Attentional Models of Multitask Pilot Performance Using Advanced Display Technology

Christopher D. Wickens and **Juliana Goh,** University of Illinois, Savoy, Illinois, **John Helleberg,** MITRE Corporation, McLean, Virginia, and **William J. Horrey** and **Donald A. Talleur,** University of Illinois, Savoy, Illinois

In the first part of the reported research, 12 instrument-rated pilots flew a high-fidelity simulation, in which air traffic control presentation of auditory (voice) information regarding traffic and flight parameters was compared with advanced display technology presentation of equivalent information regarding traffic (cockpit display of traffic information) and flight parameters (data link display). Redundant combinations were also examined while pilots flew the aircraft simulation, monitored for outside traffic, and read back communications messages. The data suggested a modest cost for visual presentation over auditory presentation, a cost mediated by head-down visual scanning, and no benefit for redundant presentation. The effects in Part 1 were modeled by multiple-resource and preemption models of divided attention. In the second part of the research, visual scanning in all conditions was fit by an expected value model of selective attention derived from a previous experiment. This model accounted for 94% of the variance in the scanning data and 90% of the variance in a second validation experiment. Actual or potential applications of this research include guidance on choosing the appropriate modality for presenting in-cockpit information and understanding task strategies induced by introducing new aviation technology.

INTRODUCTION

Two new technologies proposed for the advanced cockpit are the cockpit display of traffic information (CDTI), designed to enhance pilots' awareness of nearby traffic (Wickens, Helleberg, & Xu, 2002), and the data link communications system, designed to provide digitally uplinked communications from air traffic control to the pilot (Kerns, 1999; Navarro & Sikorski, 1999). The two systems have much in common: Both are entering a phase of preliminary in-flight testing, rely on advanced technology, and are undergoing extensive human factors evaluations. In particular, both systems will change a function that has been traditionally carried out by air traffic control (ATC) communications – and hence has involved auditory (speech) information in the cockpit – to one that is computer mediated and will involve visual (display) information. This alteration could have major implications for a single-pilot aircraft, in which visual attention is already heavily burdened by responsibilities of instrument-panel scanning and outside-world monitoring.

One function of the CDTI will be to help pilots understand where traffic outside can be spotted and thus aid them in calling out "traffic in sight" by replacing the traditional role of ATC in guiding attention through oral instructions (e.g., "watch for traffic, 10:00 high, 2 miles out"). In the case of the data link, the intention is to provide a visual text version of instructions (e.g., "climb to flight level 220") to replace the oral communications from ATC, which have proven vulnerable to working memory failures when messages are long.

The change from auditory to visual representation of traffic and communications information in an already-busy visual environment has

Address correspondence to Christopher D. Wickens, University of Illinois, Aviation Human Factors Division, 1 Airport Rd., Savoy, IL 61874; cwickens@psych.uiuc.edu. **HUMAN FACTORS,** Vol. 45, No. 3, Fall 2003, pp. 360–380.

important implications for the single pilots' limited visual attentional resources. In this two-part paper, we consider the implications of this change for two aspects of attention: the role of multiple resources (defined by auditory and visual modalities) in characterizing the pilot's divided attention and the role of optimal scanning models in characterizing the pilot's selective attention. In Part 1 we describe an experiment that addresses the first of these issues by comparing auditory and visual delivery of both traffic and communications information. In Part 2 we address the second issue by evaluating an optimal model of visual scanning in the traffic-monitoring phase of the experiment and in two related experiments for which the data are reported elsewhere. We first review the literature that bears on each of these parts.

Modality Differences in Information Delivery

At one level of abstraction, the pilot's tasks can be modeled as depending on two generic sources of visual information. In order to aviate, the highest-priority task (maintaining stability and keeping the aircraft from stalling), the pilot must process visual information from the instrument panel as well as (in good weather) from the relative orientation or attitude of the true horizon viewed outside. To navigate, the second-priority task, the pilot must also process information from the instrument panel, maps, specialized navigational instruments, and the view outside to identify both hazardous objects to be avoided (other air traffic, terrain) and objects to seek (e.g., a runway; Wickens, 2003). Against this two-task backdrop of visual (V) information-processing demands, the pilot may also have tasks of lower priority (we refer to these here as *side tasks*) that can be accomplished visually (e.g., reading checklists) or auditorily (A; e.g., listening to ATC communications). The lower-priority status of such tasks does not mean that they are unimportant – rather, that if they conflict with a higher-priority task, the latter should normally take precedence (Schutte & Trujillo, 1996).

An extensive line of dual-task research generally suggests an advantage of mixed-modality (AV) over intramodality (VV) presentation in basic laboratory tasks as well as in more applied flying and driving simulations (Parkes & Coleman, 1990; Wickens, Sandry, & Vidulich, 1983). Generally, such research can be well explained by the concept of multiple processing resources (Navon & Gopher, 1979; Wickens, 1991, 2002), whereby the visual channels are supported by resources that are somewhat independent from those involved with auditory processing. Hence multitask processing can be supported more efficiently by cross-modal presentation than by within-modal presentation (i.e., AV is superior to VV). According to such a view, the replacement of ATC voice (AV) with digital in-cockpit displays (VV) is actually a regressive move toward less efficient time-sharing performance.

In contrast to multiple-resource predictions, some research has suggested that auditory presentation of side task information can be detrimental because of the phenomenon of "preemption" (Damos, 1997; Latorella, 1998; Wickens, Dixon, & Seppelt, 2002). According to such a view, which is supported by a good deal of empirical work (Dismukes, 2001; Wickens & Liu, 1988), a discrete auditory message is more likely than a visual message to attract attention away from the ongoing visual tasks of higher priority (aviating, navigating) because (a) the auditory channel has inherent attention-capturing properties (Spence & Driver, 2000) and (b) if the message is long, it will be rapidly forgotten from working memory and hence must be attended to immediately. A more permanent visual message, such as a data link text line or an updated CDTI, has no such cognitive urgency associated with it, allowing the pilot to more leisurely complete higher-priority visual tasks, represented in the instrument panel and outside world, before turning to the side task information.

Thus the comparison of AV versus VV delivery of information reveals different patterns of effects, depending on the relative influence of multiple resources versus preemption. Both mechanisms support improved side task performance with auditory (versus visual) delivery, but preemption implies that the primary visual tasks of aviating and navigating will suffer from auditory side task delivery, whereas multiple resources implies that performance of these tasks will improve with auditory delivery. Such improvement should be directly linked to more visual scanning on the sources of information for these primary tasks: the outside world and the instrument panel.

In addition to these mechanisms, two other factors may come into play when auditory and visual side task information delivery are compared. First, single-task auditory delivery may have inherent weaknesses in supporting side task performance. In the context of the tasks examined here, the auditory modality will be a poor delivery channel for long verbal messages, for which working memory limits may be exceeded. Audition is also not an optimal channel for delivering precise spatial information, such as localizing traffic in 3-D space (Wickens et al., 1983). (We do not examine here the effectiveness of 3-D sound localization [Begault & Pittman, 1996] but, rather, consider verbal information delivered from ATC.)

The second factor is the possibility that redundant delivery of both visual and auditory information might provide the "best of both worlds" by eliminating any auditory limitation for a particular task (because the visual modality is available) and at the same time allowing the eyes to focus on the instrument panel most of the time, while relevant auditory information can be processed. Such redundancy gain has been found in the delivery of instructions (e.g., Sweller, Chandler, Tierney, & Cooper, 1990; see Wickens & Hollands, 2000, for a review), although it has been surprisingly absent when investigated in some dual-task contexts (Helleberg & Wickens, 2003; Seagull & Wickens, 2001), as if the combination may produce the "worst of both worlds" rather than the best.

In the particular task context of the aviation flight deck simulation experiment reported here, few investigators have compared the delivery of different modalities of information sources. Wickens et al. (1983) compared auditory and visual delivery of brief text messages, as well as spatial messages, to pilots in a fighter aircraft simulation and noted benefits of auditory delivery to both side task and flight task control, particularly when the auditory message pertained to verbal information. Prinzo (2003) compared auditory (ATC) and visual (CDTI) delivery of traffic location information and observed slight advantages for the latter because of its greater precision. However, because her experiment was carried out in actual aircraft, Prinzo did not collect tracking or navigational measures of whether the primary visual flight task was disrupted differentially by the two modalities. Helleberg and Wickens (2003), who compared visual, auditory, and redundant delivery of data link information in a full mission simulation, observed greater disruption of ongoing visual tasks of aviating (flight path control) and navigating (spotting and calling out "traffic in sight") by the auditory mode than by the visual data link. However, their data link communications strings were sometimes very long, either requiring writing on a clipboard (which directly required visual resources) or, in one condition for which clipboard writing was not permitted, imposing severe costs on working memory.

In summary, few studies have compared the two modalities of delivery of equivalent information in a cockpit environment. One study (Wickens et al., 1983), which employed side tasks different from the data link and CDTI tasks of interest here, found results entirely consistent with multiple-resource predictions (AV > VV for both primary and secondary tasks). Prinzo (2001) did examine the traffic-sighting task as supported by auditory versus visual delivery but did not assess dual-task interference with the highest-priority task of aviating. Helleberg and Wickens (2003) carried out such an assessment, but only in the context of a data link display, and observed that the auditory mode caused preemption of the higher-priority flight tasks, as observed by Latorella (1998). However, excessively lengthy communications messages used on some of their trials may have placed the auditory modality at a distinct disadvantage.

In the current experiment we hope to fill in the gaps of this knowledge. We had skilled pilots fly a full mission simulator while monitoring for traffic (calling out "traffic in sight"), maintaining tight flight path control, and, concurrently, either processing information about traffic (from a CDTI, from ATC call outs, or from both redundantly) or processing communications information about required flight path trajectories (from a data link display, from ATC, or redundantly). These two side tasks (traffic and communications) occurred during different, nonoverlapping phases of the flight legs. In order to determine if differences in modality effectiveness were modulated by workload, we also had the flight legs differ in workload imposed: by more or less traffic (four vs. one aircraft) and by shorter or longer communications strings (one vs. three chunks). We predict that neither side task (traffic call out, communications retention) would be hurt by auditory presentation in the dual-task context and, indeed, that they may be helped. The higher-priority visual tasks of aviating (measured by flight path tracking error) and navigating (monitoring for traffic that is not announced by ATC or data link), however, would be helped by auditory (vs. visual) delivery to the extent that multiple resource mechanisms come into play or would be harmed by it to the extent that preemption mechanisms come into play. As noted, previous data are ambiguous on this point.

