A Theory and Model of Conflict Detection in Air Traffic Control: Incorporating Environmental Constraints

Shayne Loft
University of Western Australia

Scott Bolland, Michael S. Humphreys,
and Andrew Neal
University of Queensland

A performance theory for conflict detection in air traffic control is presented that specifies how controllers adapt decisions to compensate for environmental constraints. This theory is then used as a framework for a model that can fit controller intervention decisions. The performance theory proposes that controllers apply safety margins to ensure separation between aircraft. These safety margins are formed through experience and reflect the biasing of decisions to favor safety over accuracy, as well as expectations regarding uncertainty in aircraft trajectory. In 2 experiments, controllers indicated whether they would intervene to ensure separation between pairs of aircraft. The model closely predicted the probability of controller intervention across the geometry of problems and as a function of controller experience. When controller safety margins were manipulated via task instructions, the parameters of the model changed in the predicted direction. The strength of the model over existing and alternative models is that it better captures the uncertainty and decision biases involved in the process of conflict detection.

Keywords: air traffic control, conflict detection, expertise, ecological rationality

A significant amount of research has been published over the past 4 decades concerning the mechanisms that underlie expert performance (e.g., Chi, Glaser, & Farr, 1988; Ericsson, Charness, Feltovich, & Hoffman, 2006; Simon & Chase, 1973), including a recent special issue of the Journal of Experimental Psychology: Applied (Ericsson & Williams, 2007). There are both practical and theoretical reasons for studying expertise. From a practical perspective, understanding expertise is important for the design of training programs (Haerem & Rau, 2007) and for the development of models of human performance to help evaluate design concepts or inventions that may affect expert performance and system safety (Byrne & Gray, 2003; Pew & Mavor, 1998). From a theoretical perspective, the principles and mechanisms proposed to underlie expertise can be used to evaluate the generalizability of broader theories about basic cognitive processes and capacities and thus explain human performance more generally (Ackerman, 2005; Ericsson, 2004).

Much of the expertise research has used some variant of a methodological framework referred to as the expert performance approach (Ericsson & Smith, 1991). This approach involves identifying domain-specific situations in which experts routinely outperform novices and designing representative tasks that elicit this superior performance. This approach has been used successfully to determine the knowledge representations and processing strategies that mediate expert performance in task domains as varied as computer programming, medical diagnosis, sports, chess, Scrabble, physics, and music (e.g., Chi, Feltovich, & Glaser, 1981; Kulatunga-Moruzzi, Brooks, & Norman, 2001; Larkin & McDermott, 1980).

However, understanding expertise in many work domains requires specification not just of the mental processes of experts and how they relate to performance, but also of operational task environments and their interaction with these mental processes. This is because occupations such as air traffic control (ATC), fire fighting, and military command and control require experts to adapt their decision making in the face of dynamic conditions, concurrent task demands, time pressure, and tactical constraints (Sheridan, 2002; Tsang & Vidulich, 2003). An emphasis on the role of the environment in shaping human decision processes has long been advocated by the ecological rationality approach to psychology (Brunswick, 1943; Gibson, 1966; Gigerenzer & Todd, 1999; Shepard, 1984; Simon, 1956).

A prototypical example in which the relationship between the task environment and expertise is poorly understood is conflict detection in ATC. In this article, we present a performance theory of conflict detection that identifies (a) the goals of controllers, (b) the characteristics of the ATC environment that constrain the manner in which controllers detect conflicts, and (c) the resulting uncertainty and decision biases involved in the process of conflict detection. This theory then serves as a framework for building a computational model that can fit the controller intervention decisions obtained across two experiments.

A Performance Theory of Conflict Detection

In developing our performance theory, we assumed that human decisions are made within the bounds of limited time and knowl-
edge and that the efficiency and effectiveness of human decisions are optimized through interactions between the human’s adaptive mechanisms and the local task environment (Anderson, 1990; Gigerenzer & Todd, 1999; Simon, 1956). In this manner, our performance theory specifies both the goals of the controller and the structure of the environment to which decisions are being optimized. This task analysis is at least partially consistent with what Marr (1982) called the computational level (inputs and outputs of the task), in that it provides a precise characterization and justification of the behavior that the controller is trying to achieve.

The goal of controllers is to ensure that aircraft are separated and to expedite and maintain orderly traffic flows. The safety of aircraft is the highest priority. Conflict detection requires controllers to compare the trajectories of aircraft to assess whether they will violate minimum separation. Aircraft are in conflict if they will violate both lateral (e.g., 5 nautical miles [nm]) and vertical separation (e.g., 1,000 ft [304.8 m]) standards simultaneously. Timely conflict detection is essential to allow sufficient time to resolve conflicts while minimizing any disruption to the aircraft involved.

There is a tradition of empirical studies examining the factors that influence how controllers visually search for potential conflicting aircraft pairs (e.g., Galster, Duley, Masalonis, & Parasuraman, 2001; Gronlund, Ohrt, Dougherty, Perry, & Manning, 1998) or the factors that influence the accuracy and timeliness of conflict detection after specific aircraft pairs have been attended (e.g., Bisseret, 1981; Boag, Neal, Loft, & Halford, 2006; Leplat & Bisseret, 1966; Rantanen & Nunes, 2005). However, formal theories or models describing either of these processes are rare. In this article, we focus on developing a theory of how controllers extrapolate the trajectory of aircraft to which they have selectively attended to make conflict status decisions.

Neal and Kwantes (in press) developed a model using a random walk formalism, assuming that participants sequentially sample evidence regarding the separation between aircraft from a distribution and accumulate that evidence over time. The model provided a good fit to the responses of participants on a laboratory conflict detection task. However, the model was limited in that it was tested using naïve participants and aircraft altitude was not simulated. Other researchers have established statistical relationships between environmental cues (e.g., lateral distance) and conflict decisions (Averty, Guittet, & Lezaud, 2008; Bisantz & Pritchett, 2003; Stankovic, Raufaste, & Averty, 2008). However, these models are opaque with respect to underlying cognitive processes, making it difficult for them to make plausible predictions in new situations.

Furthermore, it may be possible to improve the robustness and generality of conflict detection models by incorporating environmental constraints (Todd & Gigerenzer, 2007). A fundamental characteristic of the ATC environment that places constraints on conflict detection is the uncertainty in the trajectory of aircraft. This uncertainty can stem from both mental (perceptual–cognitive) and environmental sources. Controllers can use aircraft state information (altitude, velocity, and heading) to estimate the future relative position of aircraft using a variety of perceptual and cognitive methods (Xu & Rantanen, 2003). For example, controllers can use the distance-to-velocity ratios of aircraft to estimate the distance or time separating aircraft at crossing points (Neal & Kwantes, in press). However, estimations of the relative arrival time of objects moving on two-dimensional display screens are prone to error (Law et al., 1993). This uncertainty is further increased in ATC because controllers are required to assess the lateral and vertical separation of aircraft simultaneously when aircraft are climbing or descending through the altitudes of other aircraft (Boag et al., 2006).

One way that controllers can decrease this uncertainty is to use decision support tools. For example, bearing and range lines can be used to calculate distances and times between aircraft or between aircraft and crossing points. However, because of variations in flight environmental factors (wind shift, aircraft load, engine parameters, temperature, airline operating rules, etc.), controllers need to allow for some level of uncertainty (Averty et al., 2008; Nunes & Kirlik, 2005). One of the strengths of human memory and learning is the ability to abstract statistical regularities from task environments (Anderson, 1991; Estes, 1950; Gigerenzer, Höff- frage, & Kleinbölting, 1991), and we make use of this strength in our assumption that with experience controllers store in memory a representative sample (cf. Brunswick, 1964) of likely variance in aircraft trajectory.

