

TRAFFIC AND DATA LINK DISPLAYS: AUDITORY? VISUAL? OR REDUNDANT? A VISUAL SCANNING ANALYSIS.

John Helleberg
MITRE CAASD Corporation
McLean, Virginia

Christopher D. Wickens and Juliana Goh
University of Illinois at Urbana-Champaign
Institute of Aviation

Twelve pilots flew a full mission simulation to examine two new display technologies: digital datalink, and the cockpit display of traffic information (CDTI). Performance with both of these information sources displayed in both visual formats was compared with performance when equivalent information was displayed auditorally and redundantly combining both modalities. Flight path tracking was disrupted by higher outside traffic load, and by the visual display of both CDTI and data link information relative to the auditory display. This disruption was a consequence of the visual attention demands of these devices, which pulled attention away from viewing the outside horizon, and led to less scanning of the instrument panel. Outside traffic was detected at around a 90% rate, was generally not helped by cuing from the CDTI, and was slightly hurt by the visual attention demands of the CDTI, particularly when traffic was not depicted on the CDTI. Pilots generally protected their scanning to the instrument panel, from the increased demands of other concurrent task elements. Comprehension of longer ATC instructions suffered with the auditory display. The redundant display failed to provide better performance across the collective tasks of aviating (flight path control), navigation (traffic detection) and communications (message comprehension) than the best of the single modality conditions, and was inferior to these in terms of disrupting altitude control.

Introduction

Data link and the cockpit display of traffic information (CDTI) are two new technologies designed to replace, or alter more traditional information exchanges between the pilot and Air Traffic Controller (ATC). The conventional data link proposal is to uplink text-based instructions to the cockpit (Navarro & Sikorsky, 1999; Kerns, 1999; Helleberg & Wickens, 2002). A range of uses have been proposed for CDTI, from strategic maneuvering in a free-flight scenario to augmenting ATC traffic point outs. Data Link and CDTI are both envisioned to transfer the delivery of information from the conventional auditory channel, to a visual channel within the cockpit.

Auditory Versus Visual Delivery

Research suggests that auditory and visual systems appear to define separate processing resources (Wickens, 1991; Wickens & Liu, 1988; Vidulich & Wickens, 1986) and distributing total load across the two channels (conventional ATC) avails more opportunity for parallel processing, than does concentrating more processing within the visual system, as with the proposed technologies. Yet mitigating this cost, are two additional factors, related to compatibility and preemption.

Research has indicated that verbal and linguistic information (ATC instruction) is more compatible when delivered in the auditory, than in the visual (text)

channel (Wickens, Sandry, & Vidulich, 1983; see Vidulich & Wickens, 1986 for a summary), as long as the message is short; however, spatial information (target cueing) is less compatibly served by the auditory channel. Thus the CDTI may be better suited to convey target search and cueing information.

The second mitigating factor, preemption, appears to penalize the auditory delivery of complex information, as the auditory modality has an intrinsic attention-grabbing quality that may pull attention away from ongoing visual tasks. Research has shown that the auditory delivery of discrete information is more disruptive of ongoing visual tasks than is the visual delivery of the same information (Wickens & Liu, 1988; Latorella, 1998; Helleberg & Wickens, 2002). While part of this disruption may relate to the initial attention-capture properties of the discrete information, another part may relate to the limited capacity of working memory required to retain complex auditory information.

Visual Traffic Detection

Previous studies have found traffic detection rates of around 65% regardless of whether traffic was announced by ATC (Prinzo, 2001), or unannounced (Marshall & Fisher, 1959; Wickens, Helleberg, & Xu, 2002). The FAA has recommended "optimal scanning strategies" that they believe will enhance the visual sighting of traffic (U.S. Department of Transportation, 2001). These organizations recommend for example,

that roughly 70% of a pilot's visual attention be directed outside the cockpit, during VFR operations. However, no studies were located which associated this strategy with specifically better detection performance. Those studies that have examined traffic scanning indicate that the ratio of scanning outside to inside is closer to 1:2 than the 2:1 ratio recommended by the FAA (Helleberg & Wickens, 2002; Wickens et al., 2002). However, this observed ratio may reflect the pilots' task hierarchy of aviate, navigate, communicate, systems (Schutte & Trujillo, 1996).

