EMBRY-RIDDLE AERONAUTICAL UNIVERSITY

ERAU ATM RESEARCH LABORATORY

Annex D to Experimental Plan for Assessment and Measurement of Some Human Factors Issues in the Control of Offshore Helicopters Using Automated Dependent Surveillance

The Effect on Air Traffic Controller Performance of Different Update Rates with Automated Dependent Surveillance

Author(s):	Ian Wilson, Micah Morris
Date:	17 February 2004
Edition:	0.D
Status:	Final Report
Reference:	FAA/ADS/Helisim/Experiment2

Revision History

Date	Revision Number	Reason for Revision	Affected Pages
25 Feb 2003	0.A	Initial Document	Document
28 Feb 2003	0.B	Format and Walkthrough Update	Document
02 Feb 2004	0.C	Editorial Update	Document
17 Feb 2004	0.D	Final Report	Document

Document Approval

	Name	Signature	Date
Author 1	Ian Wilson		
Author 2	Micah Morris		
QA Check	Kenneth Fleming		
Director	Kenneth Fleming		

TABLE OF CONTENTS

EXECUTIVE SUMMARY

Embry-Riddle Aeronautical University's ATM Research Laboratory (EARL), under an FAA Research Grant has been carrying out research into the Human Factors aspects of the use of Automated Dependent Surveillance-Broadcast (ADS-B). This paper describes the second in a series of experiments that were carried out at the university to investigate these issues. The first series of experiments are described in a companion paper (see Annex C).

There are more than 600 helicopters operating in the Gulf of Mexico in good weather. Each of these aircraft may fly more than 10 flights per day. However, when the weather is poor, the number of flights is reduced to less than 2% of the good weather figure. This is because there is a limited, procedural air traffic control service during such circumstances. The disruption and delays caused by these factors has been estimated to cost the oil industry and the helicopter operators $300,000 per hour.

In theory, the provision of ADS-B surveillance and DSTs such as a URET should enable radar-like control. However, there are many Human Factors issues that need to be assessed. The procedural control system is self-limiting, and while it is restrictive for helicopter operations, it protects the controllers from overload. A radar-like service can provide far more efficient flight paths for the aircraft, but the controllers could quickly become overloaded.

In addition to the change of control paradigm, there are differences in accuracy between ADS-B and radar that could cause misidentification leading to deconfliction problems for controllers. This could also lead to problems at the sector boundaries. One way of avoiding this is to declare certain sectors as ADS-B sectors and others as radar sectors. Another method consists of mixing the two surveillance systems in the same sectors.

Many of the helicopters cannot be equipped with ADS equipment due to space, weight, and power restrictions. Therefore, there will always be a mixed equipage problem for the controllers so that some aircraft may remain under procedural control.

The present study investigates the ability of the air traffic controller to detect and resolve conflicts with varying ADS-B position update rates as displayed on their ATM display. The controller participants for this study were upper-level air traffic control students at Embry Riddle Aeronautical University. In the experiments they were evaluated on their ability to detect and resolve conflict situations using the simulated ADS-B position updates.

The results of the study indicate that there is a significant effect for update rates versus the time to conflict detection. However, this effect is most pronounced between a 180 second update rate and a 60 second update rate. When the update rate goes below 60 seconds, there is a pronounced levelling effect that may require further investigation.

1. BACKGROUND

1.1 As part of the EARL Research grant from the FAA, research is being conducted into the use of ADS-B by controllers for helicopter traffic in the Gulf of Mexico. As there is almost no radar coverage of the helicopters operating off-shore, the IFR service provided when conditions are IMC is limited to a strict procedural system. This system has been enhanced in recent years by the use of a grid of reporting points over the ocean, but it is still extremely restrictive. During normal VFR operations, as many as 6,000 flights per day can take place. When IFR procedural rules are in place, this figure drops to around 100 flights per day. Effectively, the 600 or more helicopters servicing the oil rigs and platforms in the Gulf are grounded when the weather conditions become IMC.

1.2 To assess the problem, visits were made to Houston Center and New Orleans TRACON, and a meeting of the Helicopter Safety Advisory Committee (HSAC) was attended. At Houston Center there is very little helicopter traffic when the weather is good, and the oceanic control sector provides a flight following service or procedural control service to the few helicopters that call as part of the overall oceanic control task. When the weather deteriorates and the helicopters are forced to fly IFR, up to 2 separate sectors may be manned to provide procedural control to the helicopter traffic. Similarly at New Orleans TRACON, a separate control position provides service to helicopters under certain IFR conditions.