Our efforts to diagnose the nature of attentional effects are supported by assessing visual scanning across four primary areas of interest: the instrument panel (IP) supporting aviating, the outside world (OW) supporting navigating (here operationally defined by the visual sighting of traffic hazards that need to be avoided), the CDTI supporting navigating, and the data link display supporting communicating. Similar methodology was employed by Wickens, Helleberg, et al. (2002). In Part 2 of this article, we use these scanning data to validate an expected value computational model of how selective attention is driven by the priority of the relevant tasks, supported by the different visual areas of interest. We compare how well the model, developed on previous data (Wickens, Helleberg, et al., 2002), is validated with the current data and with eye movement data from a second experiment (Helleberg & Wickens, 2003).

PART 1: EXPERIMENTAL SIMULATION – METHODS

Participants

Twelve certified flight instructors (10 men, 2 women) were recruited from the Institute of Aviation at the University of Illinois and from an aviation safety seminar sponsored by the Federal Aviation Administration to participate in the study. The pilots' ages ranged from 21 to 60 years (M = 41.7 years). Their total flight hours ranged from 200 to 3700 hr (M = 1183.8 hr), and their instrument flight time ranged from 20 to 300 hr (M = 96.8 hr). The pilots were paid $10/hr for their participation and a $10 bonus at the end of the experiment.

Equipment and Displays

Flight simulator. Pilots flew in a Frasca 142 flight simulator that was configured as a single-engine Beechcraft Sundowner. The simulator consisted of a full instrument panel and a radio stack as well as the standard aircraft controls (yoke, throttle, and rudder pedals). An Evans and Sutherland SPX 2400 visual system was used to project a 135° view of the outside world, and traffic could be displayed up to 5 nautical miles away. In addition, simulated ATC instructions were prerecorded and presented to participants via speakers attached to an 80 MHz 386 PC.

Data link displays and CDTIs. A Silicon Graphics workstation and a 20-inch (51-cm) color monitor (screen resolution 1280 × 1024) were used to display the data link text messages and the CDTI. The 2-D coplanar format of the CDTI was identical to that reported in previous research (Wickens, Helleberg, et al., 2002). The data link display subtended horizontal and vertical visual angles of approximately 12° and 8°, respectively, and the CDTI subtended horizontal and vertical visual angles of approximately 10° and 18°, respectively. Figure 1 depicts the locations of the various displays and equipment.

Head-mounted eye/head tracker. An ASL Model 501 head-mounted eye tracker with an integrated magnetic head tracker was used to track pilots' eye movements. Both pupil and corneal reflections were sampled at 60 Hz with an accuracy of better than 1°, and the head tracker allowed six degrees of freedom of movement for the head.

Task

Pilots flew six instrument flight rules (IFR) cross-country flights (each 30 min long) under visual meteorological conditions (VMCs). Each flight consisted of 11 legs, beginning with a "communications" leg followed by a "traffic" leg. This alteration sequence continued until all 11 legs were completed. Figure 2 shows a graphical representation of a typical flight used in the experiment.

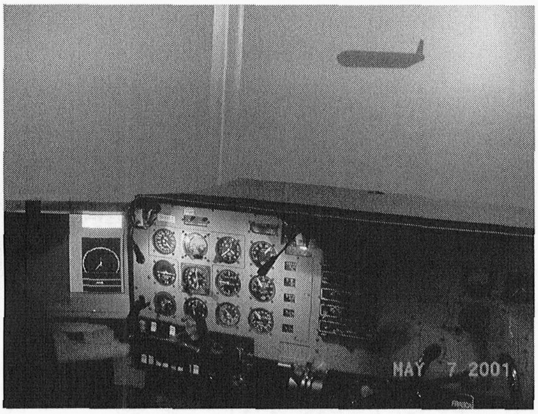

Figure 1. Frasca instruments, data link and CDTI displays, and traffic close-up. The CDTI is shown on the black display to the left of the instrument panel; the data link message appears in the white rectangle just above that.

Communications legs. Pilots flew six communications legs during each flight. The pilots began each flight at cruising altitude, after which ATC instructions (heading, altitude, and airspeed) were presented. ATC instructions were always preceded by an alerting tone as a signal to pilots. The pilots were required to first read back the ATC instructions (e.g., "turn left heading 320, climb to 8000 feet, increase airspeed to 140 knots") and then to maneuver the aircraft accordingly. The ATC instructions were presented in one of three display formats throughout each flight: (a) auditorily (verbal instructions via speakers, no data link display), (b) visually (via data link display, no verbal instruction), or (c) redundantly (both data link display and verbal instruction). Half the instructions were three parameters in length (high workload) and the other half were one parameter in length (low workload). However, all flights began with a three-parameter ATC instruction to put pilots on a specific course of flight. Pilots were allowed to have ATC instructions repeated to them, and the experimenter corrected pilots who made errors during read back or maneuvering.

Traffic legs. Each flight included five traffic legs, during which pilots encountered either one (low workload) or four aircraft (high workload), in an unpredictable order, and which were flying level at an unpredictable altitude from an unpredictable location. Under the auditory and redundant conditions, the simulated ATC provided traffic point outs (e.g., "traffic to your upper right"). To reduce predictability further, we varied the timing of these verbal point outs relative to the time at which the traffic became

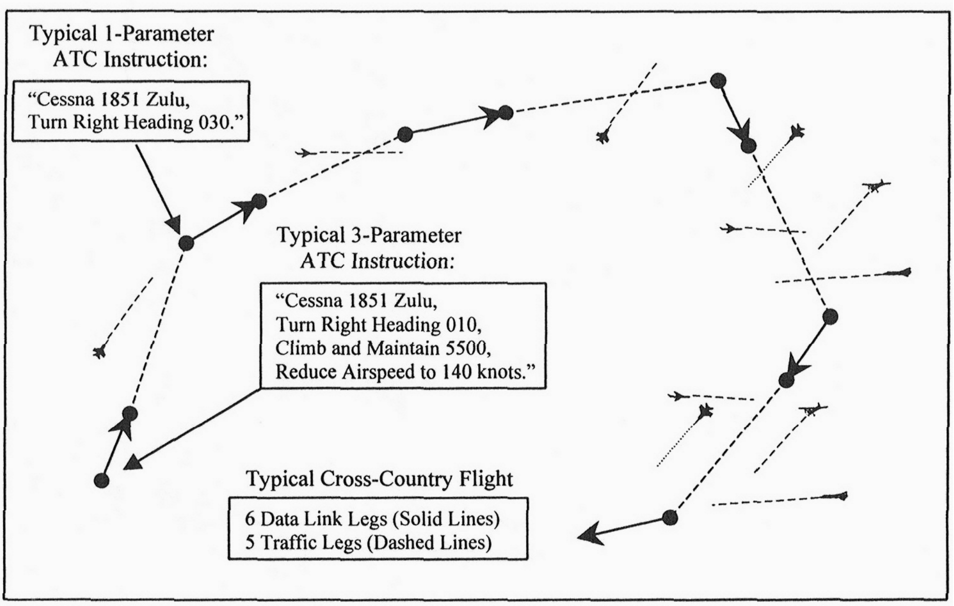

Figure 2. Typical cross-country flight.

visible in the outside world. These point outs occurred either at the time when the traffic became visible in the outside world or approximately 10 s before or after. Under the visual condition, traffic information was available only through the CDTI and from visually scanning the outside world. Pilots were told to scan the outside world and to use the CDTI to locate traffic and call them out ("traffic in sight"). Pilots were also told that they had primary responsibility (i.e., ATC did not provide traffic avoidance vectors) in maintaining separation and maneuvering away from traffic that appeared to be on a collision course with their aircraft.

During each flight, pilots encountered one aircraft that would collide with their aircraft if avoidance maneuvers were not made. This conflict aircraft could appear under either the low- or high-workload condition. All other traffic during the flight were nonconflict aircraft. In addition, in the second-to-last flight, pilots were presented with one aircraft that was neither depicted on the CDTI nor pointed out by ATC. This "rogue" aircraft came into conflict with the pilot's flight path for half the participants but was not a threat to the other half of the participants. This rogue aircraft was presented to simulate a situation in which an aircraft has its transponder off or in which there is a malfunction in the CDTI.

Procedure

Pilots were allowed either to complete all six flights in one session of 4 to 5 hr, with a mandatory break of 10 to 15 min after the first three flights (10 pilots), or to return on a second day to complete the last three flights (2 pilots). Pilots first read and signed consent forms and filled out a demographic questionnaire (e.g., age, total flight hours). They were then allowed as much time as needed to read the instructions for the experiment. This was followed by a practice trial, consisting of a data link communications leg followed by a traffic leg. Pilots familiarized themselves with the flight simulator, displays, and other equipment during the practice trial and were shown an example of traffic in the outside world. The pilots were reminded that their three main tasks were to read back ATC instructions, maneuver the plane according to ATC instructions, and scan available displays, instruments, and the outside world to locate and call out all traffic. Pilots were fitted with an eye tracker, and calibration was carried out before the start of the experiment and again each time the apparatus was taken off. After completing all six flights, pilots were compensated and thanked for their participation. There appeared to be no differences in performance between the 2 pilots who took 2 days to complete the experiment and the other 10 pilots.

Experimental Design

The main factors of interest in this study were (a) display format (visual, auditory, redundant) and (b) traffic load (one or four aircraft) during traffic legs and (a) display format (visual, auditory, redundant) and (b) communications load (one or three ATC parameters) during communications legs. The experiment was run as a complete within-subjects design, and the order of presentation of factors was counterbalanced using a Latin square. Because the communications legs and the traffic legs were not performed concurrently, the analysis was carried out in two 3 × 2 repeated-measures designs. Each pilot was presented one rogue aircraft on the second-to-last trial. Thus there were only four observations of the rogue detection for each of the three display conditions.

RESULTS: TRAFFIC DETECTION PHASE

In the following analyses, trials in which pilots maneuvered to avoid traffic, encountered conflicting traffic, or experienced rogue traffic were not included in the analyses of lateral and vertical error, visual scan parameters (percentage dwell time and mean dwell duration), accuracy of target detection, or time to detect traffic. This resulted in only 8 participants having full sets of data. Means from the available data were used to replace the missing data to increase power, except in the analyses of lateral and vertical error.