In addition to using tools, controllers can further reduce uncertainty by deferring conflict decisions to allow examination of environmental factors influencing aircraft trajectory. However, controllers work under time pressure. To manage workload, therefore, it is beneficial for controllers to perform tasks quickly (Loft, Sanderson, Neal, & Mooij, 2007; Sperandio, 1978). Controllers are likely to stop acquiring aircraft trajectory information when the expected utility of making more precise estimates cannot be justified. Conflict detection delays impose constraints because they decrease the time available to intervene and can produce a memory load because the controller has to remember to return to the potential conflict (Loft, Humphreys, & Neal, 2003). As a result of this truncation in information acquisition, some variance in aircraft trajectory is irreducible, and we argue that controllers use approximating (satisficing; Simon, 1956) mechanisms for determining the future position of aircraft.

Bisseret (1981) reported data that are consistent with our view that controllers adopt strategies that compensate for environmental constraints to ensure aircraft separation. Bisseret presented experts and trainees with pairs of converging aircraft. Conflicts were rarely missed (0.87%), but there was a high false alarm rate. In fact, only 30% of the aircraft with which experts intervened would have missed the aircraft with which experts intervened would have violated separation, assuming no variability in aircraft trajectory. Experts were more likely to make false alarms than trainees. Bisseret claimed that experts are more psychologically conservative than trainees; they aim to minimize the risk of conflict by setting a low criterion for intervention. This interpretation emphasizes the internal processes of the controller, in terms of his or her sensitivity to payoffs, motivation, and so forth. We agree that differences in decision criteria may be formed through experience with the task and may vary between controllers. However, we argue that controller intervention is also influenced by the expected uncertainty in aircraft trajectory, which is a property of both the structure of the environment and the processes by which information is sampled from the environment.

According to our performance theory, controllers apply safety margins to ensure aircraft separation. These safety margins are formed through experience and reflect two underlying processes. First, controllers hold expectations regarding variation in aircraft
trajectory. Depending on this uncertainty, their predictions of aircraft position at specific points in the future will be some distance closer or further (or lower or higher in the vertical plane) than the positions predicted by aircraft state values. Second, safety margins reflect the degree to which decision criteria are biased to favor safety over accuracy (Bisseret, 1981).

These safety margins ensure controllers do not miss aircraft conflicts to which they have selectively attended. It is counterintuitive then that Boag et al. (2006) reported a miss rate of 10% in a simulation in which aircraft pairs were presented in isolation. However, Boag et al. instructed controllers to indicate which aircraft pairs would violate separation and which would not. This required controllers to adopt a level of information sampling that led to the finest possible discrimination between conflicts and nonconflicts. In reality, intervening to ensure separation between aircraft that are not in conflict, although affecting efficiency, is far less of a concern than missing a conflict. Bisseret (1981) used more appropriate instructions, asking controllers to indicate which aircraft would require intervention to avoid a loss of separation, and subsequently reported that experts rarely missed conflicts.

The two experiments reported here were designed in consultation with subject matter experts. To control visual search, trials were presented in which a single pair of aircraft converged to a crossing point. Controllers were asked to make conflict status judgments by indicating whether they would intervene to ensure separation. A 5-nm lateral and 1,000-ft (304.8-m) vertical separation standard was used, and controllers had access to decision support tools. The purpose of the experiments was to quantify how intervention varied across the geometry of problems and as a function of controller experience. In addition, we expected variations in response time and tool use to be indicators of information acquisition strategies. In the modeling section that follows these experiments, we demonstrate that a model based on the assumptions of our theory can provide a good fit to the controller intervention decisions.

Experiment 1

Experiment 1 used controllers with different experience levels. Experts were licensed controllers. Trainees had completed 1 year of training. The minimum vertical and lateral distance of aircraft pairs were manipulated, along with the type of problem. For lateral problems, both aircraft were at level flight. On the basis of aircraft velocities and climb rates, the vertical distance when aircraft were at level flight and another was climbing, with lateral separation set a 0 nm. We also manipulated the angle of intersection for both problem types (45°, 90°, and 135°).

Intervention should decrease with increases in minimum distance (e.g., Stankovic et al., 2008). More important, though, in contrast to Boag et al. (2006), we do not expect experts to miss conflicts. We expected to replicate the finding of Neal and Kwantes (in press) that controllers are more likely to classify aircraft as conflicts with decreasing angle.

When Bisseret (1981) asked controllers to give numerical estimates of minimum distance, experts were less accurate than trainees, suggesting that experts estimated greater uncertainty in aircraft trajectory. In addition, although the cost of missing a conflict should be as important to the trainee as it is to the expert, controllers may learn through experience the benefits of being cautious (e.g., delayed decisions, although potentially more accurate, may produce higher workload). Taken together, this suggests experts will apply larger safety margins than trainees and will be more likely to intervene. We also expected experts to make faster decisions (Stankovic et al., 2008), either because they are better practiced at conflict detection or because they choose to spend less time making trajectory calculations.

It is difficult to make predictions regarding the effect of problem type because lateral and vertical distances are measured with different distance metrics. However, the ATC interface is two dimensional, representing the lateral coordinates of aircraft and routes, allowing the direct perception of spatial relationships between aircraft and the use of tools. In contrast, vertical position is displayed numerically in the aircraft data block. Thus, climbing rates must be predicted on the basis of successive readings of changing altitude and estimations of elapsed time between readings. This suggests that controllers should apply larger safety margins to vertical distance than to lateral distance and thus be more likely to intervene when aircraft are changing altitude than when aircraft are at level flight.

Method

Participants. Thirteen licensed controllers at Brisbane ATC Centre (Brisbane, Queensland, Australia) participated. They had been controllers for an average of 15.2 years and had an average age of 39.4 years (12 men and 1 woman). Seven trainees had an average age of 26 years (6 men and 1 woman). These trainees had completed 1 year of training, which consisted mainly of theory but some practice.

Conflict detection task. We used ATC-lab Advanced (Fothergill, Loft, & Neal, 2009). As illustrated in Figure 1, each trial presented a single pair of aircraft traversing a fictitious en route sector. The total area of the airspace was 260 nm × 195 nm. Each aircraft had a data block that displayed the call sign, the aircraft type, the current and cleared flight level (> indicates aircraft were at level flight, and indicates aircraft were climbing), and the velocity in nautical miles. The tracks crossed at the intersection at the center of the sector. Aircraft position was updated every 5 s. Aircraft were in conflict if they would, given their respective flight levels, velocities, and headings, simultaneously violate vertical and lateral separation in the future.

Controllers had no control over the flight levels, velocities, or headings of aircraft. Their task was to make a series of judgments. The first response panel asked controllers whether they would intervene now (or in the future) to ensure separation and provided four response options (definitely, likely, unlikely, and definitely not). When definitely or likely were selected, a second response panel was presented that asked controllers to type into a response

---

1 A subset of the data from Experiment 1 has previously been reported in Loft, Bolland, and Humphreys (2007) and Fothergill, Loft, and Neal (2009). These articles presented the expert and trainee intervention data from Experiment 1 for lateral problems only (as function of minimum lateral distance, but not as a function of angle). None of the models reported in this article were reported in these articles.
box the time (e.g., immediately, in 1 min, or in 2 min) at which they would intervene to resolve the conflict and to specify in the response box the exact method they would use to resolve the conflict. Conflict resolution methods include but are not limited to changing the flight level of aircraft, assigning flight-level requirements to aircraft (e.g., requesting pilots to reach altitude $x$ by $x$ distance or by $x$ time), and vectoring aircraft.