Traffic sighting is inherently a spatial task; therefore one might anticipate substantial benefits of the visual spatial CDTI over auditory verbal guidance. However, Prinzo, (2001) found only a modest traffic detection improvement with the CDTI over auditory cueing. Furthermore, Wickens et al., (2002), found that detection accuracy was essentially unchanged by the addition of CDTI relative to a no-CDTI condition. The muted benefits of the CDTI for traffic callout, can be explained in terms of multiple-resource competition (Wickens, 2002). That is, any advantages gained by the CDTI directing attention to find traffic, could be offset or overridden by the potential costs associated with increased head down time. However, the CDTI has other obvious benefits as a tool that pilots can use to plan traffic avoidance maneuvers, or share traffic awareness with ATC. Furthermore, a pilot who has "positively sighted" a traffic aircraft on the CDTI, has, in a sense, less urgency to positively sight it outside, than the pilot who has only an ATC verbal report. However one important circumstance where the visual head down time imposed by the CDTI will not be offset by its more precise guidance, and this is in the detection of traffic that is not known to the display-driver of the CDTI. This could be the so-called "unequipped aircraft," or transponder inoperable aircraft. In the following pages, we refer to this as the "rogue" aircraft. Its detection may be inhibited by both increased head down time and to the extent that a pilot expects the CDTI to guide attention to all traffic, may suffer still more from a sort of "complacency effect," or attentional tunneling.

The purpose of the current study is to examine data link and traffic guidance information, offered by auditory, visual, and redundant display in which scanning, message comprehension, traffic detection, and flight path tracking are all measured. The presentation of a rogue aircraft was implemented with careful sensitivity to assuring that a different pilot would see the exact same aircraft (geometry and aspect angle) in one of the other cueing conditions.

Methods

Participants

Twelve certified flight instructors, eleven males and two females, were recruited. The total flight hours ranged from 200 to 3700 hours, with a mean of 1183.8. The pilots were paid \$10.00 per hour.

Equipment and Displays

Pilots flew a Frasca 142 flight simulator. Figure 1 shows the instrument panel, location of the projection screens, data link, and CDTI display, relative to the simulator, and shows an example of the traffic. An Evans and Sutherland SPX 2400 visual system was used to project the outside visual world. A Silicon Graphics IRIS workstation with a 20-inch color monitor was used to display the data link text messages and CDTI (see Wickens, Goh, Helleberg, & Talleur 2002, for details about the displays). Visual scanning was measured with an Applied Science Laboratories Model 501 eye tracking system.



Figure 1. Frasca instruments, data link and CDTI displays, and traffic close-up.

Task

The pilot's task was to fly six simulated Instrument Flight Rules (IFR) cross-country flights in VMC, using each of three different display formats for the ATC clearances and traffic locations. Each flight consisted of 11 waypoint to waypoint legs, with the odd legs being designated as "data link" legs and the even legs being designated as "traffic" legs.

Data Link Legs. A simulated ATC issued instructions, half of which were three parameters in length (heading, altitude, and airspeed) and the other half were single parameter instructions. ATC instruction length was blocked. The ATC instructions were presented in one of three formats. An auditory format employed a

synthesized voice; a 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, which was immediately followed by the ATC instruction. The pilot was required to read back the entire ATC instruction and then execute the maneuver. The experimenter recorded any errors or clarification requests made by the pilots.