Flight Path Tracking Performance

Lateral tracking. Lateral tracking performance (deviations in degrees from target heading), shown in the top panel of Figure 3, revealed a main effect of traffic load on heading error, $F(1, 7) = 140$, $p < .01$. Although there was no effect of modality, $F(2, 14) = 1.444$, $p > .10$, the

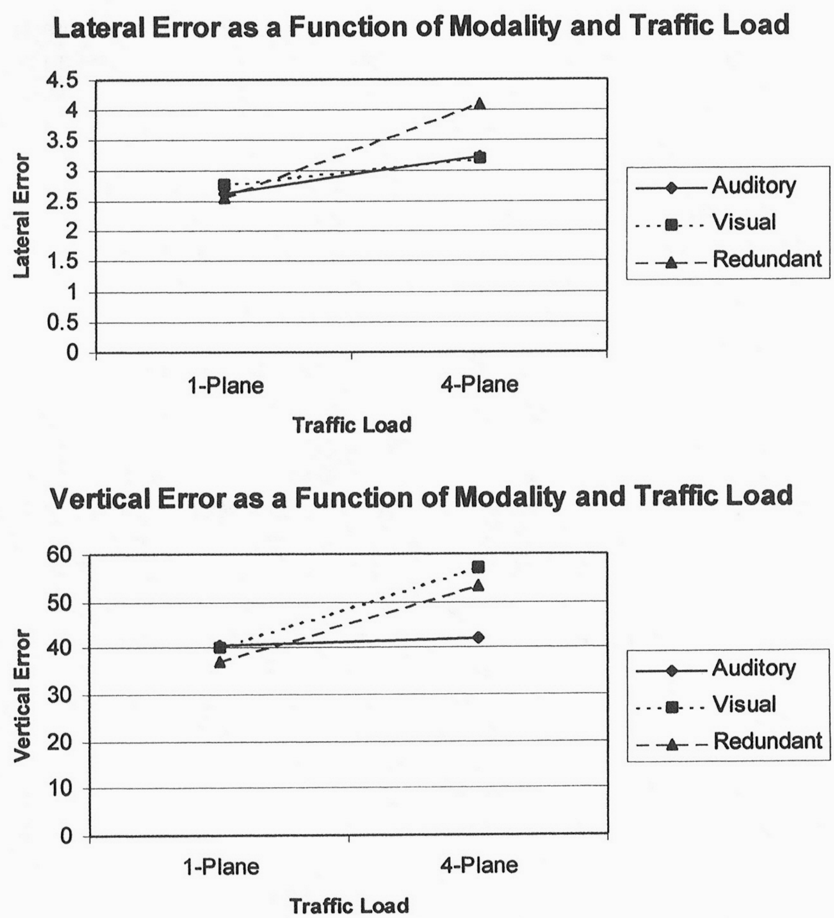

Figure 3. Top: Heading RMS error (degrees off target heading) as a function of modality and traffic load. Bottom: Vertical RMS error (feet) as a function of modality and traffic load.

Traffic Load × Modality interaction, $F(2, 14) = 3.698$, $p = .05$, revealed that the increase in heading error with high traffic load was much greater (more than twice as great) in the redundant AV condition than in the two single-modality conditions (A and V).

Vertical tracking. The results of the analysis of altitude error, shown in the bottom panel of Figure 3, revealed a significant main effect only of traffic load, $F(1, 7) = 6.06$, $p = .04$. A marginally significant effect of modality, $F(2, 14) = 2.83$, $p = .09$, suggested that vertical tracking was more disrupted by the two CDTI conditions (V and AV) than by the auditory condition. Although the Modality × Traffic Load interaction was not significant ($F = 1.3$), inspection of the data in the bottom panel of Figure 3 reveals that this cost to tracking imposed by the CDTI was observed only at high traffic load. Indeed, a separate one-way analysis of variance (ANOVA) conducted on only the high traffic load data revealed a significant effect of modality on vertical error, $F(2, 14) = 3.69$, $p = .05$.

Time in predicted conflict. In addition to flight path tracking error, another aspect of tracking performance is the extent to which pilots could avoid being in the undesirable state of a conflict with traffic aircraft, predicted within the next 45 s. Analyses of these data revealed a main effect of modality, $F(2, 22) = 7.54$, $p < .01$, suggesting that the two CDTI conditions (V and AV) substantially reduced this time, from an average of 40 s/leg (auditory condition) to an average of 28 s/leg. This variable was not influenced by traffic load via either a main effect or an interaction. Furthermore, on the infrequent occasions when a conflict was present, pilots flying with the auditory call out chose the appropriate (safest) maneuver only 50% of the time (12/24 trials), whereas pilots flying with the CDTI chose the safest maneuver 83% of the time (39/47 trials).

Visual Scanning Analysis

Percentage dwell time (PDT). The PDT was a measure of the percentage of time the scan was within each of the areas of interest (AOIs): the outside world (OW), instrument panel (IP), and CDTI. These data averaged across the six conditions and three relevant AOIs are shown in Figure 4 (top panel: low workload; bottom panel: high workload). Because the auditory modality condition had only two AOIs, a single 3 × 3 × 2 ANOVA on all three factors was not conducted. Instead, the following approach employed two separate 2 × 3 × 2 ANOVAs.

The first ANOVA (ANOVA 1), which focused only on the IP and OW as a function of modality (three levels) and traffic load (two levels), revealed the following:

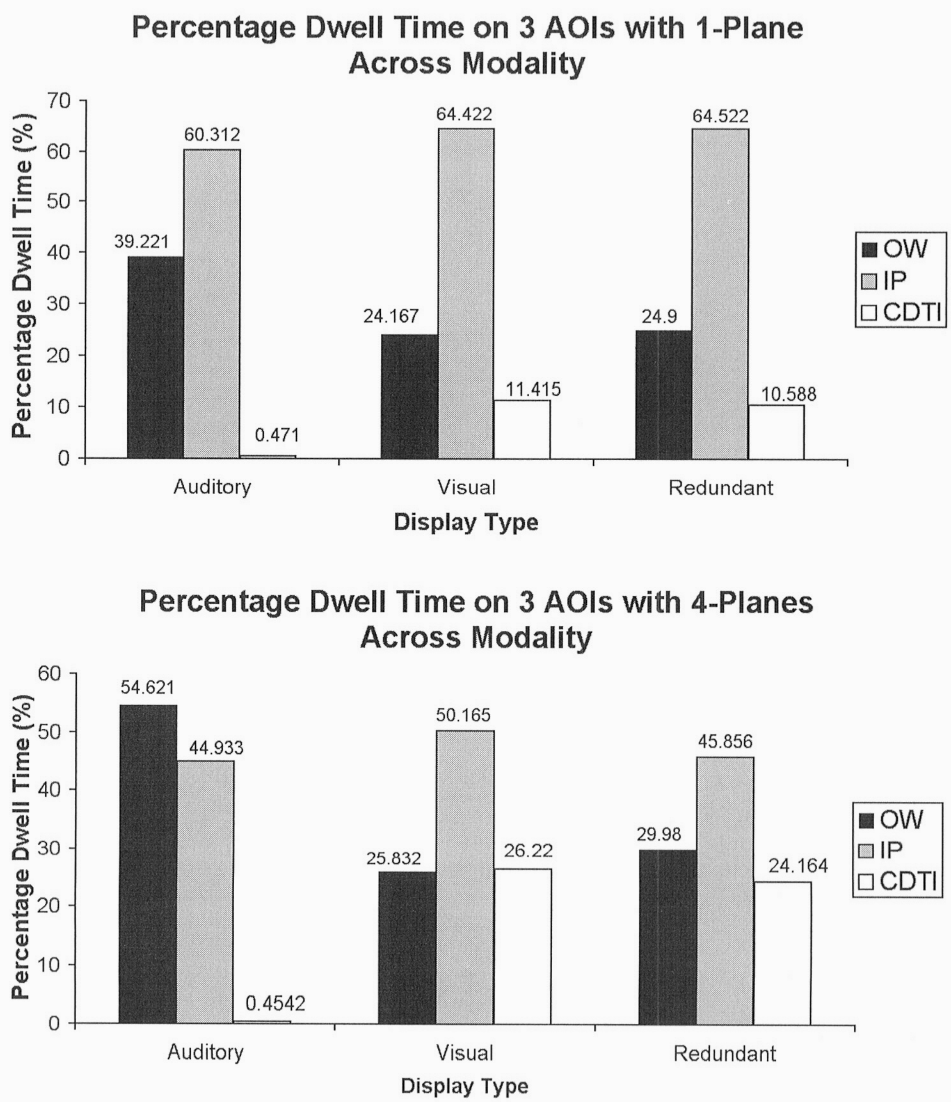

Figure 4. Top: Percentage dwell time on three areas of interest (AOIs) with one plane (low workload) across modality. Bottom: Percentage dwell time on three AOIs with four planes (high workload) across modality. (OW = outside world, IP = instrument panel, CDTI = cockpit display of traffic information.)

First, there was a main effect of AOI, $F(1, 22) = 17.043$, $p < .01$, replicating our previous findings that the IP is fixated more often (55%) than is the OW (33%).

Second, the significant AOI × Modality interaction, $F(2, 22) = 13.768$, $p < .01$, revealed that the OW fixations were about 20% higher for the auditory condition than for the two visual CDTI conditions; this was not surprising, given that in the two CDTI conditions some visual attention is reallocated to the CDTI. Of significance, however, is that this CDTI attention is not "borrowed" from the IP; the scan percentage on this AOI remained relatively unaffected by the presence of the CDTI. Thus, across all modality conditions, pilots appeared to "protect" their visual attention allocation to the IP, a finding replicating that observed by Wickens, Helleberg, et al. (2002).

Third, there was a significant interaction between traffic load and AOI, suggesting a redistribution of scan away from the IP at high traffic load, $F(1, 11) = 42.026$, $p < .01$. The negative effect of this load-imposed reallocation on tracking performance in both lateral and vertical axes can be seen in Figure 3.

Finally, there was a significant three-way interaction among traffic load, AOI, and modality, $F(2, 22) = 3.433$, $p = .05$, suggesting that although there was a reallocation of scan away from the IP with an increase in traffic load for all display types, visual attention was reallocated to the OW in the auditory condition, whereas for the two visual conditions (V and AV) visual attention to the OW remained relatively unchanged by high workload (compare the black bars in the top and bottom panels of Figure 4). This equivalence is not too surprising, given that under the auditory condition pilots could reallocate their scan to the only other AOI available, the OW, whereas in the visual conditions pilots could also reallocate their scan to the CDTI.

The second ANOVA (ANOVA 2) was a 2 × 3 analysis that focused on only the two visual CDTI conditions (visual and redundant) and considered all three AOIs. A highly significant Traffic Load × AOI interaction, $F(2, 22) = 29.604$, $p < .01$, revealed that high traffic load significantly increased CDTI scanning, $t(11) = -7.856$, $p < .01$, significantly reduced IP scanning, $t(11) = 6.201$, $p < .01$, and marginally increased OW scanning ($p = .10$).