A 20 nm $\times$ 10 nm scale maker was provided in the top left-hand corner of the screen and could be moved. A number of decision support tools were available. A range–bearing line could be selected to determine the bearing of an aircraft, its distance from a crossing point and other aircraft, and the time at which it would reach a given crossing point. Velocity vectors indicated the future positions of an aircraft from 1 min to 15 min in time. These tools used only current state information and did not consider uncertainty in flight trajectory.

A screenshot of the ATC-lab $^{\text{advanced}}$ task. The two aircraft are converging to the intersection point located at the center of the screen. VOZ515 is at flight level 370. QFA322 is climbing through the flight level of VOZ515 to a cleared flight level of 410. The minimum lateral distance of these aircraft will be 0 nm. The vertical distance between the aircraft when they are separated by 5 nm will be 400 ft (121.92 m). Two-minute velocity vectors are selected for each aircraft (pointing in the direction the aircraft are traveling). Controllers also had the option of using range–bearing lines (not shown). A scale marker is available in the top left-hand corner of the sector. Participants respond by using the response panel provided.

Materials. Ten types of aircraft were presented (e.g., 767). Each aircraft type had a velocity range (767; 420 nm to 480 nm) and a flight-level range (767; 30,000 ft to 43,000 ft [9,144 m to 13,106.4 m]). Aircraft types were randomly allocated to problems, with velocity and flight level randomly selected from the appropriate range of that aircraft type.

Group (expert or trainee) was the between-subjects factor. Problem type (lateral or vertical), minimum distance (for lateral problems in nautical miles: 0, 1, 2, 4, 6, 8, 10, 12, 14, 16, 18, and 20; for vertical problems in feet: 0, 200, 400, 800, 1,200, 1,600, 2,000, 2,400, 2,800, 3,200, 3,600, and 4,000 [meters: 0, 60.96, 121.92, 243.84, 365.76, 487.68, 609.60, 731.52, 853.44, 975.36, 1,097.28, and 1,219.20]), and angle (45°, 90°, and 135°) were manipulated within subjects.

Thus, 36 lateral problems were presented, 12 of which were in conflict (minimum distance < 5 nm). Thirty-six vertical problems
were also presented, 12 of which were in conflict (minimum distance < 1,000 ft [304.8 m] when lateral distance < 5 nm). For vertical problems, the minimum lateral distance was set at 0 nm. One aircraft was cruising and the other was climbing to a cleared flight level 4,000 ft (1,219.20 m) above the cruising aircraft. In the 0-ft minimum vertical distance case, the climbing aircraft would be at the exact same flight level at the time of lateral conflict. In the 4,000-ft minimum vertical distance case, the climbing aircraft would be 4,000 ft above the cruising aircraft at the time of lateral conflict. Time to point of closest approach refers to the time from the start of trials to when an aircraft pair reached minimum distance. Two time ranges were used (5–7 min and 13–15 min), and these were counterbalanced.

Procedure. Task instructions were read by controllers and verbally reinforced by the experimenter. The instructions informed controllers that they would be performing a conflict detection task and explained how to use the response panels. Controllers were asked to assume that they were under moderate workload conditions, and they were reminded that there are environmental factors that can affect aircraft trajectory and to take them into account as they would in the field. Controllers completed four practice trials before commencing the main task.

Results

The primary dependent measure was the probability that controllers would intervene. The majority of controllers in both experiments used the 4-point response scale categorically, choosing either definitely or definitely not. Nunes and Kirlik (2005) also found that controllers were more likely to select nonconflict (0%) or conflict (100%) probability values rather than intermediate values (10% to 90%). This pattern of responding reflects the dichotomous nature of conflict detection; controllers either intervene or they do not. On this basis, cases in which controllers responded definitely intervene or definitely not intervene were coded as 1 (intervene). Cases in which controllers responded unlikely to intervene or definitely not intervene were coded as 0 (not intervene).

We conducted a series of Group × Type × Minimum Distance × Angle mixed analyses of variance (ANOVA) on the probability of intervention, response time, and decision support tool use. Group was a between-subjects factor, and problem type, minimum distance, and angle were within-subject factors. The data were collapsed over successive levels of minimum distance. Effect sizes are given as eta-square. The values for small, medium, and large effect sizes are .10, .25, and .40, respectively (Cohen, 1988).

Probability of intervention. The probabilities that expert and trainee controllers intervened in lateral and vertical problems are presented in Figure 2. Intervention decreased with increased minimum distance, \( F(5, 90) = 88.52, p < .01, \eta^2 = .83 \). A main effect was found for group, \( F(1, 18) = 17.28, p < .01, \eta^2 = .49 \), with experts (\( M = .86, SD = .08 \)) more likely to intervene than trainees (\( M = .67, SD = .13 \)). A main effect was found for problem type, \( F(1, 18) = 88.29, p < .01, \eta^2 = .83 \), with controllers more likely to intervene in vertical (\( M = .94, SD = .12 \)) than in lateral (\( M = .64, SD = .18 \)) problems. There was a main effect of angle, \( F(2, 36) = 11.13, p < .01, \eta^2 = .38 \), with controllers more likely to intervene in problems with smaller convergence angles (45°, \( M = .83, SD = .13 \); 90°, \( M = .78, SD = .14 \); and 135°, \( M = .76, SD = .13 \)).

There were interactions between group and minimum distance, \( F(5, 90) = 7.78, p < .01, \eta^2 = .30 \), and between problem type and minimum distance, \( F(5, 90) = 57.38, p < .01, \eta^2 = .76 \). However, these interactions were qualified by a three-way interaction between group, problem type, and minimum distance, \( F(5, 90) = 4.89, p < .01, \eta^2 = .21 \).

To decompose this three-way interaction, we examined the relationship between group and minimum distance at each level of problem type. The interaction between group and minimum distance was different for lateral and vertical problems. For lateral problems, when minimum lateral distance was 4 nm or less, experts and trainees would always intervene. However, at larger minimum distances (6 nm or greater), experts were more likely to intervene than trainees. In contrast, for vertical problems experts almost always intervened, whereas the probability that trainees intervened generally decreased with increased vertical distance.

There was also a Problem Type × Angle interaction, \( F(2, 36) = 5.25, p < .05, \eta^2 = .23 \). The probability of controller intervention increased with decreased angle for lateral problems, whereas angle had no effect on controller intervention in vertical problems. A two-way interaction between minimum distance and angle, \( F(10, 180) = 1.84, p = .06 \), and a three-way interaction between problem type, minimum distance, and angle, \( F(10, 180) = 1.85, p = .05 \), approached significance. No other main effects or interactions approached significance.