Traffic Legs. Pilots were presented with five legs where traffic aircraft (one or four) would be visible in the projected outside world (and on the CDTI in the two visual conditions). During the auditory and redundant conditions the simulated ATC provided general traffic point outs, but did not provide traffic avoidance vectors. Within each flight, pilots were randomly presented with one leg where a traffic conflict would occur if the pilot did not execute an avoidance maneuver. These legs consisted of a single conflict aircraft, which would collide with the pilot's ownship if no maneuver were initiated. The four remaining traffic legs contained only non-conflicting traffic (2 legs: 1 traffic; 2 legs: 4 traffic aircraft. See Wickens, Goh, Helleberg, & Talleur 2002, for details about the traffic. Pilots were instructed to scan both visually and utilizing the CDTI for traffic that could pose a threat to their ownship and were required to callout "traffic in sight" whenever there was an airplane visible in the projected outside world. Each pilot was presented with one traffic aircraft, which was neither depicted on the CDTI nor called out by ATC. This "rogue" aircraft was intended to simulate a "transponder off" aircraft or one with malfunctioning CDTI equipment.

Experimental Design

Each of the above factors were completely balanced across the subjects and randomly presented throughout the six cross-country flights, which yielded a 3 (display format) X 2 (ATC instruction length) X 2 (traffic load) factorial, within-subjects design

Results

Traffic Detection Phase

Flight Path Tracking. Lateral tracking performance revealed a main effect of traffic load on heading error, $F(1,7) = 140$, $p < .01$. While there was no effect of modality, $F(2,14) = 1.444$, $p > .10$, the traffic load x modality interaction, $F(2,14) = 3.698$, $p = .05$, revealed that the increase in heading error with high traffic load was much greater (over twice as great) in the redundant AV condition than in the two single modality conditions (A and V). The results of the analysis of

altitude error revealed only a significant error increase with 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) compared to the auditory condition. Inspection of the data revealed that this cost to tracking was only observed at high traffic load. Indeed a separate one-way ANOVA conducted only on the high traffic load data revealed a significant cost of the visual 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 predicted within the next 45 seconds. 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 sec/leg (auditory condition) to an average of 28 sec/leg. This variable was not influenced by traffic load via either a main effect or an interaction.

Visual scanning analysis. The Percentage Dwell Time (PDT) was a measure of the percentage of time the scan was within each of the areas of interest. These data, averaged across the six conditions and three relevant AOI's, are shown in Figure 2 (a: low workload; b: high workload). Because the auditory modality condition had only two AOI's, a single 3x3x2 ANOVA on all three factors was not conducted. Instead, the following approach employed two separate 2x3x2 ANOVA's.

One ANOVA focusing only on the instrument panel (IP) and outside world (OW), as a function of modality (3 levels) and traffic load (2 levels) revealed 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%). There was also a significant AOI x modality interaction, $F(2,22) = 13.768$, $p < .01$, indicating that the OW fixations were about 20% higher for the auditory than for the two visual CDTI conditions. There was a significant interaction between traffic load and AOI, suggesting a redistribution of scan away from the IP at high traffic load, $F(1,11) = 42.026$, $p < .01$. Finally, there was a significant 3-way interaction between traffic load, AOI and modality, $F(2,22) = 3.433$, $p = .05$, suggesting that the reallocation of IP scanning, with increasing traffic, was to the OW in the auditory condition and to the CDTI in the two visual conditions.

A second analysis focused only on the two visual CDTI conditions (visual and redundant) and considered all three AOIs. A highly significant traffic load x AOI

interaction, $F(2,22) = 29.604$, $p < .01$, revealed that the consequences of high traffic load are to significantly increase CDTI scanning, $t(11) = -7.856$, $p < .01$, significantly reduce IP scanning, $t(11) = 6.201$, $p < .01$, and marginally increase OW scanning ($p = .10$). There was also slightly (4%) more OW scanning in the AV than in the V condition, although t tests revealed that this effect was not statistically significant ($p > .10$).

