

Effects of Data-Link Modality and Display Redundancy on Pilot Performance: An Attentional Perspective

John R. Helleberg and Christopher D. Wickens

*Aviation Human Factors Division
University of Illinois, Savoy, Illinois*

The effectiveness of three different data link interfaces, involving auditory, visual, and redundant presentation of Air Traffic Control (ATC) information, was evaluated in a single pilot general aviation simulator. Fifteen certified flight instructor pilots flew a Frasca flight simulator with full visual display of the outside world, through a series of ATC-instructed maneuvers, while scanning outside for traffic. ATC instructions, of various lengths, were delivered through a text based data link display, through synthesized voice (similar to the current system), or through a redundant voice-text format. Pilots read back the instructions and then complied with whatever maneuver was instructed, while monitoring for traffic. Visual scanning was measured. The results revealed that the visual display provided greatest accuracy of communications readback, and was least disruptive of both traffic monitoring and flight path tracking. The auditory-only condition was most disruptive of these tasks, in part because the pilot's visual attention was drawn into the cockpit for longer periods while note taking, when compared with the visual text display. The redundant display condition provided many of the same benefits as the visual-only display, but was not associated with better performance than that display, and was sometimes inferior to it. This resulted, in part, because the arrival of the discrete auditory communications appeared to disrupt performance of the ongoing visual flight task. Across all conditions, the pilots allocated approximately 60% of their visual attention to monitoring the instrument panel and their communications accuracy was degraded by the longer ATC instructions. The results are interpreted in terms of mechanisms of attention and working memory in information processing, and in terms of the pilot's task priority hierarchy.

On March 27, 1977 a Pan Am 747-121 (flight number 1736) and a KLM 747-206B (flight number 4805) collided while still on the runway at Los Rodeos airport on Tenerife of the Canary Islands. This disaster resulted in the largest loss of life (583 people) that has ever occurred from a single aviation accident. The probable cause, cited by Roitsch, Babcock, and Edmunds (1978), was the KLM pilot taking off without takeoff clearance. How could a senior captain make such an error? There were many contributing factors that led up to this accident. However, miscommunication of the auditory-voice loop played a significant role and is the major focus of this article.

A prototype data-link system is currently being tested in actual operations at Miami Center (Fiorino, 2002). However, a majority of the current airspace system still uses a voice-only radiotelephone system to exchange information between air traffic controllers and pilots. There are several problems with this current voice-only system. For example, Dismukes (2001) examined Aviation Safety Reporting System (ASRS) database incidents, which occurred when the aircrew's attention was diverted away from an ongoing task to attend to some type of interruption. Following the interruption, the crew typically failed to resume the ongoing task, thus leading to the reported incident. The analysis indicated that 68 of the 107 incidents were related to the crew's attention being distracted from the task at hand to engage in communication.

Communication errors listed in the ASRS database are found to result from several sources (Kerns, 1999), comprehension of voice communications, working memory demands (Baddeley, 1990; Loftus, Dark, & Williams, 1979), the vulnerabilities of the readback process (Monan, 1986), and message confusion between multiple pilots (Flathers, 1987). The current system can also lead to congestion of a single radio frequency, which can lead to substantial delays in communication (Lee, 1989; National Research Council, 1998), leading to communications that are either missed or blocked by the transmissions of other users (Kerns, 1991, 1999).

The proposed system that is intended to reduce the aforementioned problems is the implementation of a data-link system. Data link is a digital means to transmit information from the ground-based controller to the airborne pilot (Kerns, 1991, 1999; Navarro & Sikorski, 1999). The basic system will enable controllers to transmit text-based messages to aircraft equipped with this technology and, as conventionally envisioned, present this information on a visual (text) display somewhere within the cockpit, in a manner that is more "permanent" than is the transient voice message currently received from air traffic control (ATC). Pilots will also be able to acknowledge receipt of that information with a key press and transmit requests to air traffic controllers.

One of the advantages of using a data-link system is that it reduces the demands on the pilot's working memory by maintaining a lasting depiction of the air traffic controller's instruction. Thus data link may also provide pilots more flexibility to complete highly demanding flight tasks before responding to or complying with

ATC instructions, whereas with the conventional radiotelephone-only communication system an immediate response is typically required (McGann, Morrow, Rodvold, & Mackintosh, 1998).

Against these advantages of datalink may be arrayed some concerns that are directly related to pilot attention, in particular, the competition for visual attention imposed by the visual display and the resulting “head down time” imposed on a pilot who should also be looking outward or monitoring instruments (Wickens, Mavor, Parasuraman, & McGee, 1998). This issue is of sufficient importance that some (Van Gent, 1995) have suggested that the voice display characteristics of the current ATC system might be preserved in data link, using computer-synthesized voice, rather than (or redundant with) a visual text display. A second attentional issue concerns the potential disruption or interruption (i.e., attentional switching) of ongoing activity, imposed by the arrival of an ATC message, which may be greater with an auditory (conventional or synthesized voice) than a visual (data link text) message (Latorella, 1998).

RESOURCE COMPETITION VERSUS PREEMPTION

Wickens (1980) conducted an extensive literature review comparing dual-task performance studies, which examined information presentation utilizing either cross-modal (auditory–visual) or intramodal (visual–visual) displays. The general finding was that cross-modal information presentation yielded more efficient performance than did intramodal information presentation. Wickens suggested that this result may be due to different multiple processing “resources” associated with visual and auditory processing (Wickens, 1980, 1991, 2002). This resource competition perspective predicts that the proposed visual text based data-link system may lead to overloading of the pilot’s visual resources, thereby imposing a performance cost to both the ongoing primary flight task (which are primarily visual) as well as the discrete secondary communications task. This resource competition perspective suggests that, within the visually demanding cockpit environment, the best overall performance will be realized with an auditory data link presentation utilizing a computer-synthesized voice display.

In addition to resource competition, a second psychological attention mechanism affected by the format of ATC information display is the degree to which the conventional auditory versus visual (data link) modality may *preempt* ongoing (primarily visual) activity in the cockpit. Research has shown that the discrete presentation of auditory information tends to result in an alerting characteristic and is more intrusive in nature than the discrete presentation of visual information (Posner, Nissen, & Klein, 1976). Within the context of dual-task performance, where an ongoing visual task is time-shared with a discrete auditory task, this interrupting characteristic of auditory information could potentially lead to a

“preemption” effect, whereby performance of the ongoing visual task suffers as a result of the discrete auditory task temporarily diverting resources away from the ongoing visual task and reallocating those resources to the more preemptive discrete auditory task (Wickens & Liu, 1988). Such diversion is less likely with a visual discrete task (Latorella, 1998; Wickens & Liu, 1988).