Response time and tool use. Response time was measured in seconds from the start of each trial to the time that controllers made decisions. As predicted, experts (\( M = 19.82, SD = 5.76 \)) responded faster than trainees (\( M = 30.61, SD = 9.61 \)), \( F(1, 18) = 10.01, p < .01, \eta^2 = .36 \). A main effect of minimum distance approached significance, \( F(5, 90) = 2.16, p = .07 \). There was a Problem Type × Minimum Distance interaction, \( F(5, 90) = 3.91, p < .01, \eta^2 = .18 \). For lateral problems, response time was fastest for 0–1 nm problems (\( M = 20.47, SD = 8.60 \)), then increased as lateral distance increased up to 6–8 nm (\( M = 30.02, SD = 15.90 \)), and then decreased as lateral distance increased up to 18–20 nm (\( M = 23.33, SD = 10.72 \)). In contrast, response time to vertical problems did not change systematically with minimum vertical distance (\( M = 22.55, SD = 10.12 \)). No other main effects or interactions approached significance.

Controllers used velocity vectors on 76% of trials and range–bearing lines on 38% of trials. We coded tool use as 1 if controllers used either tool on a given trial. There was a trend for experts (\( M = .76, SD = .25 \)) to use tools less than trainees (\( M = .92, SD = .13 \)), but this did not reach significance, \( F(1, 18) = 2.63, p = .12 \). However, controllers were more likely to use tools for lateral (\( M = .92, SD = .15 \)) than for vertical (\( M = .71, SD = .31 \)) problems, \( F(1, 18) = 17.03, p < .01, \eta^2 = .49 \). No other main effects or interactions approached significance.

Discussion

Several key findings can be summarized. First, intervention decreased as the minimum lateral distance of problems increased. In contrast to past studies that have also controlled visual search (Boag et al., 2006), experts rarely missed conflicts (only one conflict was missed). Second, controllers were more likely to intervene, and were less likely to use tools, for vertical compared with lateral problems. Remarkably, experts almost always inter-
vened in vertical problems, even when the minimum vertical distance was four times the prescribed standard. Third, controllers were more likely to intervene in lateral problems with a smaller angle of convergence. Fourth, experts were more likely to intervene and made faster decisions than trainees. As demonstrated later, controller intervention can be closely fit by a model that assumes that controller safety margins differ as a function of expertise.

**Experiment 2**

The purpose of Experiment 2 was to test the validity of our model by inducing shifts in controller safety margins and to replicate the effect of vertical distance found in Experiment 1 using a more representative sample of conflict detection problems. One method used in the memory literature to manipulate decision making is to deliver explicit or implicit test instructions that highlight to participants the preferred conservativeness of response (Curran, DeBuse, & Leynes, 2007; Macmillan & Creelman, 1991; Van Zandt, 2000). We did not expect controllers to respond to an explicit instruction to be more or less conservative because controllers are trained to view the safety of the aircraft under their jurisdiction in black-and-white terms. Instead, we used what we thought was a more naturalistic method by instructing controllers to assume that they were under conditions of high or low workload before they commenced blocks of experimental trials.

Many theories of strategic control have proposed that operators maintain acceptable workload levels by using strategies that minimize the control activity required to meet their objectives (Loft et al., 2007; Rouse, Edwards, & Hammer, 1993; Sperandio, 1971). Consistent with this, controllers often report that they are reluctant to let potential conflicts continue without intervention under high workload, preferring to intervene and reduce monitoring requirements (Amaldi & Leroux, 1995; Kallus, Van Damme, & Dittman, 1999). Thus, we assumed that controllers would respond to higher workload instructions by increasing safety margins. We also expected controllers to make faster decisions and to be less likely to use tools following a higher workload instruction.

Minimum distance was again manipulated along with problem type. Lateral problems were identical to Experiment 1. Remarkably, experts almost always intervened in vertical problems in Experiment 1 when minimum lateral separation was set at 0 nm, even those with minimum vertical separation four times the stan-
dard. It was important to replicate this novel finding. One possibility is that controllers apply smaller vertical safety margins when minimum lateral distance is predicted to be larger. We presented problems in Experiment 2 in which both the minimum lateral and vertical distance of aircraft were varied. For these problems, one aircraft was at level flight and the other was climbing. For vertical–8 nm problems, minimum lateral distance was set at 8 nm and vertical distance when aircraft reached 8 nm varied from 0 ft to 4,000 ft (0 m to 1,219.20 m). For vertical–16 nm problems, minimum lateral distance was set at 16 nm and vertical distance when aircraft reached 16 nm varied from 0 ft to 4,000 ft.

Method

Participants. Fourteen licensed controllers at Melbourne ATC Centre (Melbourne, Victoria, Australia) participated. They had been controllers for an average of 12.2 years and had an average age of 35.8 years (12 men and 2 women). Nine trainee controllers at Melbourne ATC Centre also participated, with an average age of 27.4 years (8 men and 1 woman). Experiment 2 was conducted 1 full year after Experiment 1. Trainees had been trained for a period of 1.5 years. This training program included both theory and practice.

Materials. The materials were the same as in Experiment 1 with the following exceptions. There were three types of problems: lateral problems, vertical–8 nm problems, and vertical–16 nm problems. Aircraft in lateral problems were at level flight. For vertical–8 nm problems, one aircraft was cruising and the other climbing, and minimum lateral distance was set at 8 nm. For vertical–16 nm problems, one aircraft was cruising and the other climbing, and minimum lateral distance was set at 16 nm. Problem type and minimum distance (for lateral problems in nautical miles: 0, 1, 2, 4, 6, 8, 10, 12, 14, 16, 18, and 20; for vertical problems in feet: 0, 200, 400, 800, 1,200, 1,600, 2,000, 2,400, 2,800, 3,200, 3,600, and 4,000 [meters: 0, 60.96, 121.92, 243.84, 365.76, 487.68, 609.60, 731.52, 853.44, 975.36, 1,097.28, and 1,219.20]) was manipulated within subjects. Instruction (low or high workload) and angle of intersection (45° or 135°) was manipulated within subjects across successive levels of minimum distance.

Thus, 24 lateral problems were presented, 8 of which were in conflict. Twenty-four vertical–8 nm problems and 24 vertical–16 nm problems were presented, none of which were in conflict. For vertical problems, climbing aircraft were climbing to a cleared flight level 4,000 ft above the paired aircraft that was at level flight. In the 0-ft vertical distance case, the climbing aircraft would be at the exact same flight level as the aircraft at level flight when lateral separation reached 8 nm or 16 nm. Time to minimum distance was set at 13–15 min.

Procedure. The procedure was the same as that in Experiment 1 with the exception of the workload instructions. Every four trials, an on-screen instruction asked controllers to assume that they were under conditions of low workload. By low workload, we mean that you are not busy and under very little time pressure. You have more than enough time to accomplish your tasks. The high-workload instruction read,

When making your decisions for the next four pairs of aircraft, please assume that you are under conditions of high workload. By high workload, we mean that you are very busy and under excessive time pressure. You are struggling to find the time to accomplish your tasks.

Results

Probability of intervention. As in Experiment 1, the data were collapsed over successive levels of minimum distance. We conducted a Group × Problem Type × Minimum Distance × Instruction × Angle mixed ANOVA. Group was a between-subjects factor, and problem type, minimum distance, instruction, and angle were within-subject factors. Although the trend was in the same direction as Experiment 1, experts (M = .72, SD = .13) were not significantly more likely to intervene than trainees (M = .65, SD = .21), F(1, 21) = 1.15, p = .30. Furthermore, group did not interact with other independent variables. Thus, we collapsed across the group variable when presenting the controller intervention data in Figure 3 (see Appendix for the data broken down by group).