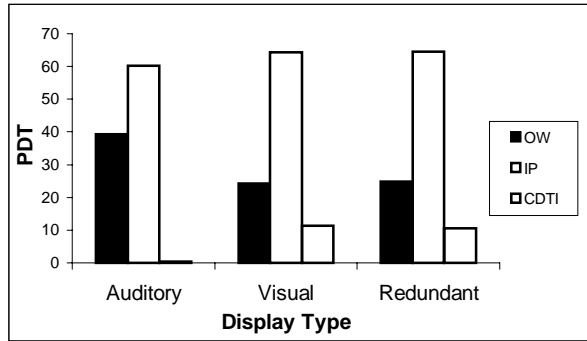


Figure 2a. Percentage dwell time on 3 AOIs with 1-plane across modality.

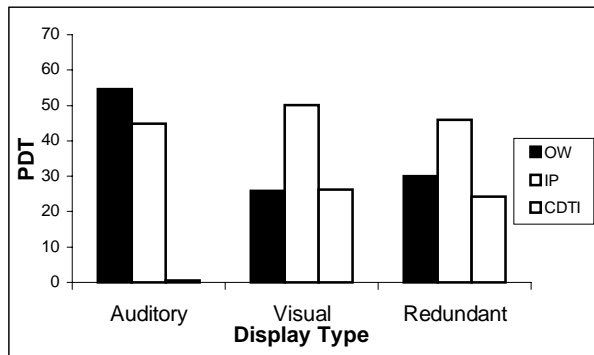


Figure 2b. Percentage dwell time on 3 AOIs with 4-planes across modality.

Rogue aircraft. In this analysis, the main effect is expressed as a cost (or benefit) relative to RT for the same aircraft, viewed by a different pilot when it appeared in a non-rogue state. The results of this analysis suggest that the increase in OW scanning associated with the high traffic load condition, appears to have improved detection of the rogue aircraft, from a 12-38.7=26.7 sec cost at low traffic load to a 48-45.5 = 2.5 sec benefit at high traffic load. This benefit for detecting rogue aircraft at high traffic load could be the result of 7% more scanning outside, or simply increased pilot vigilance, on trials with more traffic. Also, 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: an 11 second benefit; versus a

36 second cost for the V condition, and only a 5 second benefit for the AV condition.

Communications Phase

Flight path tracking. Heading was disrupted significantly, $F(1, 11) = 261$; $p < .01$, almost doubling, as a consequence of the longer communications strings. However heading control was not influenced at all by modality ($F < 1$). In contrast, vertical tracking was unaffected by communications load, but was substantially influenced by modality, $F(2,22) = 4.63$; $p < .05$. In this case, planned contrast, revealed that the effect - a benefit for auditory delivery over the two visual conditions - was in the opposite direction of the prior study (Helleberg & Wickens, 2002), a reversal attributable in part to the shorter messages employed here, reducing the need for visually guided note-taking.

Communication accuracy. The percentage of readback errors is shown in Figure 3, and reveals a main effect of length, $F(1,11) = 13.54$; $p < .05$, and of modality, $F(2,22) = 8.18$; $p < .01$, both of which can only be interpreted in the context of the significant length X modality interaction, $F(2,22) = 3.73$; $p < .05$. As is evident in the figure, when the communications length is short, there is no penalty for auditory delivery. However lengths of 3 chunks, apparently sometimes exceeded the capacity of working memory. These error rates were significantly higher in the auditory 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.

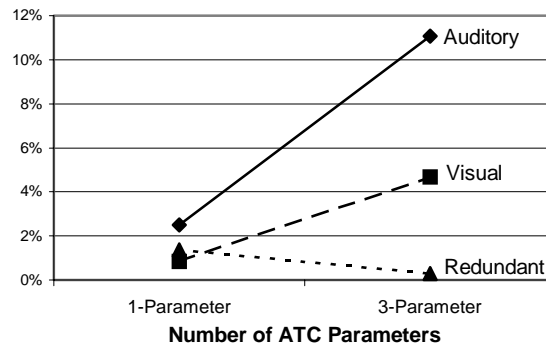


Figure 3. Number of readback errors as a function of ATC instruction length and datalink modality.