This preemption factor might offset some of the proposed benefits for an auditory display based upon multiple resources, in so far as the ongoing continuous visual task is concerned, and hence favor the proposed visual data link display, particularly given that data link is employed for communications, a lower priority task in the pilot’s standard “aviate, navigate, communicate” task hierarchy (Schutte & Trujillo, 1996; Wickens, Helleberg, & Xu, 2002). Furthermore, given the analysis of Wickens and Liu (1988) because both the aviate and navigate tasks (higher priority) are visual these are most likely to be preempted by the auditory format of a data link message.

Evidence for the role of preemption versus resource competition can be derived from a handful of studies that have contrasted auditory with visual display of communications information while pilots are engaged in flying tasks; unfortunately few of these have recorded all of the data necessary to draw strong conclusions. Van Gent (1995) evaluated a visual data link interface by using pilots in a commercial aircraft flight simulation. He observed the large extent to which the visual gaze of both pilots was shifted downward to process the visual data link interface. The investigation did not, however, contrast this directly with the baseline auditory system, nor measure performance on the primary flight task.

Mackintosh, McGann, Lozito, and Logsdon (1999) did compare the two interfaces in a flight simulation study; their results provided some evidence that the auditory display produced less interference (e.g., a result supporting the role of resource competition). McGann et al. (1998) also compared the two interfaces for data-link information. Consistent with a preemption mechanism, they noted that pilots begin to deal with the auditory data link message more immediately (i.e., begin to make entries into the mode control panel sooner), than with the visual message (delayed entry). However, the investigators did not report primary task performance data, which would enable confirmation of a preemption mechanism (e.g., greater disruption by the auditory interface).

Two studies appear to have provided all of the necessary tasks and measurements to draw some conclusions regarding the two mechanisms of preemption and resource competition in a flight environment. Wickens, Sandry, and Vidulich (1983) had pilots perform a communications task, wherein information was delivered auditorally or visually, while they flew an F/A-18 simulator. Consistent with both preemption and competition, they found that the performance of the discrete communications task was much more timely (shorter initiation and completion) when it was delivered auditorally. However, there was little benefit (but no cost) of auditory display for the ongoing flight task. This could indicate that whatever benefits of an auditory display to flight control were offered by reduced resource competition, these benefits were offset by preemption.

Finally, a study carried out by Latorella (1996, 1998) offered strong support for the role of preemption. In a flight simulation, pilots responded to data link communications messages, either delivered auditorally or visually, while engaged in an ongoing visual task (extracting information from the flight management system). Latorella observed both the better performance of the data-link task with the auditory interface (faster initiation) and the greater disruption of the ongoing visual task when the data-link message was presented auditorally. Such a result is entirely consistent with Dismukes' findings from ASRS reports: auditory communication tasks tend to interrupt tasks of higher intrinsic priority (Dismukes, 2001).

REDUNDANT DISPLAYS

Although all of the aforementioned studies have compared the effects of presenting interrupting task information with either a visual or an auditory interface, optimal performance might be achieved if a redundant display (simultaneously presenting auditory and visual information) were utilized. Some research in nonflight environments has found redundancy gain (Wickens & Hollands, 2000). However, there appears to have been only a handful of studies that have examined redundant presentation using both auditory and visual modalities within a multiple task context, typical of the pilot's cockpit environment. It is conceivable that this type of redundant display may provide the "best of both worlds" where the visual information allows pilots to avert a preemption effect and the redundant auditory information enables pilots to avoid the effects of resource competition.

Three studies of automotive displays (Dingus, Hulse, et al., 1997; Dingus, McGehee, et al., 1997; Liu, Schreiner, & Dingus, 2000) have provided collective evidence for a redundancy gain above the single-modality condition, while none showed conclusive evidence for auditory preemption. Only Van Gent (1995) in the second phase of his data-link study examined redundant presentation of data-link information. Although such presentation supported faster responding to the data-link message than did a visual-only condition, Van Gent did not measure navigation and aviation performance in a way that would allow him to assess resource competition or preemption.

LITERATURE REVIEW SUMMARY AND PROPOSED EXPERIMENT

Our review of the literature suggests that the replacement of visual for auditory delivery of information can impose a mixed blessing on attention. On the one hand, visual display of lower priority communications information may be less disruptive

of the other ongoing (higher priority) visual aspects of the flight task, given that the visual modality is less interruptive (i.e., avoids “auditory preemption” of visual aviating and navigation tasks). However, on the other hand, the visual resources necessary to process a text data-link message could well compete with the visual resources necessary to perform these other higher priority aspects of the flight task, like monitoring instruments and scanning out the windscreen for traffic. Understandably, these attentional factors, acting both for and against the text based data link display, may be amplified in single pilot operations of general aviation because there is only one pair of eyes involved, and in visual meteorological conditions (VMC), when these eyes should be actively involved in out of cockpit scanning.

The purpose of the experiment we describe here was to compare, in a single pilot VMC simulation, three different display configurations of ATC information: a text based data link display, a format presenting the same information only in synthesized voice, and in a format providing the redundant and identical display in both modalities. To impose limits on the auditory-only conditions, demonstrating the potential advantages of the two systems involving visual text (visual only and redundant), we systematically varied the length of the ATC messages, imposing progressively greater demands on working memory. Furthermore, we also assessed two versions of the auditory-only condition, one allowing pilots to write information on a clipboard if they desire (simulating the characteristics of normal flight) and the other requiring readback based only on working memory (memory condition).

We anticipated memory failures, as reflected in readback errors, to dominate the auditory-only condition, relative to the two visual conditions, particularly as message length is increased. However, to examine potential costs to the ongoing visual tasks in aviation, three further dependent variables were assessed. First, visual scanning was measured, with the expectation that visual attention may be pulled downward (at the expense of the outside world) in the visual-only condition. Second, we imposed outside traffic that had to be detected at or around the time that the data-link message arrives, hence enabling us to quantify the possible penalties of attention diversion or head down time. We considered these as penalties suffered by the navigation subtask, given the “see and avoid” navigational responsibilities associated with flying in VMC. Third, measures of flight path tracking were extracted, to evaluate the costs of visual attention that might be diverted from the instrument panel and hence the primary task of “aviate.”

METHODS

Participants

Fifteen certified flight instructors, 12 males and 3 females, were recruited from the Institute of Aviation at the University of Illinois in Urbana Champaign. The pilots’

ages ranged from 21 to 45 years, with a mean of 27.8 years. The total flight hours ranged from 240 to 3,580 hr, with a mean of 941.3 hr, and the range of instrument flight rules (IFR) flight hours was 60–500 hr, with a mean of 125.3 hr. The pilots were paid \$8.00/hr with a \$5.00 bonus for completing the full experiment. All participants flew in all of the experimental conditions, yielding a completely within-subjects design.