We found a main effect for instruction, F(1, 21) = 27.77, p < .01, η² = .57, with controllers more likely to intervene following a high (M = .75, SD = .17) than a low (M = .63, SD = .16) workload instruction. There was a main effect of problem type, F(2, 42) = 51.10, p < .01, η² = .71. A follow-up simple effect test revealed that controllers were more likely to intervene in vertical–8 nm (M = .88, SD = .15) than in vertical–16 nm (M = .48, SD = .26) problems, F(1, 21) = 69.19, p < .01, η² = .77. These main effects were qualified by an interaction between instruction and problem type, F(2, 42) = 13.19, p < .01, η² = .39. The increase in intervention as a function of instruction was largest for vertical–16 nm problems (M干预 = .21), intermediate for lateral problems (M干预 = .10), and smallest for vertical–8 nm problems (M干预 = .05).

Intervention decreased with increased minimum distance, F(5, 105) = 26.54, p < .01, η² = .56. There was an interaction between problem type and minimum distance, F(10, 210) = 22.83, p < .01, η² = .52. When the minimum lateral distance of lateral problems was 4 nm or less, controllers would almost always intervene. As minimum lateral distance then increased (6 nm or greater), controllers were less likely to intervene. In contrast, intervention in vertical problems did not vary in any consistent manner with minimum vertical distance. This differential pattern of intervention to lateral and vertical problems as a function of minimum distance is consistent with Experiment 1.

Replicating Experiment 1, controllers were more likely to intervene in problems with a smaller angle of intersection (45°, M = .76, SD = .15; 135°, M = .62, SD = .19), F(1, 21) = 31.62, p < .01, η² = .60. There was also a two-way interaction between angle and problem type, F(1, 42) = 3.69, p < .05, η² = .15. The increase in intervention as a function of angle was larger for vertical–16 nm problems (M干预 = .23) than for lateral problems (M干预 = .12) and vertical–8 nm problems (M干预 = .10). No other main effects or interactions approached significance.

Response time and tool use. We conducted a Group × Problem Type × Minimum Distance × Instruction × Angle mixed ANOVA for response time and tool use. Controllers made decisions faster following a high- (M = 26.03, SD = 14.24) than a low- (M = 32.3, SD = 17.35) workload instruction, F(1, 21) = 13.21, p < .01, η² = .39. A follow-up simple effect test revealed that controllers had faster response times following a high (M = 17.6, SD = 12.3) than a low (M = 26.0, SD = 14.2) workload instruction, F(1, 33) = 10.42, p < .01, η² = .25. There was also a main effect of problem type, F(1, 21) = 17.77, p < .01, η² = .46.
In contrast to Experiment 1, there was no significant difference in response time between experts ($M = 30.46, SD = 15.9$) and trainees ($M = 27.15, SD = 12.75; F < 1$). There was a main effect of minimum distance, $F(5, 105) = 4.64$, $p = .01, \eta^2 = .18$. This main effect was qualified by an interaction between problem type and minimum distance, $F(10, 210) = 2.63$, $p = .01, \eta^2 = .11$. For lateral problems, response time was fastest for $0–1 \text{ nm}$ problems ($M = 20.14, SD = 10.39$) compared with problems with larger minimum distance ($6–8 \text{ nm}, M = 30.12, SD = 14.79; 18–20 \text{ nm}, M = 32.26, SD = 25.01$). In contrast, response time to $8 \text{ nm}$–vertical ($M = 29.78, SD = 17.76$) and $16 \text{ nm}$–vertical ($M = 30.64, SD = 17.82$) problems did not change systematically with minimum vertical distance. No other main effects or interactions approached significance.

Consistent with Experiment 1, controllers used velocity vectors (76% of trials) more than range–bearing lines (52% of trials). Collapsing across tool type, controllers used tools less following a high– ($M = .86, SD = .20$) than following a low– ($M = .94, SD = .17$) workload instruction, $F(1, 21) = 3.90, p < .05, \eta^2 = .18$. There was no significant difference in tool use between experts ($M = .85, SD = .23$) and trainees ($M = .94, SD = .18$), $F(1, 21) = 1.29, p = .27$. There was a main effect of problem type for tool use, $F(2, 42) = 4.13, p < .05, \eta^2 = .16$. A follow-up contrast revealed that consistent with Experiment 1, controllers were more
likely to use tools for lateral ($M = .95$, $SD = .14$) than for vertical ($M = .88$, $SD = .21$) problems, $F(1, 21) = 5.03, p < .05, \eta^2 = .19$.

Discussion

Several key findings can be summarized here. First, our instructional manipulation was successful in inducing shifts in controller safety margins because controllers were more likely to intervene following a high- than a low-workload instruction. This indicates that the safety margins that we propose underlie conflict detection are under the strategic control of controllers (Loft et al., 2007; Rouse et al., 1993). In addition, controllers made faster decisions and were less likely to use tools following a high- than a low-workload instruction. Replicating Experiment 1, the probability of intervention was dependent on minimum lateral distance and did not vary significantly with minimum vertical distance. Finally, consistent with Experiment 1, intervention increased as the angle of convergence decreased. As demonstrated later, controller intervention can be closely fitted by a model that assumes that controller safety margins shifted as a function of workload instructions.

Although trends were in the same direction as Experiment 1, experts were not significantly more likely to intervene, did not make faster decisions, and were not less likely to use tools than trainees. The size of the expert–trainee effect may have depended on the nature of the training programs. The trainees used in Experiment 2 were trained for a period of 1.5 years (as opposed to 1 year in Experiment 1), and the training at Melbourne included greater practical experience than that at Brisbane. If the changes to training in Experiment 2 decreased the size of the expertise effect, there should be evidence that it was the intervention of trainees, and not the experts, that differed across experiments. Figure 4 presents intervention by experts and trainees to 45° and 135° lateral problems across the two experiments. It is clear from Figure 4 that it was the intervention of trainees, and not the experts, that differed across experiments. The effect of expertise is likely to be small, and researchers examining expertise in the future should ensure that trainee controllers are sampled in the early stages of training, and perhaps before they gain practical experience.

Modeling Controller Intervention Decisions

The aim of this section is to explore, through the use of computational simulation, how well the performance theory proposed in this article can explain the empirical results. In testing our assumption that controllers use safety margins to ensure aircraft separation, we have implemented three models. The models make different assumptions regarding the psychological processes underlying controller safety margins. For simplicity, the models assume that controllers always deemed aircraft to be in vertical conflict because we found no consistent main effect of minimum vertical distance for experts in either experiment.

Signal Detection Theory Model

Signal detection theory (SDT) has been used in a range of applications for measuring the accuracy of system diagnosis under conditions of uncertainty (Bisseret, 1981; Swets & Pickett, 1982). In implementing a SDT model to describe the process of conflict detection, we assumed that controllers sample evidence regarding the predicted minimum lateral distance between aircraft and intervene if this minimum lateral distance exceeds a criterion (Neal & Kwantes, in press). In addition, the SDT model assumes that there is uncertainty in controllers’ estimates of minimum distance. The decision algorithm can be described as follows: Intervene if min-

![Figure 4](image_url)
minimum distance \( + g(\sigma) < c \); otherwise do not intervene, where \( c \) is the criterion, and \( g(\sigma) \) is the noise or uncertainty in the estimate, following a Gaussian distribution with 0 mean, and \( \sigma \) standard deviation. The SDT model assumes controllers do not place safety margins along the projected trajectories of aircraft. This model is used as a baseline in evaluating whether uncertainty in the projected trajectory of aircraft is a factor underlying controller decision making.