Mean dwell duration (MDD). Using the same two ANOVA procedures described for PDT, analysis of the mean dwell duration – the length of time the eyes fixated on a particular AOI before leaving it – revealed a main effect of AOI: ANOVA 1, $F(1, 11) = 8.68$, $p < .01$; ANOVA 2: $F(2, 22) = 28.9$; $p < .01$. We observed longest dwells (mean = 4.22 s) on the IP, intermediate dwells (mean = 2.83 s) on the OW, and shortest dwells (mean = 1.23 s) on the CDTI. The significant Load × AOI interaction, $F(1, 11) = 17.7$, $p < .01$, indicated that increasing traffic load increased dwell duration on the CDTI by 0.8 s and on the outside world by 0.6 s but that it decreased dwell duration on the instrument panel by 1.5 s. Thus these dwell data mirror those observed for the PDT.

A significant main effect of modality in ANOVA 1, $F(2, 22) = 3.59$, $p = .05$, indicated that dwells were longer with the auditory than with the two visual displays. That is, with more places to look (in the visual conditions), pilots presumably shortened their dwells at each place.

Consequences of Scan Changes to Traffic Detection

The effects of format and traffic load on the speed and accuracy of visually sighting the traffic, as mediated by the influence of the scanning, were fairly straightforward. These effects revealed that increased traffic load led to a significant increase in traffic call-out time, from 17.1 to 31 s, $F(1, 11) = 230$, $p < .01$, as well as a significant decrease in accuracy, from 97% to 82%, $F(1, 11) = 18.14$, $p < .01$. Neither modality nor the Modality × Traffic Load interaction modified the response time effect. Although there was no significant effect of modality on call-out accuracy, $F(2, 22) = 2.27$, $p = .13$, separate planned comparisons revealed higher accuracy in the auditory conditions (94%) than in the visual condition (87%), $t(11) = 1.823$, $p = .09$.

These data indicate that the decrease in OW scanning in the two CDTI conditions did not seriously hurt detection of the announced aircraft, relative to the auditory condition. Any cost of that decreased scanning was nearly offset by the increased usefulness of the CDTI for locating the traffic with greater precision. There was, however, a hint that the auditory ATC guidance improved detection performance (a marginally significant increase in accuracy but no change in response time [RT]) relative to the visual-only condition. We now consider the influence of these OW scanning differences on detection of the rogue aircraft, a characteristic of environments in which some aircraft are not equipped with a transponder and therefore are not registered on the CDTI.

Rogue Aircraft

The rogue aircraft was experienced only once by each pilot, and so the available data were insufficient to allow us to examine the interaction between workload and display format (there were only two observations/cell for this 3 × 2 design); therefore we consider each main effect in turn in a nonstatistical fashion.

In this analysis, the main effect is expressed as a cost (or benefit) relative to RT for the same aircraft when it was viewed by a different pilot when it appeared in a nonrogue state. That is, for the two pilots, the aircraft had exactly the same conspicuity properties in terms of aspect angle, motion, and eccentricity. For one pilot it was uncued (rogue), and for the other it had either visual or auditory cuing. Most important, the results of this analysis suggest that the increase in OW scanning associated with the high traffic load condition (see Figure 4) appears to have improved detection of the rogue aircraft, from a cost of 26.7 s (12.0 – 38.7 s) at low traffic load to a benefit of 2.5 s (48.0 – 45.5 s) at high traffic load. We cannot tell the extent to which this benefit for detecting rogue aircraft at high traffic load is the result of 7% more scanning outside or of the pilots' realization of the increased importance of traffic on trials in which traffic was more dense.

As we have also seen (Figure 4), the absence of a CDTI in the auditory condition availed far more OW scanning across both traffic load conditions. This OW allocation of visual attention, in turn, supported faster detection of the rogue aircraft in the auditory condition, in which there was an 11-s benefit, versus a 36-s cost for the V condition and only a 5-s benefit for the AV condition. Note that this pattern of effects is identical to that observed for the nonrogue aircraft at high workload, as reported earlier. These data also provide some evidence that the redundant modality did help call out of the announced traffic relative to the visual-only condition, but this trend was apparent only at low traffic load.

RESULTS: COMMUNICATIONS PHASE (DATA LINK)

Data recorded during the communications phase, which initiated each flight leg, were analyzed in a fashion corresponding to that carried out in the traffic detection phase, focusing initially on the effects of modality and load on flight control, then on their effects on the central task of interest (in this case, communications comprehension), and finally on their effects on visual scanning.

Flight Path Tracking

The communications task had different effects on lateral and vertical control. Heading errors were significantly disrupted, $F(1, 11) = 261$, $p < .01$, almost doubling, as a consequence of the longer communications strings. However, heading error was not influenced at all by modality ($F < 1$). In contrast, vertical tracking (altitude error) was unaffected by communications load but was substantially influenced by modality, $F(2, 22) = 4.63$, $p < .05$, showing a benefit for auditory delivery over the two visual conditions. This effect was in the direction opposite that observed in the prior study (Helleberg & Wickens, 2003), in which the auditory display, with its requirement to engage in clipboard writing, produced greater interference with tracking. In the current data, because the message strings were all relatively short (and known in advance to be so), pilots did not need to engage in writing and could capitalize on maintaining more visual attention on flight control, to the advantage of vertical control. The effect is consistent with that observed in the traffic detection analysis: An overall auditory advantage was observed in vertical error but not lateral error. The effects support a multiple-resource, rather than a preemption, interpretation.

Communications

The percentage of read-back errors is shown in Figure 5 and reveals a main effect of load, $F(1, 11) = 13.54$, $p < .05$, and of modality, $F(2, 22) = 8.18$, $p < .01$, both of which can be interpreted only in the context of the significant

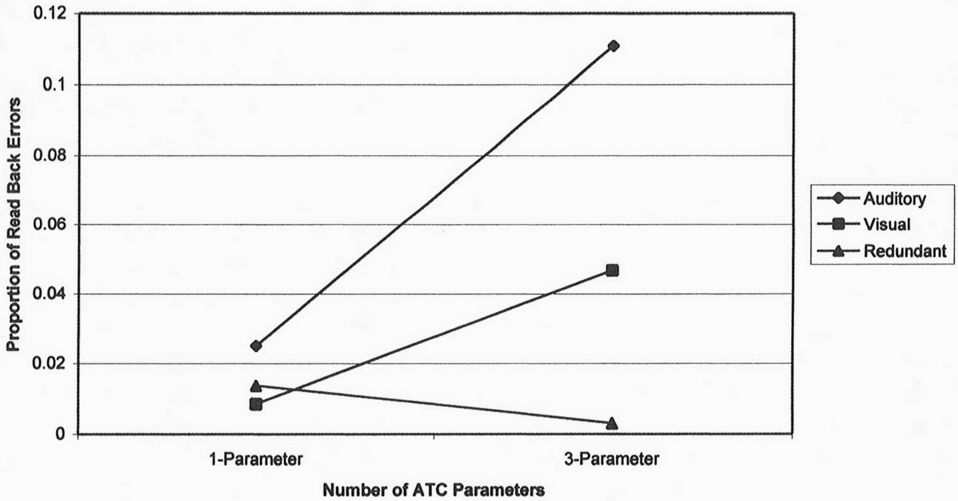

Figure 5. Proportion of read-back communication errors as a function of display and message length.

Load × Modality interaction, $F(2, 22) = 3.73$, $p < .05$. As is evident in the figure, when the communications load is short (one chunk), there is no penalty for auditory delivery (and indeed, as reported earlier, there is a reward in reduced vertical tracking error). However, loads of three chunks apparently sometimes exceeded the capacity of working memory, thereby causing read-back errors. These error rates were significantly higher in the auditory condition than in the two visual conditions (contrast with visual, $p < .02$; contrast with redundant, $p < .01$). These two visual conditions did not differ significantly from each other, although there is a hint of a redundancy benefit.

Visual Scanning

Although visual scanning was recorded during this segment of the flight legs, its interpretation is less critical because no OW traffic was presented during these phases. Helleberg and Wickens (2003) described scanning analysis during data link communications when traffic was present, and scanning analysis of the current data is reported in detail in Wickens, Goh, Helleberg, and Talleur (2002).

DISCUSSION: PART 1

The primary purpose of the current experiment was to ascertain the attentional (dual-task) effects of changing from the more conventional voice-delivered information on "secondary tasks" via ATC communications (traffic cuing and navigational instructions) to advanced-cockpit visual technology of the CDTI and data link. This shift from A to V, in an environment already heavily loaded with visual processing demands of tasks of equal (navigating) or higher (aviating) priority, was predicted to have differing effects, depending on the relative influence of multiple-resource mechanisms versus preemptive mechanisms of time-sharing performance. Both mechanisms predict a cost for the A-to-V shift for the secondary task, the modality of which was varied. Multiple resources predicts a corresponding cost to the ongoing higher-priority visual tasks (aviating, navigating). Preemption, however, predicts a benefit of the A-to-V shift to the higher-priority tasks because attention is not abruptly pulled away, or maintained away, from those visual tasks by a preemptive auditory onset while transient auditory information is processed to avoid decay from working memory. As we will discuss, the current evidence generally tends to support the role of multiple resources over that of preemption, although the influences of both mechanisms are also modulated by some task-specific compatibility influences. We catalog these latter influences by comparing only the two single-modality conditions.

Concerning performance of the higher-priority, visually supported primary tasks, there was general evidence for a visual cost (or auditory benefit), consistent with multiple resources. This cost was reflected in the aviate subtask of vertical tracking error during the traffic phase at high workload (four plane legs) and during the communications phase, and it was also reflected in the navigation subtask of traffic call out (a marginally significant effect of a 13% loss in detection rate). A particularly large (47-s) cost was evident in detection latency of the rogue aircraft, although the very low power of this difference prevented statistical comparison. These costs appeared to be mediated by the reduction in outside scanning imposed by the CDTI, which had a direct effect on the outside traffic detection tasks as well as an effect on flight path control, which we attribute to a period of "horizon deprivation." That is, in the auditory condition (in the good-weather VMCs simulated here), a view of the horizon for attitude control is always available, regardless of whether scanning is inside or out. In the visual condition, however, scans to the CDTI or data link display temporarily obscure this critical attitude information, thereby degrading flight path tracking. In summary, these collective "primary-task benefits" of auditory secondary-task display are consistent with a multiple-resource interpretation and inconsistent with a strong influence of a preemption mechanism.