1.3 The concerns over the provision of some services in the Gulf of Mexico have lead to attempts to improve the service. As stated above, the procedural system has been improved, and there have been trials for several years based on new avionics that are now available such as ADS. ADS has been used in Alaska, albeit in a far less intensely flown area, to resolve the problems of lack of radar cover and to provide the aircraft with Cockpit Display of Traffic Information (CDTI). There is some reluctance from the helicopter operators to try to equip small helicopters with avionics. This is due not only to the costs associated with them, but also the lack of space and power supplies in the cockpits of the aircraft. Therefore, a solid cost-benefit case would need to be made for the operators to equip their aircraft. The case would have to be made both to large international helicopter operators and to the small independents. Even with a strong business case, it is unlikely that all aircraft would equip these avionics. Therefore, any future system would need to accommodate mixed equipage.

2. INTRODUCTION

2.1 At present, the local air traffic services within the Gulf of Mexico are provided by several air traffic control units. These vary from the large, well-equipped TRACON at New Orleans through the smaller radar equipped airports like Houma to busy non-radar units such as Patterson and to the large number of landing sites with almost no facilities. Houston Center provides the air traffic service for the majority of the area of the Gulf of Mexico and has a dedicated oceanic sector, as well as a single helicopter oceanic sector in cases of sufficient demand for IFR service. The Houston sectors operate with their DSR displays set to show their area of responsibility that stretches from the Mexican airspace boundary North to the Gulf shore and from Brownsville in Southern Texas, east to the Mississippi delta. The 600 aircraft in the helicopter fleet carry out their 4,000-6,000 flights per day within the area covered by this display.

2.2 This paper describes an experiment that explores the issue of different ADS-B position update rates that can potentially be used on the ATM display covering the Gulf of Mexico. Although ADS has the capability to provide very high aircraft position update rates, they are more costly than lower rates. This is of course due to the costs associated with each update transmission. Therefore, a benefit case can be made if controllers are able to detect helicopter conflicts as well when the ADS-B update rate is decreased as they are when the update rate is high, given a constant number of aircraft within the display.

2.3 It is quite possible, given the fact that helicopters fly at airspeeds that are much slower than fixed-wing aircraft, that the ADS-B position update rate could be reduced to as much as once every three minutes, or .33 hertz. The experiment described below explores this idea by measuring the performance of the controllers as they are presented with varying levels of update rates. More specifically, the rates selected are an update once every 6, 20, 60, and 180 seconds.

3. STATEMENT OF THE PROBLEM

3.1 The issues raised by the requirement to provide control of helicopter traffic in the Gulf of Mexico that pertain to this study are as follows:

a There are more than 600 helicopters operating offshore in the Gulf with up to 6,000 flights per day (many of which are short platform-to-platform flights).

b Of the 600 or more airframes, only approximately 100 are equipped with avionics to allow IFR flight, although this number can be expected to rise to 200 if there was a good cost benefit case made to the operators.

c The ability of the controller to effectively and efficiently detect helicopter conflicts could be similar when the ADS-B position update rate on the ATM display is reduced, thereby making a benefit case for the lower rate

4. REVIEW OF RELATED LITERATURE

4.1 Situational Awareness is a phenomenon that has been under great scrutiny within the field of Human Factors within the past few decades. It refers to the representation that humans have of their environment (Fracker, 1988). As the success of air traffic management is dependent on the controller's ability to appropriately and concisely detect any potential conflict situations, it is very important that the factors contributing to their situational awareness are examined closely. Improved safety and decision making, in addition to reduced workload, are the results of increased situational awareness, (Regal, Rogers, & Boucek, 1988; Sarter & Woods, 1991).

4.2 Reports from the Aviation Safety Reporting system identified the circumstances under which most situational awareness errors occurred. The report found that 76.3% of errors were due to the failure to correctly perceive the information, 20.3% were due to the failure to comprehend the situation, and 3.4% were due to the failure to project the situation in the future. Most of these errors were found to be due to the failure to perceive the data, having an incorrect mental model, and in many cases, memory loss.