### Trajectory and Hybrid Models

In testing our assumption that controllers place uncertainty bounds around the projected trajectory of aircraft, we adapted the approach described in Granger, Durand, and Alliot (2001). In this approach, the position of an aircraft projected into the future can be modeled as a region of uncertainty bounded by the position at which the aircraft would be if it were flying at its minimum velocity, and a second point, representing the aircraft’s position if it were flying at its maximum velocity. Granger et al. assumed that uncertainty in aircraft position increases linearly with time. A conflict exists if there is a point in time in the future in which the projected regions of two aircraft violate lateral separation.

The trajectory and hybrid models differ from Granger et al. (2001) in two important ways. First, rather than bounding aircraft positions by actual aircraft performance, we assume that there are differences in the uncertainty of trajectory estimates between controllers. This allows the model’s decisions to be stochastic rather than deterministic. Second, we assume that uncertainty in aircraft position may not necessarily increase linearly with time, but may follow some other form of monotonically increasing function. Speculation on the exact nature of this function is beyond the scope of this article. Thus, for simplicity, we model the uncertainty in aircraft position as constant over time.

In emulating the responses of controllers, the trajectory and hybrid models project the trajectories of aircraft into the future, bounded by a region of uncertainty. This region of uncertainty represents the projected minimum and maximum positions of aircraft along the lateral plane. Because the minimum and maximum positions of the aircraft depend on velocity, we represent this uncertainty in terms of time rather than distance. If the trajectories are extrapolated \( t \) seconds into the future, the region in which the aircraft could be is calculated by determining where the aircraft will be at \( t - x \) and \( t + x \) seconds, based on its current velocity. Thus, the region of uncertainty is larger for aircraft that fly faster. The value of \( x \) is calculated for each decision using one parameter: \( x = lg(\sigma) \), where \( g(\sigma) \) is the uncertainty of the estimate, following a Gaussian distribution with 0 mean and \( \sigma \) standard deviation. The absolute value of the uncertainty is taken to ensure that it remains positive.

The trajectory model then calculates the minimum distance of separation, taking into account the potential minimum and maximum positions of aircraft. The controller intervenes if this minimum distance of separation violates the 5-nm standard (the criterion is fixed at 5 nm). Thus, the trajectory model is a one-parameter model that assumes that controller safety margins, and all the associated variability in decision making, may be expressed in terms of the uncertainty associated with controllers’ estimates of aircraft trajectory.

In contrast, the hybrid model assumes that controller safety margins reflect both variability along the projected trajectories of aircraft and changes in decision thresholds. In this manner, the hybrid model has a second parameter—a criterion \( c \). Controller intervention occurs if the minimum distance of separation, taking into account the potential minimum and maximum positions of aircraft, violates the particular criterion value that is set by the controller: Intervene if minimum distance \( < c \); otherwise do not intervene.

### Model Calibrations and Results

The generated decisions in each model are stochastic and binary. We determined an estimate of the probability of intervention by calculating the mean of a large number of decisions (1,000 model simulations). Best-fitting parameters for each model were calculated by performing an exhaustive parameter sweep and selecting the values that minimized the sum squared error between the model’s predictions and controller intervention. The SDT and hybrid models each had two free parameters \( (c \) and \( \sigma ) \) fitted to 72 lateral problem data points in Experiment 1 and 24 lateral problem data points in Experiment 2. The trajectory model had one free parameter \( (\sigma) \) fitted to the same data points.

In making decisions about lateral conflict, our performance theory assumes that experts apply larger safety margins than trainees and that controllers would adopt larger safety margins following high-workload instructions than low-workload instructions. The final values for \( c \) and \( \sigma \) reported in Table 1 are consistent with these assumptions. These values also consistently vary across experiments, with \( c \) and \( \sigma \) larger following a high-workload instruction in Experiment 2 than for experts in Experiment 1 and \( c \) and \( \sigma \) larger following a low-workload instruction in Experiment 2 than for experts in Experiment 1.

Table 2 presents the model fits as a function of root-mean-square error. The root-mean-square error index represents the average error of fit to each individual conflict detection problem, expressed in the same units as the dependent variable of the probability of intervention. The trajectory model better fit the data than the SDT model. In turn, the hybrid model provided

### Table 1

<table>
<thead>
<tr>
<th>Condition</th>
<th>Group or condition</th>
<th>Signal detection theory model</th>
<th>Trajectory model</th>
<th>Hybrid model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>( c )</td>
<td>( \sigma )</td>
<td>( c' )</td>
</tr>
<tr>
<td>Experiment 1</td>
<td>Experts</td>
<td>14</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Trainees</td>
<td>8</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>Experiment 2</td>
<td>High instruction</td>
<td>17</td>
<td>7.5</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Low instruction</td>
<td>12.5</td>
<td>6</td>
<td>5</td>
</tr>
</tbody>
</table>

*The criterion for the trajectory model was fixed at 5 nm.*
a better fit than the trajectory model. Note that for trainees in Experiment 1, the best-fitting parameter for the trajectory model was identical to that of the hybrid model. Thus, these models did not systematically differ in their predictions for trainees in Experiment 1. It is evident in Table 2 that the models better fit the Experiment 2 data than the Experiment 1 data. One reason for this is that the models were required to fit more individual problems in Experiment 1 (72 problems because of the counterbalancing of the time to point of closest approach) than in Experiment 2 (24 problems). In addition, there were smaller groups (13 experts and 7 trainees fitted separately) in Experiment 1 than in Experiment 2 (N = 23). To explore systematic differences in the performance of the three models, Figures 5, 6, 7, and 8 compare the controller intervention and each model’s output across minimum distance and angle. As can be seen in these figures, the trajectory and hybrid models predicted that the probability of intervention increased as angle decreased, whereas the SDT model did not.

### Discussion

The previous sections described three computational models derived from our theory that were used to fit data extracted from controllers. The close fit of the models demonstrates that the predictions derived from the assumption of our performance theory that controllers use safety margins to detect conflicts are consistent with controller behavior. In addition, the finding that the trajectory and hybrid models better fit the controller intervention data than the SDT model supports our more specific assumption that controllers place uncertainty in the projected trajectory of aircraft.

As illustrated in Figure 9, angle has a differential effect on conflict decisions made by the trajectory and hybrid models because of the direction of projected uncertainty in aircraft position when the aircraft are at their closest distance. Given two aircraft at their minimum distance of 10 nm, when aircraft are at a 45° approach angle, their potential minimum distances are close. In contrast, at 135° these projected possible locations are less close. Because the decision to intervene is based on whether any of the potential minimum and maximum positions of aircraft violate separation standards, more interventions will occur at a 45° angle than at a 135° angle. In this manner, the trajectory and hybrid models exhibited emergent behavior consistent with controller performance that was not explicitly programmed into the models or predicted to occur, allowing them to explain some of the less intuitive characteristics of conflict detection behavior. In contrast, to predict the effects of angle, the criterion in an explicit criterion model would need to be changed with each angle, making it less parsimonious.

The trajectory model clearly outperformed the SDT model because it provided a superior fit to the data using one less parameter. In contrast, we draw more tentative conclusions regarding the relative performance of the trajectory and hybrid models. Although the hybrid model provided a better fit than the trajectory model, it required two parameters to do so. Thus, the extra fit provided by the hybrid model came at a cost to parsimony.