Regarding performance of the secondary tasks, both mechanisms predict an auditory benefit, but in fact neither phase of flight showed such a benefit. Rather, the pattern of costs can be attributed more directly to task-specific effects (i.e., effects of modality that would be observed in single-task performance). In the communications phase, the auditory display imposed a cost to working memory for longer message strings, a cost well established in other literature (Helleberg & Wickens, 2003; Morrow, Lee, & Rodvold, 1993). In the traffic phase, the auditory display imposed a cost for precise navigation around conflicts, a cost directly attributable to the incompatibility of the auditory modality for delivering precise spatial information (Wickens et al., 1983). This cost was also seen in the inappropriate choice of conflict avoidance maneuvers fostered by the auditory condition. Furthermore, the multiple-resource benefits of the auditory modality for calling out "known traffic" (i.e., nonrogue aircraft) and availing more "eyes-out" time were probably somewhat offset by the lower precision with which this information could designate the location where the traffic could be seen in the outside world. This incompatibility also accounts for the results of Prinzo's (2003) study, revealing a visual CDTI advantage for traffic call-out latency.

We also evaluated the third (redundant, or AV) condition to see if it might provide the best of both (A and V) worlds. Somewhat surprisingly, the redundancy gain, which has been observed in some instructional literature (e.g., Sweller et al., 1990), was not seen here; instead, our data replicated patterns found in other dual-task simulations, by Helleberg and Wickens (2003), and Seagull and Wickens (2001). Thus flight path control was actually hurt by redundancy, relative to both single-modality displays in the lateral axis and to the auditory display in the vertical axis. There was no redundancy gain in outside scanning (relative to the visual-only condition), no gain in detection of either known or rogue traffic (relative to the auditory condition), and no redundancy gain in communications performance (relative to the visual condition). In short, the redundant display never outperformed the best of the two single-modality displays, and it occasionally led to the poorest performance. The primary explanation for this failure is that redundant presentation, by definition, offers twice as much "data" (even if there is no gain in "information"), relative to the single-modality conditions. As such, unless pilots adopt appropriate processing strategies for attending to such data, redundancy may act as an "attention sink" at the expense of processing other sources of information. It is yet to be determined if systematic training of attentional strategies can indeed create the "best of both worlds" with a redundant display.

In conclusion, the current data do suggest a note of caution regarding the visual attention demands of new technology and the possible negative implications of such technology for the high-priority tasks of aviating and navigating, particularly for single-pilot operations. Although the effects are not sufficiently strong to warrant reversion to radio-mediated ATC communication, with all of its additional pitfalls (Morrow et al., 1993), they do suggest the possibility of capitalizing on voice synthesis of digitally uplinked traffic and communications information. Even here, however, if such synthesis is to be used redundantly with existing text displays and CDTIs, some training in the use of redundant presentations may well be warranted, so that the "best of both worlds" can be realized.

Underlying our analysis of the data from Part 1 has been the application of models of attention. In the second part of this article we describe a model that focuses most directly on the visual scanning components as they relate to multitask management.

PART 2: COMPUTATIONAL MODEL OF ATTENTION ALLOCATION

Background

Our objective in Part 2 was to establish the extent to which the pattern of attention allocation, as indexed by visual scanning (Figure 4), could be captured by an optimal *expected value* model of information sampling (Wickens, Helleberg, Goh, Xu, & Horrey, 2001). Such a modeling effort is important for at least two reasons. First, to the extent that optimal scanning can be associated with good target detection, it may be established as a "gold standard" toward which training can be directed. Second, many human factors professionals have realized the importance of computational models of human performance in predicting performance in human-system interaction before a system is fully implemented, thereby obviating the need for expensive human-in-the-loop simulation (e.g., Pew & Mavor, 1998). The current model is designed to support this goal.

The model is based on the plausible assumption that four factors drive the acquisition of visual information: the salience (S) of events that might capture attention (Yantis, 1993); the effort (E) required to redirect attention from one location to another (i.e., visual saccade, head rotation), which will inhibit information access (Wickens, 1993); the expectancy (E) that a given location in the visual field will contain information (Senders, 1964); and the value (V) of information to be obtained at that location for the task or tasks at hand (Sheridan, 1970). As such, it is referred to as the SEEV model (Wickens et al., 2001). Previous researchers have modeled and empirically validated the role of expectancy (Senders, 1964), combined with value (Carbonnell, Ward, & Senders, 1968), in driving either visual scanning (Carbonnell et al., 1968; Ellis & Stark, 1986; Moray, Richards, & Low, 1980; Senders, 1964) or other measures of information sampling and acquisition (Kvalseth, 1977; Sheridan & Rouse, 1971). Furthermore, both Sheridan and Rouse (1971) and Kvalseth (1977) have accounted for the role of effort as an inhibitory process in information acquisition (see also Gray & Fu, 2001). In a comprehensive review of such models, Moray (1986) noted that the participants' level of expertise (e.g., Carbonnell et al., 1968; Moray et al., 1980) corresponded with the degree to which predictions of optimal expected value models of scanning were validated.

The current modeling effort builds on those described earlier in three respects. First, we define value not by the value of events that are detected along a visual channel but by the relative value or importance of a task served by the channel (e.g., in the current context, these values can be generated by the pilot's task priority hierarchy of "aviate-navigate-communicate"; Schutte & Trujillo, 1996). Second, by collecting data on a relatively large number of participants, we have faith that the measure used in establishing criterion validity of the model (attention allocation, or PDT, in an area of interest, as shown in Figure 4) is itself stable and more representative of the population than is the case with the smaller sample sizes used in evaluating other scanning models of expert participants (e.g., Carbonnell et al., 1968; Moray et al., 1980). Third, we cross-validate the model on two independent data sets, a process that was not carried out in the prior modeling efforts on attention allocation in vehicle control environments.

The optimal scanning form of the model is based only on the expectancy and value components of the larger SEEV scanning model. This is because, optimally, highly salient events should not be scanned, nor should effort be allowed to inhibit scanning over longer distances, as long as information with the maximum expected value is sought. Thus the model predicts that a visual channel or area of interest (AOI) will be sampled to the extent that it contains frequently changing information (high bandwidth) and that such information is relevant to a task of high importance or value.

The structure of the expected value model is shown in Figure 6, which represents three AOIs on the left: the instrument panel (IP), the outside world (OW), and the CDTI. Visual scanning as measured by percentage dwell time (PDT) may be allocated to any of these AOIs according to some proportion. Optimally, that proportion will be related to (a) the bandwidth (BW) along the channel (Senders, 1964) and (b) the degree of relevance (R) of each source of information to (in this example) each of the two most critical tasks, aviating and navigating. This BW × R product in turn must be weighted or multiplied by (c) the value (V) of the task supported by the AOI in question. Because some AOIs contain information relative to both tasks (e.g., the OW supports both the view of the horizon, for aviating, and the view of traffic to be avoided, for navigating), and because some tasks are served by more than one AOI (e.g., navigating is served by both the OW and CDTI view of traffic), the prediction of the amount of attention allocated to an AOI must sum the R × V product across all tasks served by an AOI, as shown by the formula at the bottom of Figure 6. Finally, because the model predicts the relative allocation of visual attention within a condition (PDT), all of the predicted values from the equation within a condition are summed, and each value within that condition is treated as a proportion of that sum.

The challenge to such a modeling effort lies not so much in its structure (Figure 6) but in assigning coefficients of bandwidth, relevance, and value. Here, we have done so by rank ordering low ordinal values across tasks and AOIs. This procedure has the advantage of allowing model users to easily reach a consensual agreement on what these values should be. For example, there is consensus in aviation that the task of aviating is more important (higher V) than navigating, which in turn is more important than communicating (Schutte & Trujillo, 1996). Hence it is easy to assign values of 3, 2, and 1 to these tasks, respectively. Correspondingly, a simple evaluation of visual dynamics reveals that the instrument panel (containing six instruments) shows more and faster-changing elements than does the view outside (containing only the horizon, also contained in the instrument panel), and this in turn shows more rapid changes than does the map information presented in the CDTI (which does not represent dynamic attitude information). Hence BW levels of 3, 2, and 1 can be assigned to these AOIs, respectively. Finally, the relevance of each AOI to each task must be assigned through a simple cognitive task analysis. The

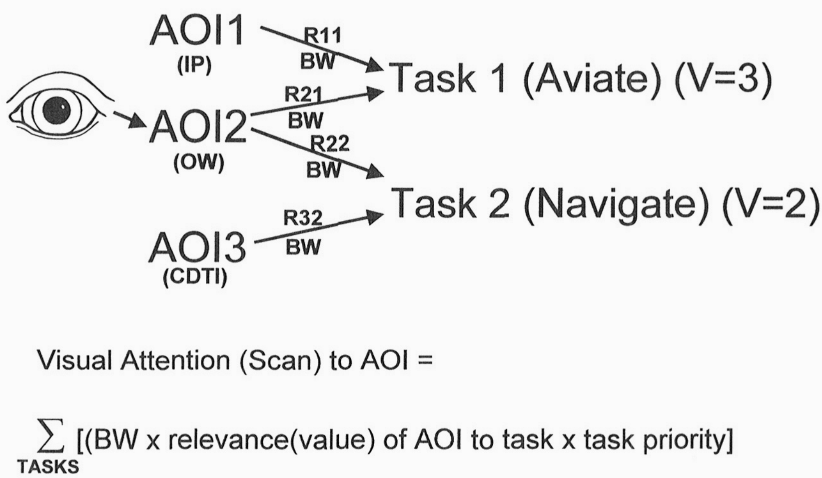

Figure 6. The expected value model of visual scanning or attention allocation. AOI = area of interest, IP = instrument panel, OW = outside world, CDTI = cockpit display of traffic information, BW = bandwidth, R*xy* = relevance of AOI to Task Y. The higher the value (V) of a task, the more important that task is.

horizon, for example, is more relevant for aviating than it is for navigating. More details of this assignment process are provided in Wickens et al. (2001).

In Table 1 we show the coefficient values for bandwidth, relevance, and task importance across the AOIs and tasks that were employed to predict the scanning data shown in Figure 4. In the table, input bandwidth, characterizing an AOI, is multiplied by the relevance of the relevance variable in the matrix below. This product is then multiplied by the correspondingly relevant task priority value in the far right column.