4.3 Applying these findings to the realm of air traffic management, it seems that it is critical that displays provide enough information to the controller so that the controller is able to interpret the situation, and then correctly project into the future what will occur, particularly with respect to conflicts. Therefore, displays and/or technologies that improve this predictive ability are very important.

4.4 One of the major advantages of ADS-B is its ability to update the location of aircraft at a predetermined rate. It is possible, for instance, to update the location of aircraft every fraction of a second. The drawback to this is in the fact that it costs more to do this than if the update rate were decreased. It might be possible, given the fact that helicopters fly at much slower flight speeds than fixed wing aircraft, to establish a less expensive and lower update rate for this particular ADS-B application.

4.5 As conventional radar has been the most common form of ATM technology studied in the past, it is interesting to note that controllers must often perform conflict detection under varying update rates. In most cases radar is updated every six seconds. However, some times the radar beacon can spin anywhere from five to fifteen revolutions per minute. This would equate to an update rate of anywhere from four to twelve seconds. Other research involving vigilance tasks in ATM have incorporated update rates of 30 times per minute, or every two seconds (Hitchcock et al., 1999).

4.6 The major area of focus of this research involves the determination of whether differences in air traffic controller conflict detection capabilities exist between different levels of aircraft position update rate. If this is true, then to the it may be possible to reduce ADS-B transmissions under the particular conditions outlined in this research.

5. METHODS

Participants

5.1 The current study consisted of a sample of upper level air traffic management students from Embry-Riddle Aeronautical University. The criterion for participation included the completion of at least two courses in air traffic control incorporating practice using conventional radar displays with some exposure/knowledge of ADS technology. A total of 6 participants were observed during this experiment.

Simulations and Equipment

5.2 The ability of the air traffic controller to effectively maintain safety separation for the number of aircraft within a sector was tested using various ADS update rates. The performance of the controller was measured against the number of conflicts the controller was effectively able to identify. In order to conduct this research, an ATC simulator with software capabilities to vary the Update Rate of traffic within the sector was required. The simulator recorded the time at which the controller identified conflicts that had been designed into the simulations prior to testing.

6. DESIGN

Experimental Design

6.1 The dependent variable under study was the controller's ability to identify aircraft that were in conflict. Traffic was presented at varying Update Rates to investigate how the controller responded to different update rates. The four UR conditions consisted of 6, 20, 60, and 180 seconds.

6.2 In addition to the UR independent variable, the Time to Loss of Separation (TLS) was also manipulated. As in the first experiment, described in a companion paper (need reference here), TLS levels of 10 minute, 12.5 minute, and 15 minute were incorporated.

6.3 Each simulation lasted 20 minutes. Six conflicts occurred in random locations on the ATM monitor during each condition. As there were three levels of TLS incorporated into each simulation, there were two conflicts per level of TLS per simulation. Background aircraft that were not in conflict were randomly scattered within the area being observed.

6.4 Several aspects of this study were controlled. The various conflicts included two different conflict angles to minimize any possible learning effects from a constant conflict angle. Based on the results of the first experiment, that involved the manipulation of various traffic densities in a similar ADS environment, a traffic density of 125 nm range was incorporated into each simulation. Aircraft were randomly scattered throughout the entire simulation screen and six conflicts per simulation were involved. While the time of occurrence of the six conflicts varied within each simulation, the time of the occurrence did not vary between simulations. This again minimized the learning effect for individual subjects, while maintaining a constant situation across subjects.

6.5 Objective measures of Conflict Detection (CD) were taken. In addition, subjective workload measures were recorded by the administration of the NASA-TLX.

6.6 The participants were studied in a completely within-subjects format. Therefore, all participants were tested in each experimental condition. In order to reduce learning effects still further, the traffic patterns varied randomly for each condition. Finally, the administration of the various simulations was randomized for each participant.

7. PROCEDURE

7.1 After completing a short informed consent form, participants were briefed as to their objectives during the experiment.

7.2 Participants were next asked to complete a 20-minute practice trial. The practice trial required the controller to detect conflicts that were occurring at an ADS UR of 6 seconds. The purpose of this practice trial was to familiarize the participants with the display, so that they received sufficient practice.