Although the hybrid and trajectory models capture the controller data well, there are at least three limitations. First, we do not know the variability in vertical trajectory that experts estimate because intervention did not decrease significantly with minimum vertical separation. There must be some limit to uncertainty placed in the vertical plane. If experts intervened with all aircraft in altitude transition, they would be proactive but overloaded with control tasks (Hollnagel & Woods, 2005; Rouse et al., 1993). Future research needs to present scenarios in which minimum vertical separations are greater than 4,000 ft (1,219.2 m). A second limitation is that the models attempt to capture the whole decision-making process, including the use of tools. A more complete process model may segregate the various decisions involved in, first, deciding whether to use tools and then having separate models for decisions made through perceptual judgments and those made through use of tools (that may be more accurate). Finally, the models assume that controllers use a single strategy. The strategy implemented in our model is a variant of the cognitive motion extrapolation strategy (DeLucia & Liddell, 1998), in that it assumes that controllers assess the distance between aircraft at their point of closest approach. An alternative strategy that controllers report involves assessing the distance or time separating the aircraft at the crossing point (Neal & Kwantes, in press). Controllers may use different strategies, and the predictions of the models may differ in subtle ways. Additional data will be required to calibrate and test such models.

### General Discussion

The decision-making components underlying expertise in ATC conflict detection are not well understood. Building theories of expertise in such work domains requires understanding of experts’ goals and the environmental context in which they make decisions. We developed a performance theory of conflict detection and used this theory as a framework for a model that provided a good fit to controller intervention data. In this section, we evaluate the evidence that supports our theory, identify further research questions that need to be addressed, and provide a more general evaluation of the implications for the human performance modeling and expertise literatures.

In developing our theory, we started with specifying the structure of the environment and the goals of controllers (Marr, 1982). Drawing on the ecological rationality tradition (Simon, 1956), we argued that mental uncertainty and the structure of the environment necessitate that controllers both approximate aircraft trajectory and set decision criteria to favor safety over

![Table 2: Model Fits (Root Mean Square Error) for the Signal Detection Theory Model, Trajectory Model, and Hybrid Model](image-url)
accuracy (Bisseret, 1981) to ensure aircraft separation. The two models based on this theory predicted controller intervention across minimum distance and angle by assuming that experts apply larger safety margins than trainees. In Experiment 2, these models predicted intervention across minimum distance and angle by assuming controllers applied larger safety margins following high- compared with low-workload instructions and, in doing so, demonstrated that these safety margins are under the strategic control of controllers.

Remarkably, experts almost always intervened in vertical problems, even those with minimum vertical separation four times the required standard. This empirical demonstration that controllers apply greater uncertainty bounds to projected aircraft climb rates than to speed rates is crucial. This finding is consistent with reports that ratings of complexity by controllers when detecting conflicts are typically larger when aircraft are transitioning through the altitudes of other aircraft (Boag et al., 2006; Lamoureux, 1999). In addition, Stankovic et al. (2008) reported that vertical separation accounted for little variance in conflict judgments after lateral separation had been entered into their regression model. One likely explanation is that the ATC interface and tools provide limited assistance for projection in the vertical plane. The fact that our controllers used tools less often for vertical than for lateral problems is consistent with this notion.

In the preceding sections of this article, we have argued that controller safety margins differ as a function of the individual experiences of controllers—with respect to their memory for

![Figure 5. Probability of intervention by experts in lateral problems in Experiment 1 (top left panel) and the output of the signal detection theory (SDT) model (top right panel), hybrid model (bottom left panel), and trajectory model (bottom right panel). Error bars represent standard errors of proportion. nm = nautical mile.](image-url)
the dynamics of the environment. However, it is also likely that safety margins are influenced by the manner in which aircraft trajectory information is acquired. A significant part of expertise may be learning to trade off information cost against information utility adaptively (Fu & Gray, 2006; Gray & Fu, 2004). In support of this, we found controllers made faster decisions and were less likely to use tools following high-than low-workload instructions. As a result of acquiring less information, controllers may have been less certain of aircraft trajectory and more likely to intervene. In contrast, no such differences in response time would be expected if differences in intervention as a function of workload instructions were driven solely by memory-driven estimates of how aircraft would transit through sectors or purely by shifts in decision criteria (Bisseret, 1981).

It is useful to compare our approach to statistical models (Averty et al., 2008; Bisantz & Pritchett, 2003; Stankovic et al., 2008). In these approaches, conflict probability is predicted as a function of variables such as minimum lateral distance, minimum vertical distance, and time to crossing point. These variables can account for up to 50% of variance in conflict detection judgments (Stankovic et al., 2008). However, these models do not describe how these environmental cues are used to make decisions and, as such, are dependent on the data on which they are trained. For example, we have demonstrated that changes in intervention with angle may be explained in terms of the

**Figure 6.** Probability of intervention by trainees in lateral problems in Experiment 1 (top left panel) and the output of the signal detection theory (SDT) model (top right panel), hybrid model (bottom left panel), and trajectory model (bottom right panel). Error bars represent standard errors of proportion. nm = nautical mile.
direction of the propagated uncertainty that differentially influences the potentially perceived minimum distance. Because variables such as minimum lateral distance and time to crossing point were held constant across our manipulations of angle, it is difficult to see how the statistical approach can explain such sensitivities in controller behavior. Another benefit of exploring explicit process models is that certain predictions can be made beyond the current data set that can be tested to further constrain the models. For example, if uncertainty in aircraft position was set to follow some form of monotonically increasing function, the hybrid and trajectory models would predict more intervention with increased time to crossing point.

In efforts to improve the capacity of air traffic management systems (Kopardekar & Magyartis, 2003; Majumdar, Ochieng, McAuley, Lenzi, & Lepadatu, 2004), human performance models are being developed to allow analysts to make predictions regarding the likely effects of different types of design concepts or interventions on controller performance, workload, and system safety (Corker, Gore, Fleming, & Lane, 2000; Ravinder, Remington, & Lee, 2005; Remington, Matessa, Freed, & Lee, 2003). However, these models do not yet include components that describe the cognitive processes involved in conflict detection. The theory presented here provides the initial information needed to develop a controller conflict detection performance model.

We focused on the processes that controllers use to make conflict judgments once they have selectively attended to an aircraft pair. Conflict detection also requires controllers to

![Figure 7. Probability of intervention by controllers in lateral problems in Experiment 2 following a high-workload instruction (top left panel) and the output of the signal detection theory (SDT) model (top right panel), hybrid model (bottom left panel), and trajectory model (bottom right panel). Error bars represent standard errors of proportion. nm = nautical mile.](image-url)
search ATC displays for potential conflict pairs (Galster et al., 2001; Metzger & Parasuraman, 2001; Remington, Johnston, Ruthruff, Gold, & Romera, 2000). A crucial next step in the development of our performance theory is to combine our model with the implementation of a monitoring–scanning controller component (e.g., see Leiden, Kopardekar, & Green, 2003; Niessen, Eyferth, & Bierwagen, 1999). An understanding of how visual search processes interact with the aircraft trajectory estimation processes studied in this article will require aircraft pairs to be embedded in the context of surrounding (filler) aircraft and a representative range of additional control tasks (e.g., accepting or handing off aircraft or communicating with pilots).