Equipment and Display

Flight simulator. Pilots flew a Frasca 142 flight simulator configured as a single engine Beechcraft Sundowner, which included the full set of instruments on the instrument panel along with a radio stack. Figure 1 shows a schematic of the layout of the Frasca instrument panel along with the locations of the data link display and clipboard (the superimposed eye scanning data is discussed later). An Evans and Sutherland SPX 2400 was used to project a 135° view of the outside world. This system was capable of depicting traffic at a range of up to 5 nm away from the participant's aircraft. Figure 1 also depicts the locations of the projection screens,

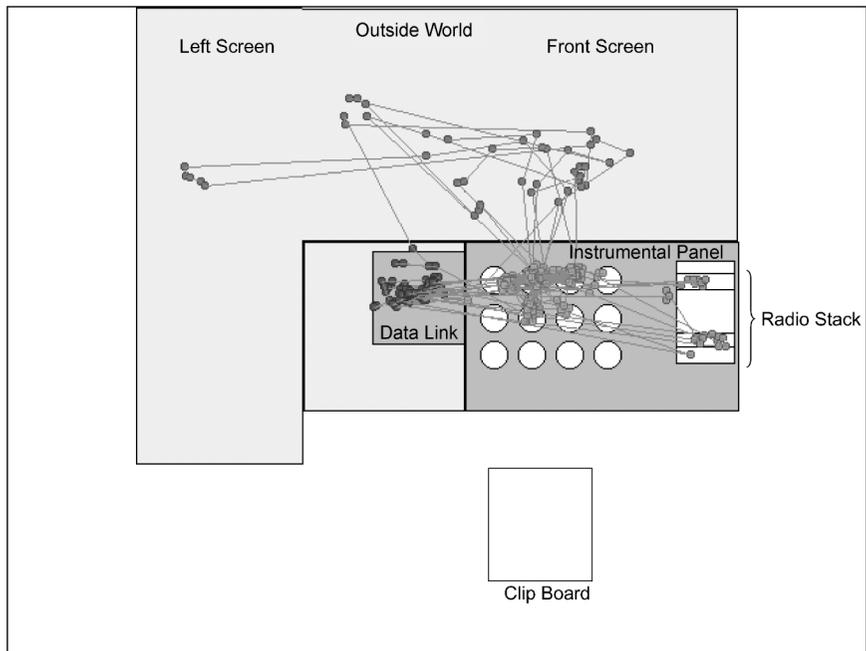


FIGURE 1 Instrument panel, data link display, and clipboard locations with sample eye-scanning data superimposed.

relative to the Frasca simulator. A voice synthesizer was used to play prerecorded simulated data link ATC instructions. A silicon graphics IRIS workstation with a 20-in. color monitor having a screen resolution of 1280×1024 pixels running at 60 Hz was used to display the data link text messages. The data link display presented the visual messages within a window of 12° of horizontal visual angle and 8° of vertical visual angle. Pilots were provided with a standard pencil and clipboard to write down the ATC instructions during the auditory condition. The experimenter recorded any errors or clarification requests that the pilots made.

Head-mounted eye-head tracker. Pilots were fitted with the eye-head tracker for the second half of the experiment (flight numbers 4, 5, and 6). But during the final flight (flight number 7, the memory condition) the eye/head tracker was removed. The eye-head tracker was an Applied Science Laboratories Model 501 head mounted eye tracking system with integrated magnetic head tracker. The eye-tracking system is sampled at 60 Hz with an accuracy of better than 1° of field view. The head tracker tracks head position with 6° of freedom of movement for the head.

Task

The pilot's task was to fly seven simulated IFR cross-country flights in VMC, each of which took approximately 25 min to complete, using each of three different display formats (auditory, visual, and redundant) for the ATC clearance information. There were six unique flight scenarios, with the first scenario that each pilot flew being repeated for the seventh and final flight (memory condition). The three display conditions were presented two times each in a random order, with eye movements recorded during the second presentation. The final flight (memory condition) used an auditory format, but the pilots were not allowed to write down the clearances. This condition was included to illustrate the cost to communication efficiency that results from the limits of working memory and does not represent a likely implementation of data link.

Figure 2 shows a graphic depiction of a typical flight used in this experiment. Each flight consisted of 11 waypoint to waypoint legs, with traffic aircraft interspersed between some of the waypoints. Each flight began with the aircraft at cruising altitude and required the pilot to make an initial maneuver to a new heading to complete the first leg. Each simulated ATC message consisted of between two and six of the following parameters: vectors (heading, altitude, and airspeed parameters), radio frequencies (communication and transponder), and an altimeter calibration setting. As noted, the auditory format employed a synthesized voice; the visual format employed a text based data link display; and a redundant format presented the same information using both formats simultaneously. Pilots were notified of an incoming ATC instruction by an alerting tone (PC speaker

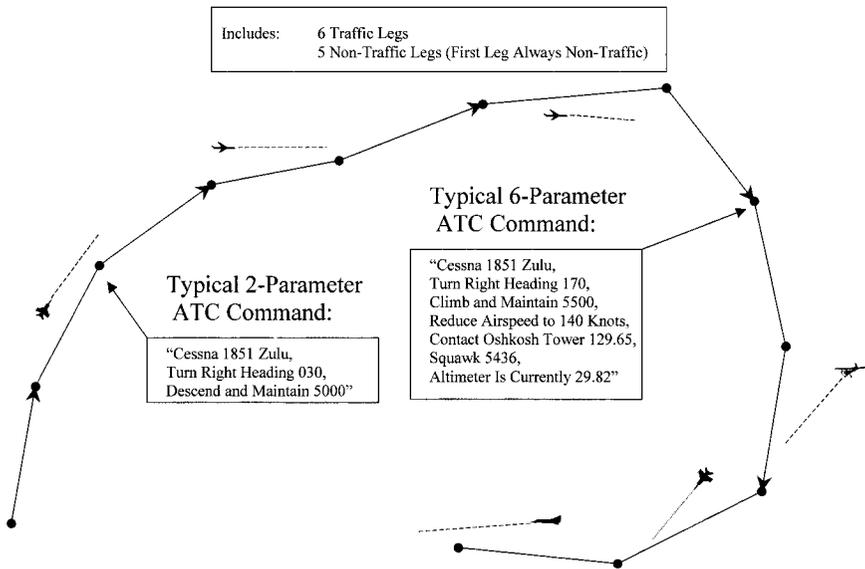


FIGURE 2 Typical cross-country flight.

beep) and the ATC instruction would be presented immediately following the alerting tone. The pilot was required to read back the entire ATC instruction including directions. These ATC instructions were issued at the beginning of the experiment and once the pilot reached each of the waypoints. The prerecorded auditory and/or text ATC instructions were automatically triggered when the pilot reached each waypoint, and could be repeated upon request of the pilot.