Visual scanning. As in the traffic phase, there was a large effect of AOI, with the IP receiving most attention (around 75%), the outside world next (around 15%) and the side display (here the data link display) receiving the least (around 8% in the two visual conditions). In addition, a significant interaction between modality and AOI, $F(2,22) = 6.97$; $p < .01$,

revealed that the consequence of adding the visual display, was to draw some attention to it, but at the sole expense of attention to the outside world. That is, full attention to the instrument panel was "protected" from the addition of the visual in-cockpit display.

Discussion

Traffic Detection Phase

The CDTI showed the greatest benefit in its role as a quasi-tracking display for assuring self-separation, its likely role for near term use. In contrast, CDTI traffic detection benefits were altogether absent. The auditory-only cueing did not hamper the speed of traffic call outs, and actually improved its accuracy somewhat, relative to the visual-only condition. Our scanning data suggest that the advantage of the auditory condition is in part due to the increased "eye out" time availed by not having a CDTI to look at. More direct evidence for this CDTI cost is provided by the disruption of flight path tracking performance, particularly at high workload, where the increased duration of scanning to the CDTI resulted in less sampling of the horizon attitude information, thus disrupting flight path tracking. One anticipated and observed effect of the CDTI "attention sink" was the decreased ability to detect the rogue aircraft. While the statistical power was low, the finding appeared to be fairly strong: a large 36 sec cost in the visual condition, that was not only absent in the auditory condition, but was replaced there by an 11 second benefit (relative to detection of the control aircraft). Interestingly, the redundant condition supported rogue detection in a manner more similar to the auditory-only than to the visual-only condition. We found little evidence here that redundancy offered "the best of both worlds" replicating Helleberg and Wickens (2002) In fact, across measures, performance with the redundant display was statistically equal to or worse than that of the single modality displays.

With regard to the traffic load manipulation, we have seen that higher traffic load expectedly induced more scanning to traffic-relevant sources of information (the OW and CDTI; Wickens, Xu, Helleberg, & Marsh, 2001); as a consequence, IP scanning diminished and flight control degraded accordingly. However, the reallocation of more visual attention outside did not appear to compensate for the higher traffic load, as traffic detection performance deteriorated despite the greater "eyes-out" time. When this finding is viewed in the context of the pilot's "ANCS" task hierarchy (Schutte & Trujillo, 1996; Wickens, Xu, Helleberg, & Marsh, 2001), the strategy adopted at high workload may be "reasonably optimal." That is, while the pilots did not entirely "protect" the aviate subtask from the

demands of higher traffic workload, neither did they overly sacrifice the aviate task, and they therefore adhered to prescriptions to maintain "aviating" as the highest priority task.

Communication Phase

The auditory data link display, fared better than in the prior evaluation (Helleberg & Wickens, 2002). In the current study, the messages were short, and note taking was not allowed (a source of visual competition in the prior study). Thus pilots could keep eyes out (using horizon information for altitude control), and increase scanning to the instrument panel, to better coordinate their maneuvers. This is consistent with multiple resource theory, regarding the advantages of offloading visual information to auditory channels (Wickens, 2002). However, working memory limitations were evident. Short message comprehension (1 chunk of information) did not suffer from auditory delivery, but with longer messages (3 chunks), readback accuracy dropped from 98% to around 89%. Again, the redundant condition failed to offer "the best of both worlds." The availability of the visual data link did mitigate the working memory limitations. However, the auditory channel did not reduce the flight path tracking error, relative to the auditory-only condition; nor did it reduce the amount of data link display scanning, relative to the visual only display. Pilots appear "compelled" to cross check the visual display with the auditory message, a finding consistent with that reported in the prior study by Helleberg and Wickens (2002).