Model Fitting and Validation

Our modeling effort involved the fit of data for three experiments. In an initial model-fitting experiment (Wickens et al., 2001; Wickens, Helleberg, et al., 2002), pilots flew a mission similar to that described in Part 1 but under both a free-flight condition, in which they were supported by a CDTI to detect and avoid traffic, and a non-free-flight baseline condition, in which pilots had no CDTI but were guided around traffic conflicts by verbally issued ATC vectors. These two different conditions were crossed factorially with flight legs in which there either was or was not conflict traffic that needed to be avoided. The four different conditions, coupled with three AOIs (in free flight) or two AOIs (in baseline; there was no CDTI), generated 10 different predicted data points for the proportion of time that visual attention would remain within the AOI in question (Table 2). The mean PDT data across pilots served as the predictive criterion of the three-parameter model.

TABLE 1: Parameter Values for Experiment Described in Part 1: Traffic Density and Modality

Parameter	AOI: IP	OW	CDTI	
Bandwidth (B)				
Visual (1)	2	1	0.5	
Visual (4)	2	2	2	
Auditory (1)	2	1	—	
Auditory (4)	2	2	—	
Relevance (R)				Priority (V)
Aviate (V)	3	1	0	3
Navigate (V)	1	2	2	2
Aviate (A)	3	1	—	3
Navigate (A)	2	4	—	2

Note. The values of 1 and 4 in the bandwidth listing correspond to the traffic density.

TABLE 2: Parameter Values Assumed for Experiment 1a (Free Flight) and Experiment 1b (Baseline)

Parameter	AOI: IP	OW	CDTI	
Bandwidth (B)				
Free flight (conflict)	3	2	1	
Free flight (nonconflict)	2	1	0.5	
Baseline (conflict)	3	2	—	
Baseline (nonconflict)	2	1	—	
Relevance (R)				Priority (V)
Aviate (free flight)	2	1	0	3
Navigate (free flight)	1	2	2	2
Aviate (baseline)	2	1	—	3
Navigate (baseline)	1	2	—	1

Figure 7 plots the model-predicted PDT against the obtained percentage dwell values, averaged over pilots for the 10 AOI-condition combinations for the first experiment, yielding a linear correlation of $r = +.885$, or accounting for 78% of the variance. The scatter plot shows three clusters of data points. Those in the lower left represent the relatively scarce scan to the CDTI, those in the center represent the more frequent scan to the OW, and those in the upper right represent the most frequent scan, to the IP. It may be asked whether a simple one-parameter model, in which scanning is based purely on AOI bandwidth and all other values

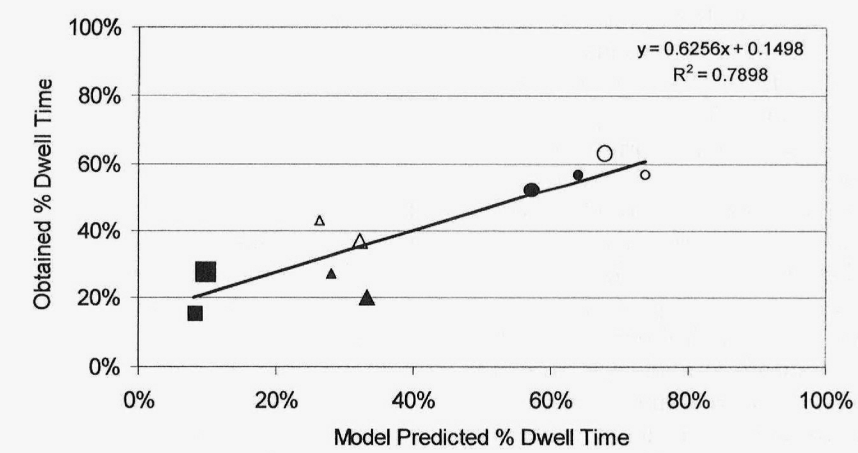

Figure 7. Model prediction versus scan (percentage dwell time) performance for the free-flight experiment. Squares = CDTI, triangles = OW, circles = IP; small = conflict, large = nonconflict; solid symbols = free flight, open symbols = baseline.

ing bandwidth and priority actually increased the model fit slightly (above that of the three-parameter version), accounting for 82% of the variance. Thus, in our initial validation, we found that bandwidth clearly dictated attention allocation (Senders, 1964), with little independent evidence for effects of relevance and task priority.

We then cross-validated the same model on the independently collected data set (Figure 4) from the "traffic" experiment described in Part 1. Only visual and auditory conditions were used because the experimental variables were slightly different in this experiment: We needed to change some of the coefficients and yet preserve the ordinal modeling assumptions of the first fit. For example, we considered the OW and CDTI bandwidth in low-traffic legs (one airplane/leg) to be less than that in high-traffic legs (four/leg), as shown in Table 1. The BW values in equivalent conditions were set to the same value that they had been in the first experiment (compare Tables 1 and 2). We also assumed that the total relevance of all sources of visual information for traffic was constant across conditions. Hence, when the CDTI was eliminated in the auditory condition, the relevance of the outside world for navigating was increased from 2 to 4. Further assumptions on setting these coefficients are described in detail in Wickens et al. (2001).

The fit between model predictions and obtained mean PDT data in the present experiment are shown in Figure 8, in which each data point corresponds to a condition (bar) in Figure 4. The six bars from the redundant condition were not included, nor were the two bars for the CDTI in the auditory condition (for which there were no fixations). Thus 10 data points were modeled. In this case, the fitting to the data was even more accurate than that from the initial experiment shown in Figure 7, now showing an R^2 of 90% ($r = .93$).

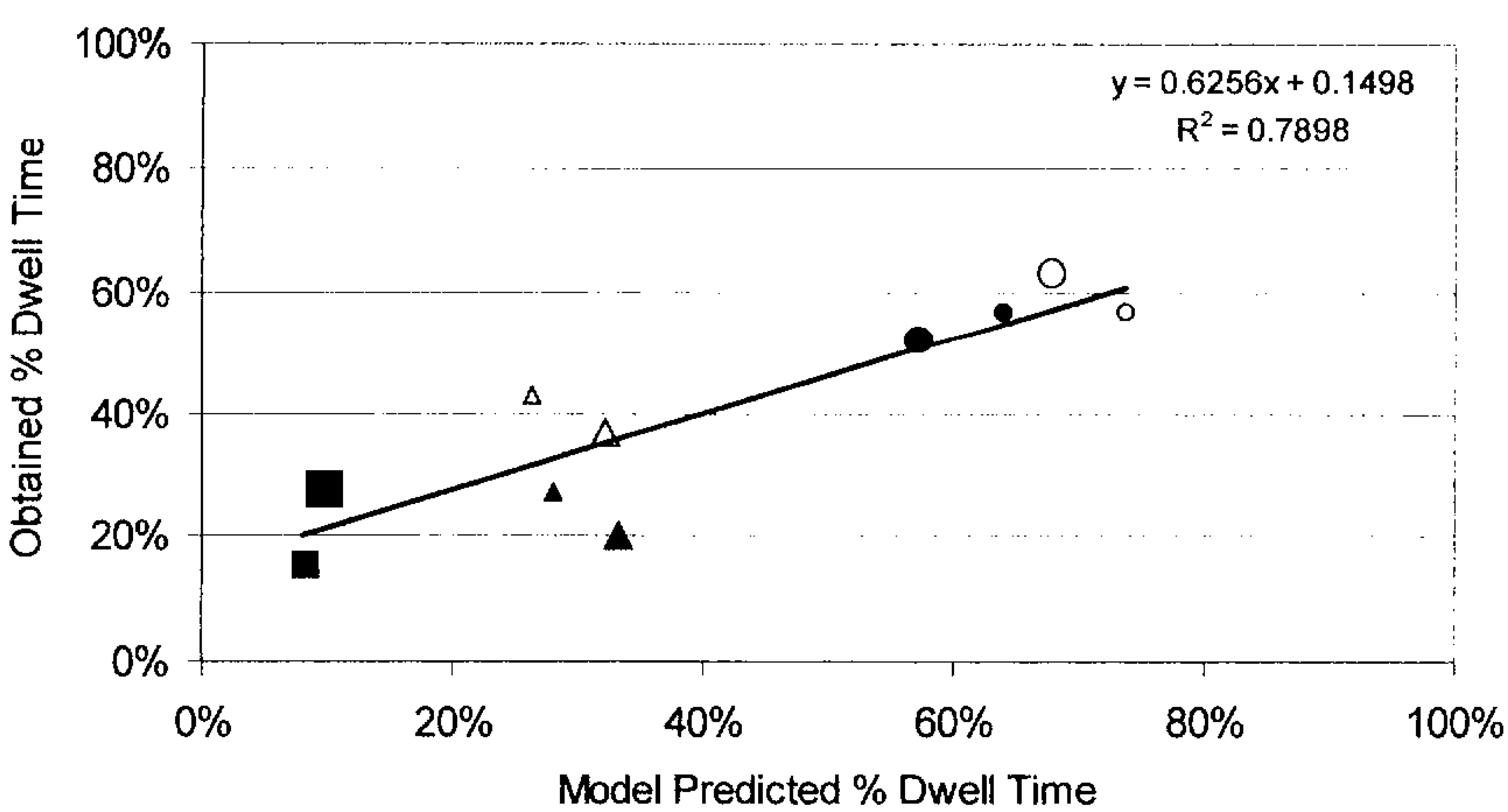

Figure 7. Model prediction versus scan (percentage dwell time) performance for the free-flight experiment. Squares = CDTI, triangles = OW, circles = IP; small = conflict, large = nonconflict; solid symbols = free flight, open symbols = baseline.

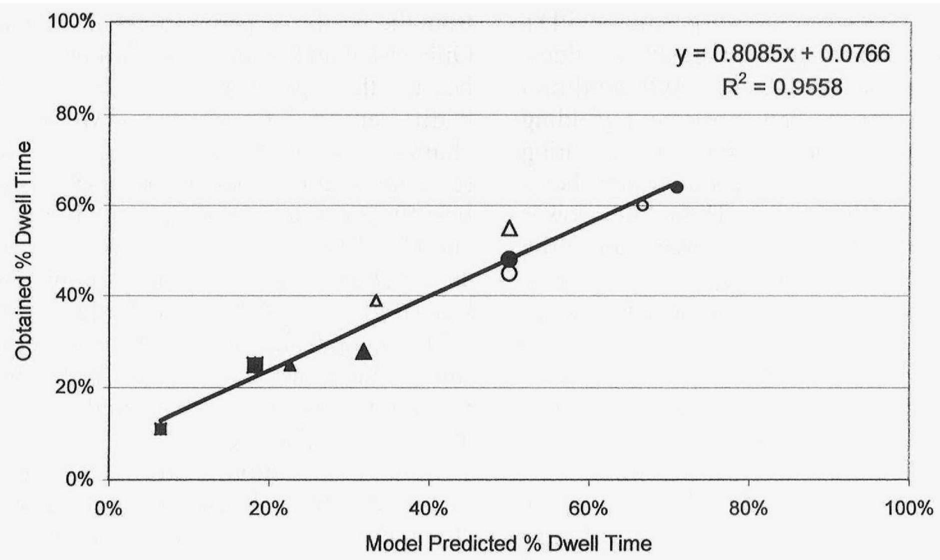

Figure 8. Model fit of traffic experiment. Squares = CDTI, triangles = OW, circles = IP; small = 1 traffic, large = 4 traffic; solid symbols = visual CDTI, open symbols = auditory.