During each 11-leg simulated cross-country flight (see Figure 2) the pilot was randomly presented with six legs where a traffic aircraft would be visible in the projected outside world, at unpredictable and widely dispersed locations. The first leg of each flight did not have traffic, and the remaining four traffic-free legs were randomly dispersed among legs 2 through 11. The ATC instructions did not alert the pilots to any of the traffic aircraft. During the flight pilots were required to scan the projected outside world for traffic, and were required to callout "traffic in sight" whenever there was an airplane visible in the projected outside world.

Experimental Design

The traffic could be presented at three different times relative to the ATC instructions (see Figure 3). Fifteen traffic aircraft were presented "simultaneously" with the ATC instruction (4 sec after alerting tone). Eleven traffic aircraft were

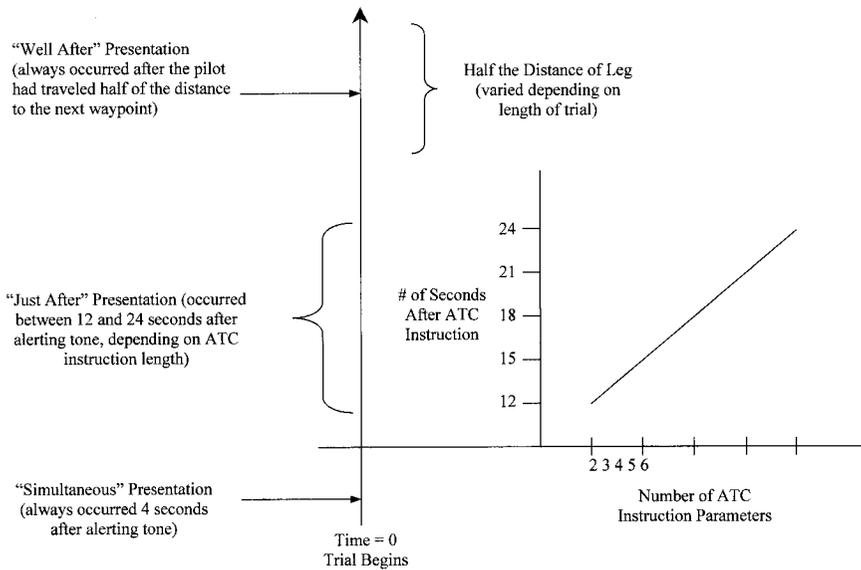


FIGURE 3 Timeline of typical trial showing the timing of traffic presentation.

presented “just after” the ATC instruction was estimated to have been processed while the pilots were executing the changes dictated by the instruction (between 12 and 24 sec after alerting tone). Ten traffic aircraft were presented “well after” the ATC instruction was assumed to have been processed while the pilots were maintaining a steady course.

The first leg of each flight scenario was always a six-parameter ATC instruction, while the remaining 10 legs consisted of two presentations of each ATC instruction length in a random order. There were a total of 77 ATC instructions per participant (11 of which were in the memory condition). This yielded a total of 1,155 legs of data that were collected (165 of which were from the memory condition).

Each of these factors was completely balanced across the participants and randomly presented throughout the six cross-country flights, which yielded a 3 (Display Format) \times 3 (Traffic Onset Time) \times 5 (ATC Instruction Length) factorial, within-subjects design. The order of presentation for each unique cross-country flight was counterbalanced using a Latin square.

Procedure

Participants could fly any of all seven experimental scenarios in two experimental sessions. Upon arrival, participants read and signed the informed consent form

and filled out a short demographic questionnaire. The participants were then given the experimental instructions to read. Once the participants understood the experimental task, they were asked to fly two short (two-leg) practice flights to familiarize them with the experimental procedure, the flight dynamics of the simulator, and the two data link display formats for the ATC instructions (the redundant condition was described, but not practiced). The practice flights also contained an example of the traffic that would be depicted in the outside world. The participant then proceeded to fly the first, second, and third experimental flights, with a short break between flights. Pilots could then take a longer break and continue the experiment, flying the remaining four flight scenarios, or return on another occasion to complete the last four experimental flights. For flight scenarios four, five, and six, the eye-tracking apparatus was fitted on the pilot and used to collect the eye-scanning data. However, no eye-scanning data were collected during the final flight scenario (memory condition). After completing the experiment the pilots were debriefed, paid for their time, and thanked for their cooperation.

RESULTS

Aviation: Flight Path Deviations

During each leg the pilot's flight path errors (root mean square [RMS] deviations from the ATC instructed heading, altitude, and airspeed) were measured. The pilot's most important subtask, flight path maintenance (aviate), was clearly disrupted by the incoming ATC instructions. A 3 (Display Format) \times 5 (ATC Instruction Length) within-subjects analysis of variance (ANOVA) on the heading deviation data (shown in Figure 4), revealed that there was a significant cost of increasing ATC instruction length, $F(4, 56) = 156.40, p < .01$. There was also a significant main effect of display format, $F(2, 28) = 15.94, p < .01$, on heading RMS error, such that pilots made the least amount of heading deviations in the visual condition followed by the redundant and auditory conditions respectively. Pairwise comparisons showed that the visual display format resulted in significantly less heading error than the other two display formats ($p < .01$), and that the redundant display format resulted in marginally less heading error than the auditory display format ($p = .08$).

There was a similar pattern of results for both the altitude and airspeed RMS errors. Figures 5 and 6 show that when ATC instruction length increased, altitude and airspeed RMS errors also increased. A pair of 3 (Display Format) \times 5 (ATC Instruction Length) within-subjects ANOVAs revealed that there was a significant main effect of ATC instruction length on altitude, $F(4, 56) = 190.88, p < .01$, and airspeed, $F(4, 56) = 99.38, p < .01$, RMS errors, respectively.