Our workload instructions were not intended to substitute for workload as experienced in the field, as would traditionally be experimentally induced by introducing secondary tasks or increasing traffic count (Galster et al., 2001; Remington et al., 2000). Changes in intervention probability that we observed as a function of our workload instructions are likely to reflect demand characteristics. This was intended because it allowed us to test the model’s validity by inducing shifts in controller safety margins. One final limitation is that controllers were not nearly as familiar with the sector as they would be in the field, and as such our model is more likely to explain performance when controllers are working on novel airspace. The types of problems that controllers encounter are highly contextualized,

Figure 8. Probability of intervention by controllers in lateral problems in Experiment 2 following a low-workload instruction (top left panel) and the output of the signal detection theory (SDT) model (top right panel), hybrid model (bottom left panel), and trajectory model (bottom right panel). Error bars represent standard errors of proportion. nm = nautical mile.
and the internalization of regularities in ATC environments is likely to have a strong memory component (Loft, Humphreys, & Neal, 2004; Loft, Neal, & Humphreys, 2007). With increasing experience, controllers may draw on their memory for aircraft trajectory or their memory for how they have previously prioritized their resources for specific air traffic configurations.

Understanding expertise is especially relevant for organizations designing training programs (Ericsson, 2005; Haerem & Rau, 2007). A recent special issue of the Journal of Experimental Psychology: Applied (Ericsson & Williams, 2007) presented several articles that focused on the theme of how best to capture expert performance. Work domains such as ATC present a unique set of challenges. One challenge we faced was that past research defined expertise in terms of the sophistication of problem representations and information-processing strategies (Chi et al., 1981; Ericsson & Charness, 1994; Larkin & McDermott, 1980). An implicit assumption made here is that expertise can only be attained with complete knowledge of the task environment. In many work domains, however, information-seeking costs and time pressure means that experts do not necessarily build complex representations of their environment. In this manner, work environments such as ATC provide a major source of constraint on the development of theories of expertise and, as we have seen in this study, can be a source of novel empirical predictions.

Although ATC provided a useful test for our theory, we have no reason to believe that the basic principles in our theory are specific to the ATC domain but rather are relevant to any domain in which individuals make decisions under uncertainty and time pressure. Continued investigation into how the properties of operational task environments interact with the cognition of experts should provide further opportunity to develop evidence-based practices and superior professionals in relevant fields of application.

References


(Appendix follows)
### Appendix

Probability That Controllers Intervened in Experiment 2 as a Function of Expertise, Problem Type, Instruction, and Angle

<table>
<thead>
<tr>
<th>Lateral problems</th>
<th>0/1 nm</th>
<th>2/4 nm</th>
<th>6/8 nm</th>
<th>10/12 nm</th>
<th>14/16 nm</th>
<th>18/20 nm</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Experts (N = 14)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High instruction</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>45</td>
<td>1 (.0)</td>
<td>1 (.0)</td>
<td>1 (.0)</td>
<td>.79 (.08)</td>
<td>.79 (.08)</td>
<td>.36 (.10)</td>
</tr>
<tr>
<td>135</td>
<td>1 (.0)</td>
<td>1 (.0)</td>
<td>.93 (.05)</td>
<td>.71 (.09)</td>
<td>.43 (.10)</td>
<td>.21 (.08)</td>
</tr>
<tr>
<td>Low instruction</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>45</td>
<td>1 (.0)</td>
<td>.93 (.05)</td>
<td>1 (.0)</td>
<td>.64 (.10)</td>
<td>.57 (.10)</td>
<td>.21 (.08)</td>
</tr>
<tr>
<td>135</td>
<td>1 (.0)</td>
<td>1 (.0)</td>
<td>.86 (.07)</td>
<td>.64 (.10)</td>
<td>.29 (.09)</td>
<td>.14 (.07)</td>
</tr>
<tr>
<td><strong>Trainees (N = 9)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High instruction</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>45</td>
<td>1 (.0)</td>
<td>1 (.0)</td>
<td>.89 (.06)</td>
<td>.89 (.06)</td>
<td>.67 (.10)</td>
<td>.56 (.10)</td>
</tr>
<tr>
<td>135</td>
<td>1 (.0)</td>
<td>1 (.0)</td>
<td>.93 (.05)</td>
<td>.67 (.10)</td>
<td>.33 (.10)</td>
<td>.11 (.06)</td>
</tr>
<tr>
<td>Low instruction</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>45</td>
<td>1 (.0)</td>
<td>1 (.0)</td>
<td>.89 (.06)</td>
<td>.78 (.08)</td>
<td>.33 (.10)</td>
<td>.22 (.08)</td>
</tr>
<tr>
<td>135</td>
<td>1 (.0)</td>
<td>1 (.0)</td>
<td>.56 (.10)</td>
<td>.33 (.10)</td>
<td>.22 (.08)</td>
<td>.11 (.06)</td>
</tr>
<tr>
<td>8 nm vertical problems</td>
<td>0/200 ft</td>
<td>400/800 ft</td>
<td>1,200/1,600 ft</td>
<td>2,000/2,400 ft</td>
<td>2,800/3,200 ft</td>
<td>3,600/4,000 ft</td>
</tr>
<tr>
<td>Experts (N = 14)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High instruction</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>45</td>
<td>.86 (.07)</td>
<td>1 (.0)</td>
<td>.79 (.08)</td>
<td>.79 (.08)</td>
<td>.86 (.07)</td>
<td>.79 (.08)</td>
</tr>
<tr>
<td>135</td>
<td>.93 (.05)</td>
<td>.79 (.08)</td>
<td>.79 (.08)</td>
<td>.86 (.07)</td>
<td>.79 (.08)</td>
<td>.86 (.07)</td>
</tr>
<tr>
<td>Low instruction</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>45</td>
<td>.89 (.06)</td>
<td>1 (.0)</td>
<td>.78 (.08)</td>
<td>.89 (.06)</td>
<td>.89 (.06)</td>
<td>.78 (.08)</td>
</tr>
<tr>
<td>135</td>
<td>.89 (.06)</td>
<td>.67 (.10)</td>
<td>.78 (.08)</td>
<td>.78 (.08)</td>
<td>.89 (.06)</td>
<td>.78 (.08)</td>
</tr>
<tr>
<td>Trainees (N = 9)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High instruction</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>45</td>
<td>.67 (.10)</td>
<td>.78 (.08)</td>
<td>.67 (.10)</td>
<td>.67 (.10)</td>
<td>.67 (.10)</td>
<td>.67 (.10)</td>
</tr>
<tr>
<td>135</td>
<td>.56 (.10)</td>
<td>.44 (.10)</td>
<td>.44 (.10)</td>
<td>.44 (.10)</td>
<td>.44 (.10)</td>
<td>.44 (.10)</td>
</tr>
<tr>
<td>Low instruction</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>45</td>
<td>.56 (.10)</td>
<td>.44 (.10)</td>
<td>.44 (.10)</td>
<td>.44 (.10)</td>
<td>.44 (.10)</td>
<td>.44 (.10)</td>
</tr>
<tr>
<td>135</td>
<td>.33 (.10)</td>
<td>.11 (.06)</td>
<td>.11 (.06)</td>
<td>.11 (.06)</td>
<td>.11 (.06)</td>
<td>.11 (.06)</td>
</tr>
</tbody>
</table>

Note. Standard errors of proportion are in parentheses. nm = nautical miles.

Received April 1, 2008
Revision received March 27, 2009
Accepted March 30, 2009

124 LOFT, BOLLAND, HUMPHREYS, AND NEAL