7.3 Following the practice trial, a 10-minute break was provided, after which the four experimental trials were administered. Each experimental trial lasted 20 minutes. The NASA-TLX was administered following each experimental trial. There was a 10-minute break offered after every experimental trial.

7.4 At the conclusion of the final experimental trial, a debriefing session was provided. The debriefing session provided an explanation to the participant as to the purpose of their participation, as well as time for questions to be answered.

7.5 A schedule for the experiment is as follows for each participant:

BRIEF	08:00-08:10
PRACTICE	08:10-08:30
BREAK	08:30-08:40
EXP 1 RUN	08:40-09:00
NASA-TLX	09:00-09:05
BREAK	09:05-09:15
EXP 2 RUN	09:15-09:35
NASA-TLX	09:35-09:40
BREAK	09:40-09:50
EXP 3 RUN	09:50-10:10
NASA-TLX	10:10-10:15
BREAK	10:15-10:20
EXP 4 RUN	10:20-10:40
NASA-TLX	10:40-10:45
DEBRIEF	10:45-10:55

8. METRICS

Objective Metrics

8.1 Participants were required to detect conflicts in varying levels of ADS-B UR. In order to measure this, the time was recorded when the controller selected the aircraft that were in conflict.

METHODOLOGIES – Objective Metrics

8.2 Conflict detection was recorded when the controller clicked on the two aircraft in conflict. The amount of time taken by the controller to detect the conflict situation was measured from the time that the aircraft began a heading that would result in the conflict a to the time that the controller clicked on the two aircraft.

Subjective Metrics

8.3 In order to gain a subjective workload assessment of the controller participants, the NASA-TLX was administered following every experimental simulation.

METHODOLOGIES – Subjective Metrics

8.4 The NASA-TLX asks participants to rate demand levels on a scale ranging from LOW to HIGH. The demand domains consist of MENTAL, PHYSICAL, TEMPORAL, PERFORMANCE, EFFORT, and FRUSTRATION. Pair-wise comparisons of these demand domains are then made by the participant, during which time the demands that provided the most significant sources of variation during the tasks at hand are circled.

9. HYPOTHESES/INTENDED RESULTS

Hypothesis 1:

9.1 It was expected that there would be a difference in conflict detection times so that greater levels of UR would result in lower conflict detection times. Thus, a UR of 6 seconds would result in a lower conflict detection time than that of 180 seconds.

Hypothesis 2:

9.2 It was expected that lower times to loss of separation would result in shorter conflict detection times.

Hypothesis 3:

9.3 An interaction effect for the UR and TLS variables was expected, so that the level of UR which would result in lower conflict detection times would be dependent on the level of TLS incorporated.

10. ANALYSIS OF RESULTS

9.1 Update Rate

A factorial repeated-measures analysis of variance (ANOVA) was conducted for the ADS position Update Rate variable, the Time to Loss of Separation variable, and the interaction of these two variables. The following tables provide a summary of the analysis.

Descriptive Statistics

	Mean	Std. Deviation	N
Update Rate 6 seconds	288.17	131.218	6
Update Rate 20 seconds	315.17	94.294	6
Update Rate 60 seconds	221.67	201.210	6
Update Rate 180 seconds	664.33	91.914	6

Tests of Within-Subjects Effects

Dependent Variable: Conflict Detection Time (secs.)

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	1104814.226(a)	9	122757.136	4.539	<.001
ADS Update Rate	838634.336	3	279544.779	10.336	<.001
Time to Loss of Sep.	402176.045	2	201088.023	7.435	.001
ADS UR * TLS	101200.575	6	25300.144	.935	.448
Error	2028416.762	12	27045.557		
Total	19425958.000	23			
Corrected Total	3133230.988	32			

a R Squared = .353 (Adjusted R Squared = .275)