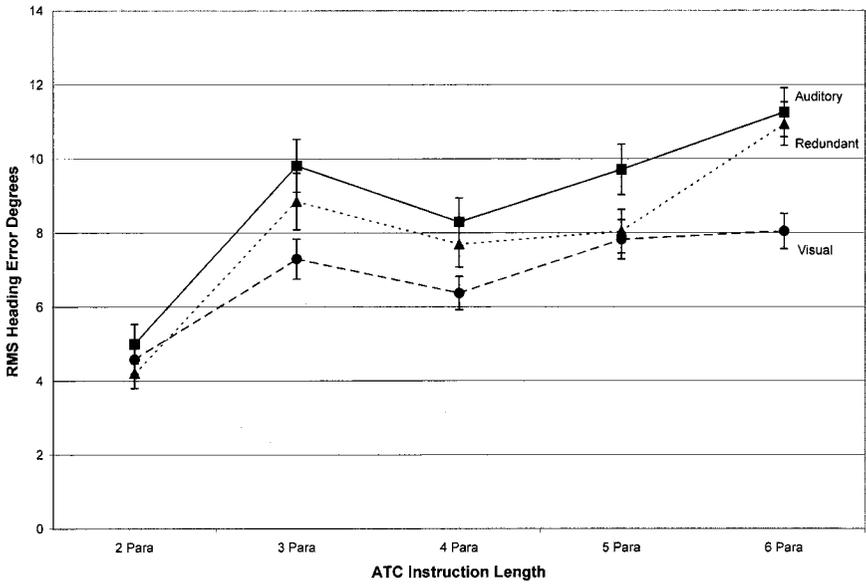


FIGURE 4 Heading root mean square error as a function of display modality and air traffic control instruction length.

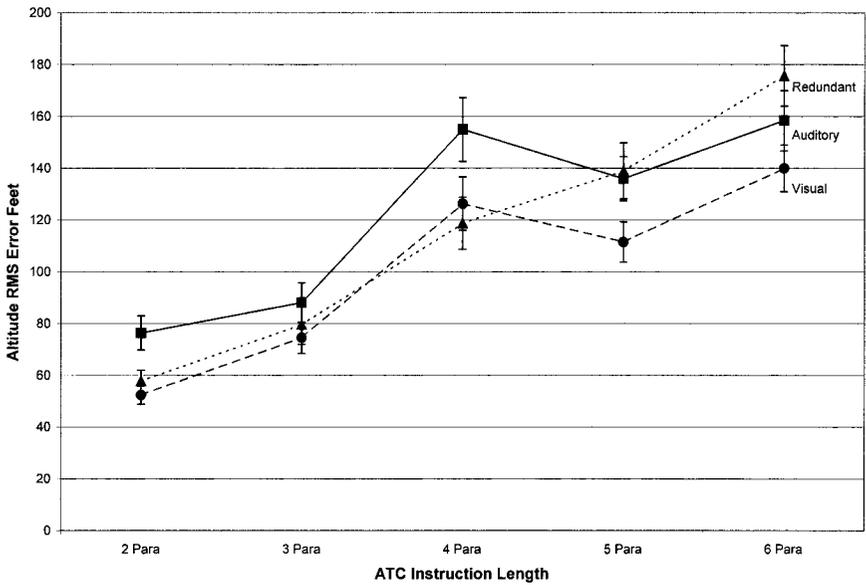


FIGURE 5 Altitude root mean square error as a function of display modality and air traffic control instruction length.

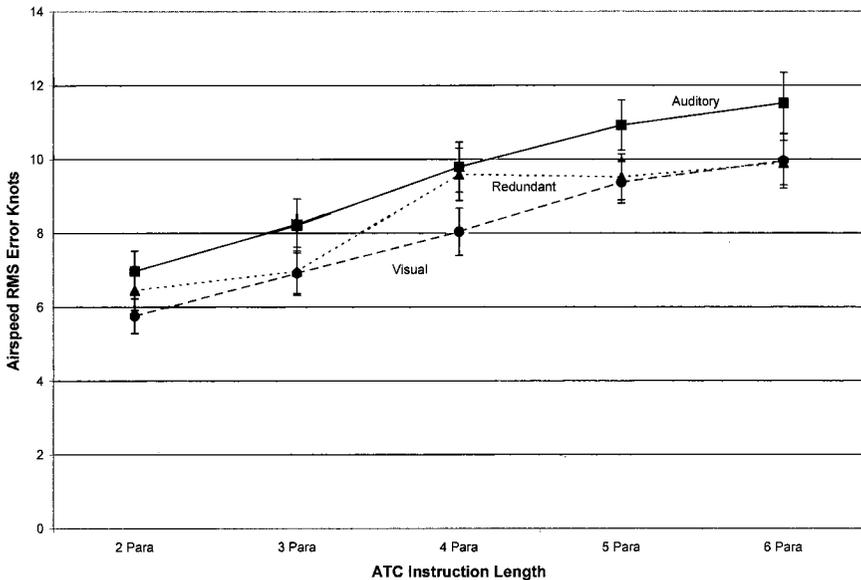


FIGURE 6 Airspeed root mean square error as a function of display modality and air traffic control instruction length.

Figures 5 and 6 also show that, similar to the heading RMS error, there was a significant main effect of display format for both altitude, $F(2, 28) = 9.54$, $p < .01$, and airspeed, $F(2, 28) = 6.34$, $p < .01$, RMS errors. Pairwise comparisons showed that the visual display format resulted in significantly less altitude error than did the other two display formats ($p < .01$). Similarly, pairwise comparisons showed that for the airspeed error the visual display format performed significantly better than the auditory display format ($p < .01$) and that the redundant display format was marginally better than the auditory display format ($p = .06$). There were no significant interactions between ATC instruction length and display format for either measure.

Navigation: Traffic Detection Latency

Traffic detection accuracy was 79%, and no significant differences were found across the independent variables. Traffic detection latency was calculated by measuring the time from when the traffic was generated in the outside world until the pilot detected and called out "traffic in sight."

Table 1 provides the mean traffic detection latency across all four of the display formats (as well as the ATC readback errors discussed a little later). The memory

TABLE 1
Mean Traffic Detection Latency and Proportion of Readback Errors as a Function
of Modality

<i>Modality</i>	<i>M</i>	<i>SE</i>	<i>Proportion</i>	<i>SE</i>
Auditory	38	2.0	0.10	0.008
Visual	30	1.8	0.03	0.004
Redundant	34	2.2	0.04	0.005
Memory	36	3.2	0.20	0.016

condition was included in the table to illustrate the similarity in traffic detection times that occurred across all display formats, which used auditory ATC instruction information, and to show the cost to communication efficiency that results from the limits of working memory. However, the data from the memory condition ($N = 90$) were not directly relevant for drawing conclusions related to the hypotheses of this study and therefore were removed prior to any further statistical analyses. The remaining data ($N = 519$) were subjected to a series of ANOVAs.