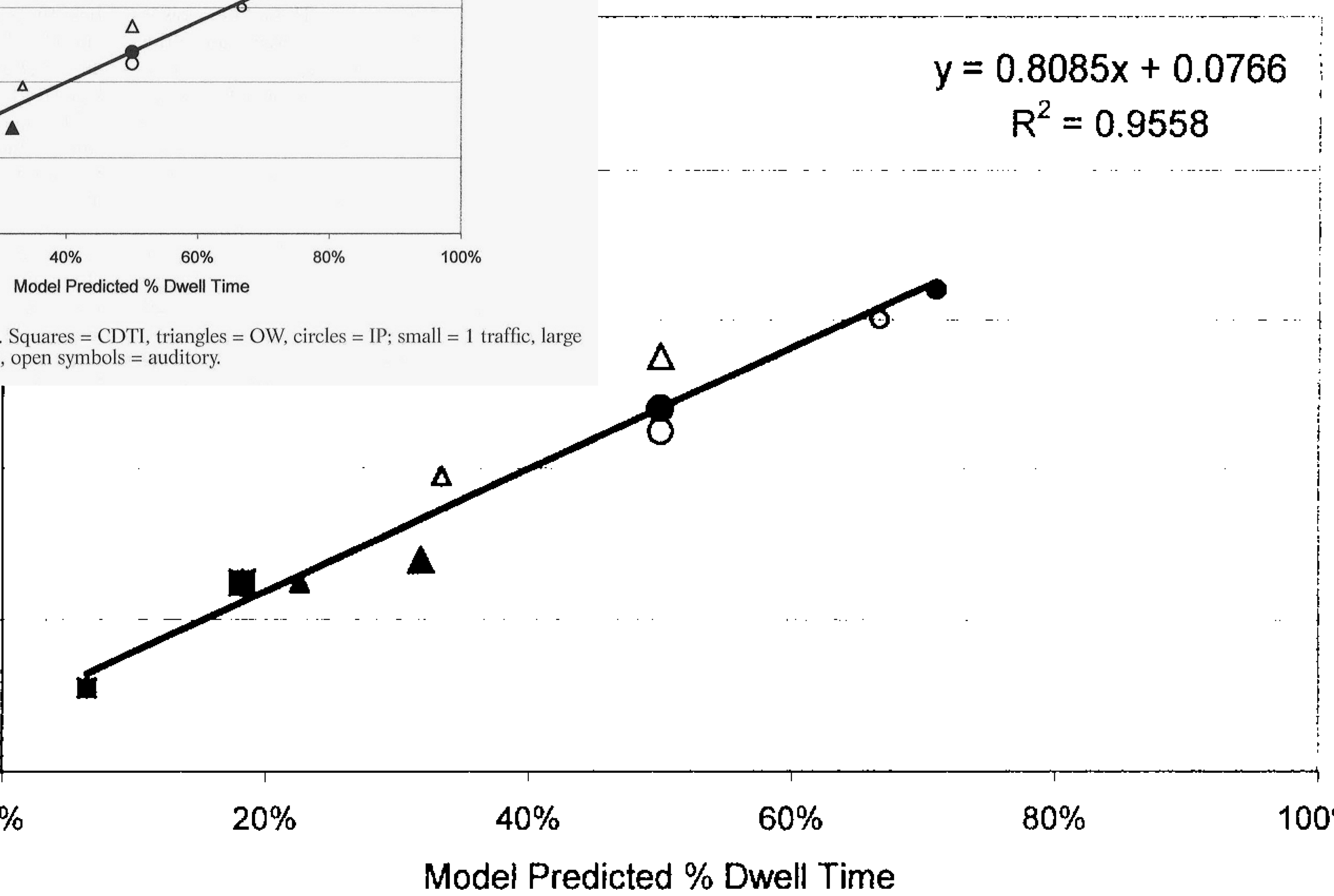

Figure 8. Model fit of traffic experiment. Squares = CDTI, triangles = OW, circles = IP; small = 1 traffic, large = 4 traffic; solid symbols = visual CDTI, open symbols = auditory.

As before, the model was run again with a set of simplified assumptions. When the discriminating parameters of either bandwidth or relevance were removed from the predictive equation, the variance accounted for dropped to values of 78% and 81%, respectively. When only bandwidth remained, the model fit was 81%. Single-parameter models of only relevance and only priority accounted for still less variance (71% and 23%, respectively).

The model's ability to predict scanning behavior of individual pilots was examined by computing correlations for the 8 pilots for whom full scanning data were available. These correlations ranged from .42 to .95 with a mean of .85. Evidence that the model fit was related to degree of experience was provided by the positive correlation (r = .55) between the correlation fit values and the pilot's hours of flight experience.

Finally, the model was cross-validated a second time, this time using the visual scanning data collected by Helleberg and Wickens (2003), who contrasted a data link display with auditory communications and a clipboard, using five different communications loads defined by the number of "chunks" of relevant information in the message. In that experiment, traffic was monitored in the outside world without the aid of a CDTI during the period when communication information could be delivered, and the communications AOI (clipboard or data link display) replaced the CDTI as the third AOI. Communications was now added as a third task with a priority of 1, which was lower than that of navigation (2) or aviating (3), and the bandwidth of the communications task was assigned five ordinal values as the message length (number of "chunks" of information) increased. The model coefficient values are shown in Table 3.

Figure 9 presents the correlation between model predictions and obtained PDT data from this third "communications experiment." As with the second experiment, the model fit is again strong, accounting for 95% of the variance in PDT data. As in Figure 7, Figure 9 reveals three discrete clusters, which are associated with (from left to right) the communications display, the OW, and the IP. Although it is true that much of the variance is accounted for by the three clusters, as in the first experiment, a quick view of the data points in Figure 9 also suggests the role of communications load (bandwidth) in accounting for shared variance between predicted and obtained scores within each cluster. As with the other two data sets, we removed single parameters from the model, and we observed that removal of task relevance (bandwidth and priority remaining) dropped the model fit to 72% of the variance. In this case, removal of bandwidth (relevance and value remaining) had no

TABLE 3: Parameter Values for Experiment 4: Communications Experiment

Parameter	AOI				
	IP	OW	Com DL	Com Clip	
Bandwidth (B), visual/red					
Com Load 2	4	2	0.5	0	
Com Load 3	4	2	1	0	
Com Load 4	4	2	1.5	0	
Com Load 5	4	2	2	0	
Com Load 6	4	2	2.5	0	
Bandwidth (B), auditory					
Com Load 2	4	2	0	0.5	
Com Load 3	4	2	0	1	
Com Load 4	4	2	0	1.5	
Com Load 5	4	2	0	2	
Com Load 6	4	2	0	2.5	
Relevance (R)					Priority (V)
Aviate	3	1	1	1	3
Navigate	1	2	1	1	2
Communicate	0	0	2	2	1

Note. Com DL = data link communications, Com clip = clipboard communications.

effect on the model fit. The maximum fit of the three single-parameter models was 72%.

DISCUSSION: PART 2

The simple expected value model appears to have done a good job of accounting for at least 90% of the variance in obtained scanning measures in the two cross-validation experiments. Although it may be argued that these fits are less impressive, given three parameters to fit 10 (traffic experiment) or 15 (communications experiment) data points, four important counterarguments can be offered to suggest that the "performance" of the model is not trivial. First, the two cross-validation predictions were made a priori, without adjusting the coefficients derived from Experiment 1 to maximize the fit. As the

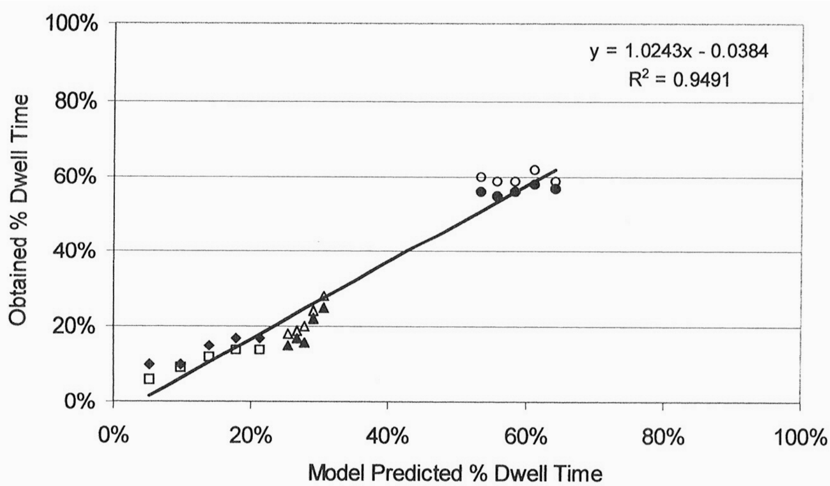

Figure 9. Model fit of communications experiment. Solid symbols = auditory, open symbols = visual; circles = IP, triangles = OW, squares = data link, diamonds = clipboard. The five communicative load points are not separately labeled but fall in monotonic order.

preceding discussion suggests, every effort was made to assign the ordinal coefficient values on a basis that could be justified independently of model fit. Second, both the parameters themselves, as well as the architecture of the model, would appear to be "cognitively plausible" based on a rationale task analysis of how people should behave. This rationality is inherent in the foundations of the model of expected value decision making (e.g., Edwards, 1987; Lehto, 1997).

Third, we note that each parameter of the model appears to contribute independently to accounting for variance in observed data. The role of bandwidth is obvious, as defined by the three clusters in each graph. However, in each cross-validation fit, removal of each of the other two parameters either substantially reduced (three cases) or left unchanged (in one case) the goodness of fit of the model. Fourth, we note that a second form of validation was established, in the traffic experiment, by our finding that better model fits were obtained by pilots with greater experience, who presumably had better internalized the important parameters driving the model.

Several possible extensions should be pursued in the modeling effort, and there are limitations of the current effort. For example, sensitivity analysis should be explored across a range of parameter values. It also would be of importance to determine if independent analysts would arrive at the same ordinal coefficient assignments that we did. It should also be noted that our operational definition of *navigating* in the current context was a fairly restricted one, involving traffic avoidance rather than, for example, route planning or flying from navigational instruments.