Multiple Comparisons

Dependent Variable: Conflict Detection Time (secs.)
Tukey HSD

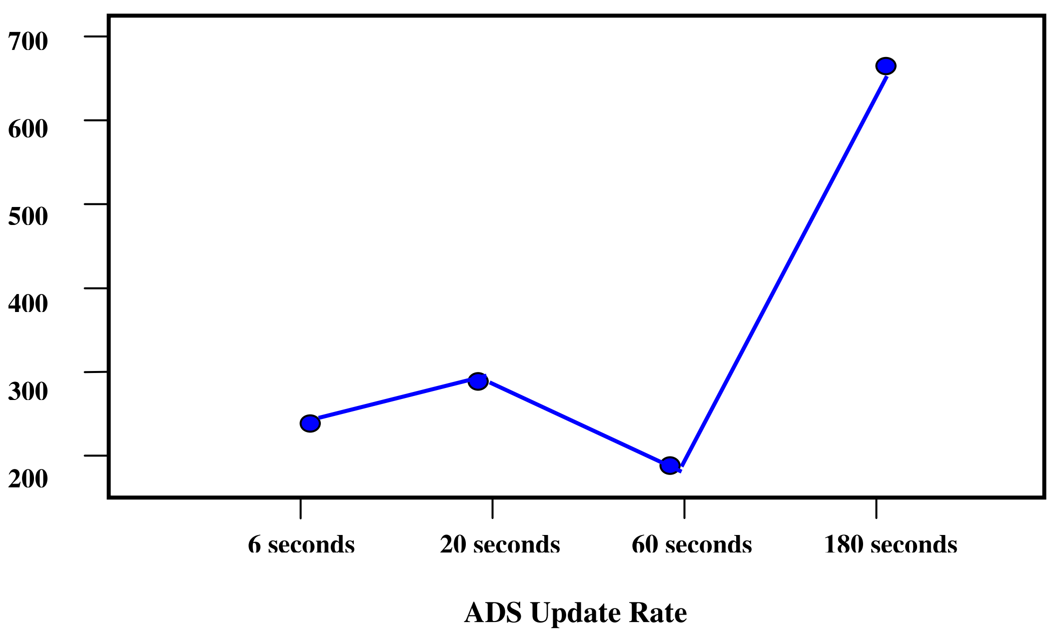
(I) ADS Update Rate	(J) ADS Update Rate	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
20 sec	6 sec	27	46.497	1.000	-120.14	124.21
60 sec	6 sec	66.5	46.854	.054	-1.41	244.81
60 sec	20 sec	93.5(*)	43.572	.037	5.18	234.16
180 sec	6 sec	376.16(*)	75.743	.002	88.59	486.63
180 sec	20 sec	349.16(*)	73.758	.001	91.77	479.38
180 sec	60 sec	442.66	73.983	.121	-28.49	360.30

Based on observed means.
* The mean difference is significant at the .10 level.

The tables above display some interesting results. There are significant differences between the levels of the independent variables and value of the dependent variable, and these are in the directions that were hypothesized. That is, there is a significant main effect between the ADS update rate and the time to conflict detection and the same thing is true of the time to loss of separation and the time to conflict detection. Both of these variables are significantly different at the .001 level or greater. This means that we can be over 99 percent sure that these results are not due to chance. However, further analysis of the data in the pair wise comparisons indicate some counterintuitive results. In particular, the pair wise comparisons for the ADS update rates are quite interesting. The 60 second rate provides the lowest mean time for conflict detection, but it also has a very high standard deviation. The 6 and 20 second update rates have higher means but significantly lower standard deviations. The pair wise test also shows that there is no significant difference between the 6 and 20 second update rates. What appears to be happening is similar to that observed in the first experiment. That is, there seems to be a natural limit to the effectiveness of the update rate. It is clearly significant from 180 seconds down to the 20 second update interval but perhaps no further. There may be an optimal update rate that is centered around the 60 second interval, but the high standard deviation for this particular update rate may also be an indication of high between subjects variation. Both of these hypotheses should be studied with further experiments.

A visual picture of the time to conflict detection versus the update rate gives a good indication of the discussion above. Therefore, The following plot displays the means for the conflict detection time across the four levels of ADS Update Rate.

Estimated Marginal Means of Conflict Detection Times Across the Four Levels of ADS Update Rate



As the graph clearly shows, the differences between the 6, 20, and 60 second update rates are insignificant when compared with that of the 180 second update rate.

10.2 Time to Loss of Separation

The findings for the time to loss of separation variable were similar to those in experiment 1. That is, there appears to be a significant effect in the mid-range of the comparison between the two variables, while the extremes indicate some sort of a leveling effect as far as the differences between the variables are concerned.

The following table provides a summary of the differences between the levels of the Time to Loss of Separation Factor.

