS, S. R., MCGREEVY, M. W., AND HITCHCOCK, R. J. 1987. Perspective traffic display format and airline pilot traffic avoidance. *Human Factors 29*, 4, 371–382.

GIBSON, E. 1969. *Principles of Perceptual Learning and Development.* Appleton-Century-Crofts, East Norwalk, CT.

HE, Z. J. AND NAKAYAMA, K. 1992. Surfaces versus features in visual search. *Nature (London) 359*, 6392, 231–233.

HEALEY, C. G., BOOTH, K. S. AND ENNS, J. T. 1996. High-speed visual estimation using preattentive processing. *ACM: Transactions on Computer-Human Interaction 3*, 2, 107–135.

HOCHBERG, J. E. AND MCALISTER, E. 1955. Relative size vs. familiar size in the perception of represented depth. *American Journal of Psychology 68*, 294–296.

HOLWAY, A. F., AND BORING, E. G. 1941. Determinants of apparent visual size with distance variant. *American Journal of Psychology 54*, 21–37.

HUNT, S. M. J. 1994. MacProbe: A Macintosh-based experimenter's workstation for the cognitive sciences. *Behavior Research Methods, Instruments & Computers 26*, 3, 345–351.

JOHNSON, M. 1987. *The body in the Mind: The Bodily Basis of Meaning, Imagination, and Reason.* University of Chicago Press, Chicago, IL.

JOHNSTON, J. C., HORLITZ, K. L. AND EDMISTON, R. W. 1993. Improving situation awareness displays for air traffic controllers. In *Proceedings of the Seventh International Symposium on Aviation Psychology*, R. S. Jensch and D. Neumeister, Eds. Ohio State University, Columbus, Off.

KELLMAN, P. J. 2002. Perceptual learning. In *Stevens' Handbook of Experimental Psychology, 3rd edition, Vol. 3: Learning, Motivation and Emotion*, R. Gallistel, Ed. Wiley, & NY.

KELLMAN, P. J. AND KAISER, 1994. Perceptual learning modules in flight training. In *Proceedings of the 38th Annual Meeting of the Human Factors and Ergonomics Society*. 1183–1187.

LAKOFF, G. 1993. The contemporary theory of metaphor. In *Metaphor and Thought A. Ortony*, 2nd ed. A Ortong, Ed. Cambridge University Press, Cambridge, UK.

LAKOFF, G. AND JOHNSON, M. 1980. The metaphorical structure of the human conceptual system. *Cognitive Science 4*, 195–208.

NICHOLS, S. AND PATEL, H. 2002. Health and safety implications of virtual reality: A review of empirical evidence. *Applied Ergonomics 33*, 3, 251–271.

RAYNER, K., MCCONKIE, G. W. AND ZOLA, D. 1980. Integrating information across eye movements. *Cognitive Psychology 12*, 2, 206–226.

REMINGTON, R. W., JOHNSTON, J. C., AND RUTHRUFF, E. 2000. Visual search in complex displays: Factors affecting conflict detection by air traffic controllers. *Human Factors 42*, 5, 349–366.

ROSKE-HOFSTRAND, R. J. AND MURPHY, E. D. 1998. Human information processing in air traffic control. *Human Factors in Air Traffic Control*, M. W. Smolensky, and E. S. Stein, Eds. Academic Press, San Diego, CA. 65–114.

SMITH, J. D., ELLIS, S. R. AND LEE, E. C. 1984. Perceived threat and avoidance maneuvers in response to cockpit traffic displays. *Human Factors 26*, 33–48.

ST. JOHN, M., COWEN, M. B., SMALLMAN, H. S., AND OONK, H. M. 2001. The use of 2D and 3D displays for shape-understanding versus relative-position tasks. *Human Factors 43*, 1, 79–98.

STENGER, A. J., ZIMMERLIN, T. A., THOMAS, J. P., AND BRAUNSTEIN, M. L., 1981. Advanced computer image generation techniques exploiting perceptual characteristics. US AFHRL Technical Report. TR 80–61.

TREISMAN, A. AND GELADE, G. 1980. A feature-integration theory of attention. *Cognitive Psychology 12*, 1, 97–136.

WOLFE, J. M. 1994. Guided Search 2.0: A revised model of visual search. *Psychonomic Bulletin & Review 1*, 202–238.

WOLFE, J. M. 1998. What can 1, 000, 000 trials tell us about visual search? *Psychological Science 9*, 1, 33–39.

WOLFE, J. M., CAVE, K. R., AND FRANZEL, S. L. 1989. Guided search: An alternative to the feature integration model for visual search. *Journal of Experimental Psychology: Human Perception & Performance 15*, 419–433.

WOLFE, J. M. AND HOROWITZ, T. S. 2004. What attributes guide the deployment of visual attention and how do they do it? *Nature Reviews Neuroscience 5*, 1–7.

Received May 2004; revised December 2005 and August 2006; accepted April 2006