GENERAL DISCUSSION

The current experiment has provided an understanding of the implications of new technology for dual-task performance and for the manner in which such technology drives the allocation of visual attention between flight-critical tasks. Also, for that allocation we have provided a plausible computational model that appears to be well validated by the existing scanning data.

The current research has both theoretical and practical implications. In terms of theoretical implications, the data from Part 1 point to the dominating role of a multiple-resource explanation for auditory-visual differences in a multitask scenario, relative to a preemption explanation. Other factors that would appear to enhance the prominence of preemption (and diminish that of visual resource competition) were not in evidence here. These include circumstances in which separation between visual channels is reduced in the VV condition (e.g., a head-up display; Wickens, Dixon, et al., 2002) or in which the auditory messages are more abrupt, less predictable, and longer (as in Helleberg & Wickens, 2003). Part 2 also provides data that weigh in on the general argument as to whether human operators are "optimal" or "suboptimal" in supervisory control and resource allocation (see Moray, 1986). The current data suggest, comfortingly, that the skilled pilots can be well categorized as being optimal. In a more general sense, the current data point to the value of theoretical models in accounting for behavior in this complex real-world task.

There are also several practical implications of the data. Part 1 suggests that visual in-cockpit technology should be adopted with caution for single-pilot operations and that other design or training features should be considered for adoption in order to mitigate the attentional implications of that technology. This is particularly true for CDTIs, which may not be "knowledgeable" of all outside traffic (e.g., rogue aircraft). Indeed, design and training should, perhaps, be performed in conjunction if redundant display modalities are chosen, so that the "best of both worlds" of redundancy can be realized.

Part 2 provides the possibility of a model that can act as a "gold standard" against which scanning strategies of novices may be compared and that can be used to diagnose nonoptimal patterns. The model also offers the potential to assess when design features may, by increasing the effort of information access, lead to serious departures from optimal scan patterns.

In conclusion, pilot attention is likely to remain one of the critical "limited resources" that will be challenged by new technology. As a consequence, it is important that the nature of such attention be modeled so as to allow prediction of the implications of new technology.

ACKNOWLEDGMENTS

The experiment reported here was supported by Grant NASA NAG 2-1120 from NASA Ames Research Center. Dr. David Foyle was the scientific/technical monitor. Also, thanks go to Ron Carbonari and Roger Marsh for their help in programming the experimental scenarios.

REFERENCES

Begault, D. R., & Pittman, M. T. (1996). Three-dimensional audio versus head-down traffic alert and collision avoidance system displays. *International Journal of Aviation Psychology, 6,* 79–93.

Carbonnell, J. F., Ward, J. L., & Senders, J. W. (1968). A queuing model of visual sampling: Experimental validation. *IEEE Transactions on Man Machine Systems, MMS-9,* 82–87.

Damos, D. L. (1997). Using interruptions to identify task prioritization in Part 121 air carrier operations. In *Proceedings of the 9th International Symposium on Aviation Psychology* (pp. 871–876). Columbus, OH: Ohio State University.

Dismukes, K. (2001). The challenge of managing interruptions, distractions, and deferred tasks. In *Proceedings of the 11th International Symposium on Aviation Psychology* [CD-ROM]. Columbus, OH: Ohio State University.

Edwards, W. (1987). Decision making. In G. Salvendy (Ed.), *Handbook of human factors* (pp. 1061–1104). New York: Wiley.

Ellis, S. R., & Stark, L. (1986). Statistical dependency in visual scanning. *Human Factors, 28,* 421–438.

Gray, W. D., & Fu, W. T. (2001). Ignoring perfect knowledge in-the-world for imperfect knowledge in-the-head: Implications of rational analysis for interface design. *CHI Letters, 3,* 112–119.

Helleberg, J., & Wickens, C. D. (2003). Effects of data link modality and display redundancy on pilot performance: An attentional perspective. *International Journal of Aviation Psychology, 13,* 189–210.

Kerns, K. (1999). Human factors in air traffic control/flight deck integration: Implications of data-link simulation research. In D. J. Garland, J. A. Wise, & V. D. Hopkin (Eds.), *Handbook of aviation human factors* (pp. 519–546). Mahwah, NJ: Erlbaum.

Kvalseth, T. (1977). Human information processing in visual sampling. *Ergonomics, 21,* 439–454.

Latorella, K. (1998). Effects of modality on interrupted flight deck performance. In *Proceedings of the Human Factors and Ergonomics Society 42nd Annual Meeting* (pp. 87–91). Santa Monica, CA: Human Factors and Ergonomics Society.

Lehto, M. (1997). Decision making. In G. Salvendy (Ed.), *Handbook of human factors and ergonomics* (2nd ed., pp. 1201–1248). New York: Wiley.

Moray, N. (1986). Monitoring behavior and supervisory control. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and performance* (Vol. 2, pp. 40-1–40-51). New York: Wiley.

Moray, N., Richards, M., & Low, J. (1980). *The behaviour of fighter controllers* (Tech. Report). London: Ministry of Defense.

Morrow, D., Lee, A., & Rodvold, M. (1993). Analysis of problems in routine controller-pilot communication. *International Journal of Aviation Psychology, 3,* 285–302.

Navarro, C., & Sikorski, S. (1999). Datalink communication in flight deck operations: A synthesis of recent studies. *International Journal of Aviation Psychology, 9,* 361–376.

Navon, D., & Gopher, D. (1979). On the economy of the human processing systems. *Psychological Review, 86,* 254–255.

Parkes, A. M., & Coleman, N. (1990). Route guidance systems: A comparison of methods of presenting directional information to the driver. In E. J. Lovesey (Ed.), *Contemporary ergonomics 1990* (pp. 480–485). London: Taylor & Francis.

Pew, R. W., & Mavor, A. (1998). *Modeling human and organizational behavior.* Washington, DC: National Academy of Sciences.

Prinzo, O. V. (2003). *Pilot visual acquisition of traffic:* An in-flight evaluation of a cockpit display of traffic information. *International Journal of Aviation Psychology, 13,* 211–231.

Schutte, P. C., & Trujillo, A. C. (1996). Flight crew task management in non-normal situations. In *Proceedings of the Human Factors and Ergonomics Society 40th Annual Meeting* (pp. 244–248). Santa Monica, CA: Human Factors and Ergonomics Society.

Seagull, J., & Wickens, C. D. (2001). *Auditory displays for alerting medical practitioners: A simulated patient care scenario* (Tech. Report ARL-01-1/FED-LAB-01-1). Savoy, IL: University of Illinois, Aviation Research Lab.

Senders, J. (1964). The human operator as a monitor and controller of multidegree of freedom systems. *IEEE Transactions on Human Factors in Electronics, HFE-5,* 2–6.

Sheridan, T. (1970). On how often the supervisor should sample. *IEEE Transactions on Systems Science and Cybernetics, SSC-6,* 140–145.

Sheridan, T. B., & Rouse, W. B. (1971, June). Supervisory sampling and control: Sources of suboptimality in a prediction task. In *Seventh NASA Annual Conference on Manual Control* (NASA Special Publication SP-215, n.p.). Washington, DC: U.S. Government Printing Office.

Spence, C., & Driver, J. (2000). Audiovisual links in attention: Implications for interface design. In D. Harris (Ed.), *Engineering psychology and cognitive ergonomics* (n.p.). Aldershot, UK: Ashgate.

Sweller, J., Chandler, P., Tierney, P., & Cooper, M. (1990). Cognitive load as a factor in the structuring of technical material. *Journal of Experimental Psychology: General, 119,* 176–192.

Wickens, C. D. (1991). Processing resources and attention. In D. Damos (Ed.), *Multiple-task performance* (pp. 3–34). London: Taylor & Francis.

Wickens, C. D. (1993). Cognitive factors in display design. *Journal of the Washington Academy of Sciences, 83,* 179–201.

Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomic Science, 3,* 159–177.

Wickens, C. D. (2003). Aviation displays. In P. Tsang & M. Vidulich (Eds.), *Principles and practices of aviation psychology* (pp. 147–199). Mahwah, NJ: Erlbaum.

Wickens, C. D., Dixon, S., & Seppelt, B. (2002). *In-vehicle displays and control task interference: The effects of display location and modality* (ARL-02-7/NASA-02-5/GM-02-1). Savoy, IL: University of Illinois, Aviation Research Lab.

Wickens, C. D., Goh, J., Helleberg, J., & Talleur, D. A. (2002). *Modality differences in advanced cockpit displays: Comparing auditory and vision for navigational communications and traffic awareness* (ARL-02-8/NASA-02-6). Savoy, IL: University of Illinois, Aviation Research Lab.

Wickens, C. D., Helleberg, J., Goh, J., Xu, X., & Horrey, W. J. (2001). *Pilot task management: Testing an attentional expected value model of visual scanning* (ARL-01-14/NASA-01-7). Savoy, IL: University of Illinois, Aviation Research Lab.

Wickens, C. D., Helleberg, J., & Xu, X. (2002). Pilot maneuver choice and workload in free flight. *Human Factors, 44,* 171–188.

Wickens, C. D., & Hollands, J. (2000). *Engineering psychology and human performance* (3rd ed.). Upper Saddle River, NJ: Prentice Hall.

Wickens, C. D., & Liu, Y. (1988). Codes and modalities in multiple resources: A success and a qualification. *Human Factors, 30,* 599–616.

Wickens, C. D., Sandry, D., & Vidulich, M. (1983). Compatibility and resource competition between modalities of input, output, and central processing. *Human Factors, 25,* 227–248.

Yantis, S. (1993). Stimulus-driven attentional capture. *Current Directions in Psychological Science, 2,* 156–161.

Christopher D. Wickens is a professor of psychology at the Institute of Aviation, Human Factors Division, University of Illinois. He received his Ph.D. in experimental psychology at the University of Michigan in 1974.

Juliana Goh is a research assistant at the Institute of Aviation, Human Factors Division, University of Illinois, where she received her master's degree in psychology in 2002.

John Helleberg is a research scientist at MITRE Corporation. He received his master's degree in psychology at the University of Illinois in 2001.

William J. Horrey is a research assistant a the Beckman Institute of Science and Technology, University of Illinois, where he received his master's degree in psychology in 2002.

Donald A. Talleur is an associate aviation education specialist and assistant chief flight instructor in the Professional Pilot Division and Human Factors Division of the Institute of Aviation at the University of Illinois, where he received a master's degree in psychology in 2001.

Date received: January 15, 2002
Date accepted: March 24, 2003