Multiple Comparisons

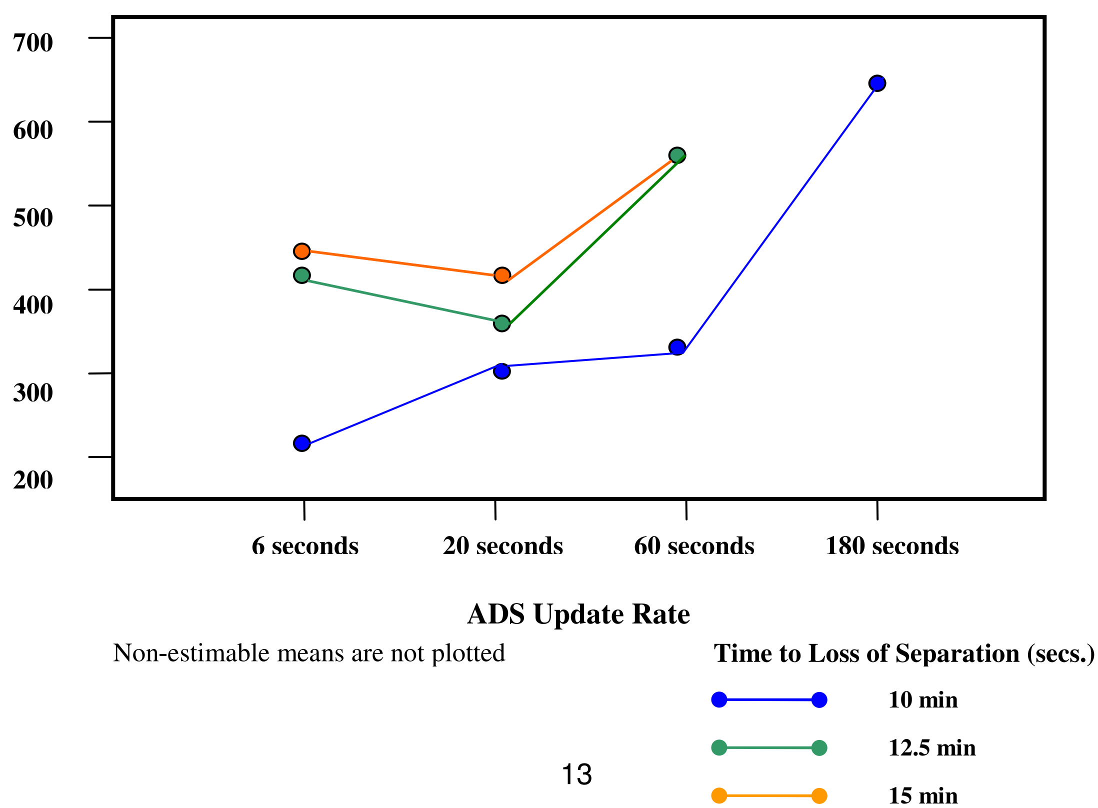
Dependent Variable: Conflict Detection Time (secs.)
Tukey HSD

(I) Time to Loss of Separation (secs.)	(J) Time to Loss of Separation (secs.)	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
12.5 min	10 min	66.75	43.291	.277	-36.77	170.26
15 min	10 min	101.75	43.291	.055	-1.77	205.26
15 min	12.5 min	35.00	44.759	.715	-72.02	142.02

Based on observed means.

The following plot displays the means for the conflict detection time for the three levels of TLS across the four levels of UR.

Estimated Marginal Means of Conflict Detection Times (seconds)



10.3 Workload

A one-way analysis of variance was conducted on the NASA-TLX workload measures collected from four of the participants during the experiment. The test first asks participants to place a marking on a low-to-high scale line that represents the magnitude of that particular factor for the task that was just performed. This is done for the factors of Mental Demand, Physical Demand, Temporal Demand, Performance, Effort, and Frustration. These factors are operationally defined on a separate sheet. The participants are then asked to fill out a pair wise comparison sheet, which lists the previously mentioned factors side by side, and the participant is required to circle the member of each pair that provided the most significant source of variation in the task just performed. The responses are then entered into a computer software program, which generates a raw rating and a weight assigned to that rating for each factor. The raw rating and the weight are then multiplied, and an adjusted rating is computed. The actual workload rating for that individual during the task is the sum of the adjusted ratings.