A 3 (Display Format) \times 3 (Traffic Onset Time) between-subjects ANOVA showed that there was a significant main effect of display format, $F(2, 398) = 4.16$, $p < .02$. Table 1 shows that the visual format resulted in the fastest detection latencies followed by the redundant and auditory formats, respectively, suggesting there is a cost to traffic detection resulting from the auditory presentation of ATC instructions. There was also a significant main effect of traffic onset time, $F(2, 398) = 58.30$, $p < .01$, indicating that traffic detection latency decreased as the traffic onset time occurred later, relative to the presentation of the ATC instruction. There was no interaction between display format and traffic onset time.

Communication: ATC Readback Errors

Several categories of communication errors were recoded into a single measure indicating the proportion of errors out of the total message length, made for each trial (see Helleberg & Wickens, 2000). The recoding involved giving "partial credit" for communications that, for example, had the correct value, but an incorrect direction. Table 1 depicts the total proportion of communication errors as a function of all four of the display formats. The table shows that the proportion of errors was greatest in the memory condition (i.e., without the note taking clipboard). However, again the data from the memory condition ($N = 165$) were not directly relevant for drawing conclusions related to the hypotheses of this study and therefore were removed prior to any further statistical analyses. For the remaining three conditions, Table 1 shows that the greatest number of errors occurred with the auditory display format, $F(2, 28) = 17.52$, $p < .01$, followed by

the redundant and visual display formats, whose performance is equivalent). There was a generally monotonic trend indicating that as the length of the ATC instructions increased so did the proportion of communication errors across all three display formats, $F(4, 56) = 2.22, p = .08$. There was no interaction between ATC instruction length and display format.

Eye-Scanning Strategy: Percentage Dwell Time

Figure 1 provides an actual example of the type of eye-scanning data that were captured on each leg of the experimental flights four, five, and six. Each dot represents a single fixation and the connecting dashed lines indicate where the eyes moved on subsequent fixations. There were five “areas of interest” (AOIs) where the pilot could fixate. These were the “outside world (OW), instrument panel (IP), radio stack (RS), data link (DL), and clip board (CB). Percentage dwell time (PDT) was the percentage of time that a pilot’s eyes remained in a particular AOI relative to the other AOIs (Wickens et al., 2002).

Figure 7 presents the PDT as a joint function of ATC instruction length, data link display format, and the AOI into which a fixation fell. The instrument panel AOI is shown at the top and the outside world AOI is shown in the middle of the figure. At the bottom of the figure is what we have labeled as the *ATC Information AOI*. This is defined as the DL display for the visual and redundant display formats, but as the CB for the auditory display format. We consider these sources as functionally equivalent, because both (CB and DL displays) represented regions within which ATC information was visually represented.

A 2 (AOI) \times 3 (Display) \times 5 (ATC Instruction Length) ANOVA carried out only on the IP and OW scans revealed the clear dominance of the IP (59%) over the OW (21%) in the allocation of visual attention, $F(1, 11) = 30.04, p < .01$. The ANOVA also revealed a significant effect of display format, $F(2, 22) = 9.15, p < .01$, revealing a lower percentage of time spent on both AOIs, when the auditory presentation was used (IP = 56%, OW = 19%), compared to the two visual display formats (visual and redundant: IP = 60%, OW = 22%). Display format did not interact with the other variables. Finally the ANOVA revealed a significant interaction between AOI and ATC instruction length, $F(4, 44) = 4.62, p < .01$, such that longer instructions led to a lower percentage of time spent on the outside world, but no change in the percentage of time spent on the IP.

Although the ATC information AOI data, shown by the three lines at the bottom of the figure, were not included in the ANOVA (because this involved mixing two different data sets), the visual trend shown by these data are consistent in two respects with the picture painted for the OW and the IP. First, the ATC information source demanded the least amount of visual attention of the three AOIs (consistent with the role of “communication” as third on the task priority hierarchy). Second,

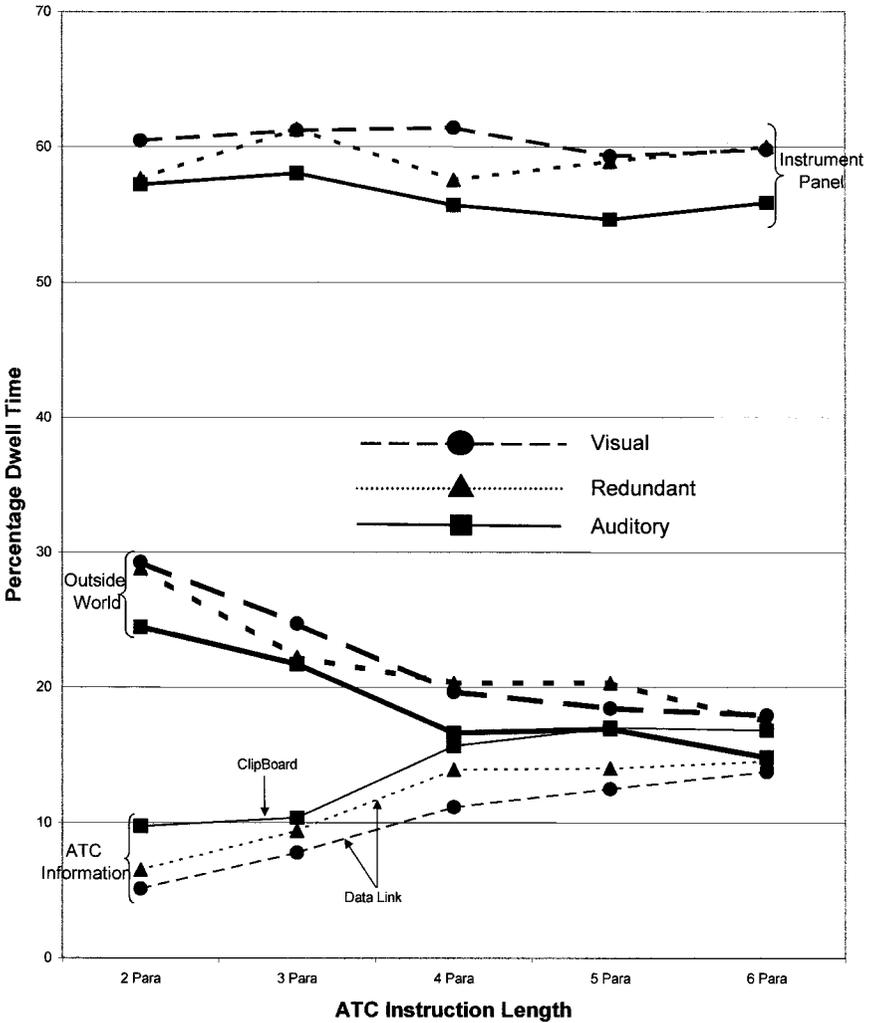


FIGURE 7 Percentage dwell time as a function of areas of interest, display modelling, and instruction length.

it is apparent that the drop in visual attention to the OW as message length is increased is a consequence of the greater amount of visual attention allocated to the ATC information source with longer ATC instructions; that is, the two bottom sets of data in Figure 7 show a complementary relationship with each other as message length is increased while fixations on the IP are relatively unchanged by message length.