General Discussion of Modality Effects

The change from auditory to visual appears to be a mixed blessing. The added head down time associated with the two visual displays imposed a cost to flight path tracking that was likely attributable to the loss of the outside visual horizon. The detection of cued traffic was only slightly hindered by the added head down time. However, this added head down time did hinder the detection of the "rogue" or transponder-off aircraft. From an attentional perspective, these results are consistent with a multiple-resource model (Wickens, 2002), that posits separate resources for auditory and visual processing, and therefore accounts for the advantage of offloading some task demands to the underutilized auditory channel, thereby availing more resources for the high demand visual tasks of flight control (aviating) and traffic awareness (navigating). There was no evidence of an auditory preemption effect. In fact the auditory-only condition was the least disruptive of ongoing visual tasks. This may be due to the somewhat predictable length and timing of the ATC messages. Redundantly offering information in

both modalities failed to provide better performance than that of the best of the single modality channels. This suggests that explicit training may be required in order to realize a redundancy gain.

The results have also shown that cueing or attention guidance, provided in either modality, appears to maintain detection performance at a level of around 90%, clearly better than the un-cued rate of around 60% revealed in other studies (Wickens, Helleberg, & Xu, 2002; Marshall & Fisher, 1959). However, this detection level still remains far from perfect, and appears to bring with it some attentional tunneling cost. Finally, we note that the overall ratio of OW to inside scanning remains, as observed in prior research (Wickens et al., 2002), as one that departs from FAA guidelines to favor the OW.

Acknowledgements

This research was supported by grant NASA NAG-2-1120 from NASA Ames Research Center. Dr David Foyle was the scientific technical monitor. The authors wish to thank Roger Marsh, Jonathan Sivier, and Ron Carbonari for their programming assistance.

References

Helleberg, J. & Wickens, C. D. (2002, in press). Effects of data link modality and display redundancy on pilot performance: An attentional perspective. The International Journal of Aviation Psychology.

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.

Latorella, K. (1998). Effects of modality on interrupted flight deck performance. Proceedings of the 42nd Annual Meeting of the HFES. Santa Monica, CA: Human Factors & Ergonomics Society.

Marshall, P. R., & Fisher, R. B. (1959). Determination of the daytime conspicuity of a small airplane in a terminal area (Technical Development Report 382). Indianapolis, IN: Federal Aviation Agency.

Navarro, C., & Sikorski, S. (1999). Datalink communication in flight deck operations: A synthesis of recent studies. The International Journal of Aviation Psychology, 9(4), 361-376.

Prinzo, O. V. (2001). Pilot visual acquisition of traffic: Operational communications from OpEval-1

(DOT/FAA/AM-01/9). Washington, DC: Federal Aviation Administration, Office of Aviation Medicine.

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

U.S. Department of Transportation (2001). Federal aviation regulations (91.113). Washington, DC: Federal Aviation Administration.

Vidulich, M. A., & Wickens, C. D. (1986). Causes of dissociation between subjective workload measures and performance. Applied Ergonomics, 17, 291-296.

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, in press). Multiple resources and performance prediction. Theoretical Issues in Ergonomic Science, 3(2) 159-177.

Wickens, C. D., Goh, J., Helleberg, J., & Talleur, D. (2001). Modality differences in advanced cockpit displays: Comparing auditory vision and redundancy for navigation communication and traffic awareness. (Technical Report ARL-02-8/NASA-02-6). Savoy, IL: University of Illinois, Aviation Research Lab.

Wickens, C. D., Helleberg, J., Goh, J., Xu, X., & Horrey, B. (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., & 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.

Wickens, C. D., Xu, X., Helleberg, J., & Marsh, R. (2001). Pilot visual workload and task management in freeflight: A model of visual scanning. Proceedings of the 11th International Symposium on Aviation Psychology. Columbus, OH: Dept. of Aerospace Engineering, Applied Mechanics, and Aviation, Ohio State University.