The following tables provide a summary of the analysis of participants' workload ratings across the four levels of Update Rate. No significant differences were found at the .10 alpha level between the levels of the variables studied for the subjective workload measures taken.

Descriptives

Workload

					95% Confidence Interval for Mean			
	N	Mean	Std. Deviation	Std. Error	Lower Bound	Upper Bound	Minimum	Maximum
6 seconds	4	2.61825	2.112296	1.056148	-.74288	5.97938	.845	5.027
20 seconds	4	1.12825	.325876	.162938	.60971	1.64679	.890	1.600
60 seconds	4	.96100	.399354	.199677	.32554	1.59646	.444	1.355
180 seconds	4	2.01100	2.112552	1.056276	-1.35054	5.37254	.533	5.110
Total	16	1.67963	1.523507	.380877	.86781	2.49144	.444	5.110

ANOVA

Workload

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	7.245	3	2.415	1.051	.406
Within Groups	27.571	12	2.298		
Total	34.816	15			

11. CONCLUSION

11.1 This experiment was designed to determine the impact of differing levels of ADS position Update Rate and Times to Loss of Separation, on the efficiency of the controller's conflict detection. The experiment preceding this one examined differences in controller performance when various levels of traffic density were presented to controllers with different levels of time to loss of separation.

11.2 Based on the results of this experiment, there was a tendency for increased conflict detection time as a function of decreased Update Rate, but the pair wise comparison indicated that this effect might level off at some point around or below the 60 second update rate.

11.3 The significant finding that lower TLS results in lower conflict detection time is consistent with the results of the first experiment.

11.4 The interaction effect was not significant. It is possible that this finding could be attributed to the relatively low sample used. Regardless, this experiment determined that the level of UR at which conflict detection is optimized is not dependent on the level of TLS involved.

11.5 The lack of significance for the subjective workload measures gathered from the NASA-TLX across the four levels of Update Rate are also surprising, but nonetheless consistent with the first experiment,

11.6 Therefore, it appears that there is at least the potential for decreases in UR that would still allow the controller to effectively detect conflict situations. It must be noted, however, that this is in an environment that involves slowly moving helicopters flying in a low density environment (18 aircraft over125 nm).

ANNEX D. A – ACRONYMS AND ABBREVIATIONS

AAR	Airport Acceptance Rate
ADS-B	Automated Dependent Surveillance – Broadcast
ANOVA	Analysis of Variance
ATC	Air Traffic Control
ARTCC	Air Route Traffic Control Center
ATD	Added Time to Destination
CD	Conflict Detection
CDTI	Cockpit Display of Traffic Information
CI	Confidence Interval
CM	Communication Mode
CPDLC	Controller-Pilot Data Link Communication
CQ	Conflict Quantity
CR	Conflict Resolution
CRE	Conflict Resolution Efficiency
DST	Decision Support Tool
EARL	ERAU ATM Research Laboratories
ERAU	Embry-Riddle Aeronautical University
FAA	Federal Aviation Administration
FPD	Flight Path Deviation
HF	Human Factors
HSAC	Helicopter Safety Advisory Committee
HSD	Honestly Significant Difference
IFR	Instrument Flight Rules
MTCD	Medium-Term Conflict Detection
RI	Radar Inaccuracy
TRACON	Terminal Radar Approach Control
TV	Traffic Volume
URET	User Request Evaluation Tool
VFR	Visual Flight Rules

ANNEX D. B – REFERENCES

[Reference 1.] Fracker, M. L.; 1988; A theory of situation assessment: Implications for measuring situation awareness; In Proceedings of the 32nd Annual Meeting of the Human Factors Society; Santa Monica, CA: Human Factors Society; 102-106.

[Reference 2.] Hitchcock, E. M., Dember, W. N., Warm, J. S., Moroney, B. W., & See, J. E.; 1999; Effects of cueing and knowledge of results on workload and boredom in sustained attention; Human Factors; Vol. 41(3); 365

[Reference 3.] Regal, D., Rogers, W., & Boucek, G.; 1988; Situational awareness in the commercial flight deck: Definition, measurement, and enhancement (SAE Technical Paper Series 881508).

[Reference 4.] Sarter, N. B., & Woods, D. D.; 1991; Situational awareness: A critical but ill-defined phenomena. The International Journal of Aviation Psychology, 1, 45-57.