DISCUSSION

An overall comparison of the three key performance variables and the attention allocation visual scanning measures, as a function of the three display formats, is represented in Table 2, in which a “+” indicates that the display format supported good performance relative to the others, a “-” indicates poor performance relative to the others, and a “0” indicates intermediate performance relative to the other display formats. This judgment of relative preference is based jointly on the statistical significance of the differences reported earlier, and upon their magnitude. Furthermore, within the rows depicting the amount of visual scanning a “↑” indicates that relatively more visual scanning occurred and a “↓” indicates relatively less visual scanning occurred within that display condition.

On examining the table, several overall features are evident: first, the visual display format yields the best performance, the auditory display format yields the worst performance, and performance with the redundant display format is intermediate. Second, for each of the aviation subtasks (aviation, communication, and navigation) the pattern of visual scanning measures is consistent with the ordering of the display formats; that is, a greater allocation of visual resources to the IP and OW leads to better flight path tracking and traffic detection performance, whereas allocating fewer visual resources to the ATC information source resulted in fewer communication errors. The overall pattern of effects revealed in the table can be interpreted within the context of three information processing mechanism as described hereafter: preemption or attention switching, multiple resources, and

TABLE 2
Summary of Results, Depicting the Relative Difference Between Measures

<i>Aviation Subtask</i>	<i>Display Format</i>		
	<i>Auditory</i>	<i>Redundant</i>	<i>Visual</i>
<i>Aviation</i>			
Flight path control	-	0	+
IP scanning	↓	↑	↑
<i>Navigation</i>			
Target detection	-	0	+
OW scanning	↓	↑	↑
<i>Communication</i>			
ATC instruction read back	-	0	+
CB and DL scanning	↑	↓	↓

Note. Judgments based upon statistical significance and magnitude relative to the other display formats. Scanning rows indicate relative amount of allocated visual attention. IP = instrument panel; OW = outside world; CB = clipboard; DL = data link; “+” = good performance; “-” = poor performance; “0” = intermediate performance; “↓” = less visual scanning; “↑” = more visual scanning.

working memory constraints. We use these mechanisms to account for performance on each of the three aviation subtasks.

Aviation Subtask

We have operationally defined this highest priority task of aviating in terms of the pilots' ability to maintain the prescribed flight path in vertical, lateral, and air-speed components, as achieved by allocating visual attention to the IP. The results (Figures 4–6) suggest that the visual format supports the best performance, the auditory format supports relatively poor performance, and the redundant format supports intermediate performance. The poorer performance seen within the auditory display format can be at least partially explained by a decrease in visual attention allocation to the IP (approximately 56%) and a corresponding increase in visual attention allocated to the CB (approximately 14%) for note taking. With the auditory only data link display, pilots were not only required to allocate visual attention to the task of reading back the ATC instructions (as was also required in both visual conditions), but in addition to this they were also required to allocate visual attention to note taking while the ATC instruction was being played.

The manual task of note taking could also have contributed to the disruption of the ongoing flight path maintenance task. The task of aviating did require the participants to respond manually (i.e., deflect simulated control surfaces through the yoke, rudder pedals, and throttle). However, this task also had a significant requirement for visual attention as is indicated by the clear majority of visual attention being dedicated to the IP. Furthermore, the auditory-only condition also yielded the worst performance on the traffic detection task, which was a visual task and did not require any manual responding.

Although resource competition had an important influence, it was not the only factor affecting the pilot's task of aviating. It is evident that the differences in performance across the two visual display formats (visual-only and redundant) are observed despite the essential equivalence of visual attention allocated to the IP (approximately 60% for each), as well as the data link display (approximately 11% for each), across these two visual display formats. We attribute this pattern of effects on the aviate subtask to the auditory preemption or interruption effect within the redundant condition, by which discrete auditory events have a tendency to interrupt ongoing continuous visual tasks, as described in basic dual task paradigms by Wickens and Liu (1988), and observed directly in the data-link simulation by Latorella (1998). The fact that this pattern of disruption was entirely echoed by the visual attention allocation measure, suggests that this preemption is a qualitatively different phenomenon from the visual resource competition, which, as we describe later, appears to underlie the disruption of visual traffic target detection. Further evidence for the role of an auditory preemption mechanism,

which is distinct from visual resource competition, is the fact that the preemption effect appears to be common to both formats that employ the auditory display.

Navigation Subtask

We have defined the traffic detection task as a navigation subtask because of the importance of hazard (traffic) awareness in safe navigation. It is apparent that this subtask suffered, as a consequence of the greater head down time and visual resource demand (relative to the data link display) imposed by the clipboard note taking in the auditory-only condition. Our analysis also revealed that dwells on the CB (2.1 sec) were longer than those on the visual display (1.3 sec), and correspondingly, the auditory display availed less scanning out the window (19%) than did the two visual conditions (22%). The traffic detection RT cost for the auditory display format was not large, but was nevertheless in a direction that supports the visual display of data-link information, relative to conditions in which pilots must use note taking to compensate for auditory working memory constraints. We were somewhat surprised that the redundant format did not yield better performance (and more visual resources outside) than the visual format, given that the redundant format should allow pilots to continue their outside scan, and only look down to refresh their memory for the longer messages. It is possible that this failure of redundancy to offer advantages might have resulted from the preemption or interruption characteristics of the auditory display, described earlier in the context of the aviation subtask. In addition, pilots using the redundant display may have been reluctant to rely on their working memory for the ATC instruction read back and instead may have chosen to consult the visual display on each leg. This head down time coupled with the interrupting nature of the auditory information may have prevented the realization of the “best of both worlds” with the redundant display.

Communication Subtask

The failure of the auditory display to support good performance on the readback task, relative to the two formats using visual displays (redundant and visual), is not surprising, and indeed documents one of the major reasons why the data-link system has been advocated in the first place (Kerns, 1999). Less expected, however, was the failure of pilots to be able to use the clipboard and note taking facilities to fully compensate for the frailties of their working memory with the auditory display. Furthermore, the use of the clipboard demanded more visual resources (14%) than did the readout of the visual data link display (11%), a cost imposed on both outside world and IP scanning, which hampered traffic detection

and flight path maintenance. It is likely that pilots could have performed the communication task with equal accuracy across display conditions, but because of the task priority hierarchy, they were hesitant to take visual resources away from the primary tasks of aviating and navigating to reallocate those resources to the lowest priority task of communication.

CONCLUSION

It appears that the proposed visual data link display is indeed a format that supports best overall performance for general aviation pilots. Despite the requirement that this format imposes for head down activity, the permanence of the visual display allows pilots to allocate that head down time in a manner that is more flexible and less disruptive of ongoing visual tasks, than that which results when there is a need to process (and take notes on) the auditory transmission of data. The failure of the redundant use of both channels to support “the best of both worlds” is surprising. This failure may result in part from the preemptive aspects of the auditory component of the redundant display. It is however possible that a redundant format could provide advantages if training was offered as to the appropriate use of each display modality within the redundant pair, an issue that requires further investigation.

ACKNOWLEDGMENTS

This research was supported by a Federal Aviation Administration (FAA) grant DTFA 98-G-022. Dr. Dennis Beringer was the scientific/technical monitor. The authors acknowledge the invaluable assistance of Roger Marsh and Ron Carbonari for developing the software for this simulation.

REFERENCES

- Baddeley, A. D. (1990). *Human memory: Theory and practice*. Hove, UK: Lawrence Erlbaum Associates, Inc.
- Dingus, T. A., Hulse, M. C., Mollenhauer, M. A., Fleischman, R. N., McGehee, D. V., & Manakkal, N. (1997). Effects of age, system experience, and navigation technique on driving with an advanced traveler information system. *Human Factors*, 39(2), 177–199.
- Dingus, T. A., McGehee, D. V., Manakkal, N., Jahns, S. K., Carney, C., & Hankey, J. M. (1997). Human factors field evaluation of automotive headway maintenance/collision warning devices. *Human Factors*, 39(2), 216–229.
- Dismukes, K. (2001). The challenge of managing interruptions, distractions, and deferred tasks. In *Proceedings of the 11th International Symposium on Aviation Psychology*. Columbus, OH: The Ohio State University.

- Fiorino, F. (2002, October 14). Controller-pilot data link goes live in Miami. *Aviation Week and Space Technology*, 40–41.
- Flathers, G. W. (1987). *Development of an air ground data exchange concept: Flight deck perspective* (NASA Contractor Rep. 4074). Hampton, VA: NASA Langley Research Center.
- Helleberg, J., & Wickens, C. D. (2000). *Pilot expectancy and attentional effects for hazard awareness: Effects of data link modality on pilot attention and communication effectiveness* (Tech. Rep. ARL-00-7/FAA-00-4). Savoy, IL: University of Illinois, Aviation Research Laboratory.
- Kerns, K. (1991). Data link communications between controllers and pilots: A review and synthesis of the simulation literature. *The International Journal of Aviation Psychology*, 1, 181–204.
- 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: Lawrence Erlbaum Associates, Inc.
- Latorella, K. A. (1996). Investigating interruptions: An example from the flightdeck. In *Proceedings of the 40th Annual Meeting of the Human Factors and Ergonomics Society* (pp. 249–253). Santa Monica, CA: Human Factors and Ergonomics Society.
- Latorella, K. (1998). Effects of modality on interrupted flight deck performance: Implications for data link. In *Proceedings of the 42nd Annual Meeting of the Human Factors and Ergonomics Society* (pp. 87–91). Santa Monica, CA: Human Factors and Ergonomics Society.
- Lee, A. T. (1989). Human factors and information transfer. In *Proceedings of the Second Conference, Human Error Avoidance Techniques* (pp. 43–48). Warrendale, PA: SAE International.
- Liu, Y. C., Schreiner, C. S., & Dingus, T. A. (2000). The effect of advanced traveler information display modality on driver performance. In *Proceedings of the IEA 2000/HFES 2000 Congress*. Santa Monica, CA: Human Factors and Ergonomics Society.
- Loftus, G., Dark, V., & Williams, D. (1979). Short-term memory factors in ground controller/pilot communication. *Human Factors*, 21, 169–181.
- Mackintosh, M. A., McGann, A., Lozito, S., & Logsdon, E. (1999). Information transfer in data link communication for ATC clearances. In *Proceedings of the Tenth International Symposium on Aviation Psychology* (pp. 136–142). Columbus: The Ohio State University.
- McGann, A., Morrow, D., Rodvold, M., & Mackintosh, M. A. (1998). Mixed-media communication on the flight deck: A comparison of voice, data link, and mixed ATC environments. *The International Journal of Aviation Psychology*, 8, 137–156.
- Monan, W. P. (1986). *Human factors in aviation operations: The hearback problem* (NASA Contractor Rep. 177398). Moffett Field, CA: NASA Ames Research Center.
- National Research Council (1998). *The future of air traffic control human operators and automation*. Washington, DC: National Academy Press.
- Navarro, C., & Sikorski, S. (1999). Data link communication in flight deck operations: A synthesis of recent studies. *International Journal of Aviation Psychology*, 9(4), 361–376.
- Posner, M. I., Nissen, J. M., & Klein, R. (1976). Visual dominance: An information processing account of its origins and significance. *Psychological Review*, 83, 157–171.
- Roitsch, P. A., Babcock, G. L., & Edmunds, W. W. (1978). *Human factors report on the Tenerife accident*. Washington, DC: Air Line Pilots Association (ALPA) Engineering and Air Safety Department.
- Schutte, P. C., & Trujillo, A. C. (1996). Flight crew task management in non-normal situations. In *Proceedings of the 40th Annual Meeting of the Human Factors and Ergonomics Society* (pp. 244–248). Santa Monica, CA: Human Factors and Ergonomics Society.
- Van Gent, R. N. (1995). *Human factors issues with airborne data link* (NLR Tech. Publication 95666L). Amsterdam, Netherlands: National Aeronautics Laboratory.
- Wickens, C. D. (1980). The structure of attentional resources. In R. Nickerson (Ed.), *Attention and performance* (Vol. 8, pp. 239–257). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

- Wickens, C. D. (1991). Processing resources and attention. In D. Damos (Ed.), *Multiple-task performance* (pp. 3–34). London: Taylor & Francis.
- Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomic Science*, 3(2) 159–177.
- 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.). New York: 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., Mavor, A. S., Parasuraman, R., & McGee, J. P. (Eds.). (1998). *The future of air traffic control: Human operators and automation*. Washington, DC: National Academy Press.
- Wickens, C. D., Sandry, D. L., & Vidulich, M. (1983). Compatibility and resource competition between modalities of input, central processing, and output. *Human Factors*, 25, 227–248.

Manuscript first received April 2001