Using the Solution Space Diagram in Measuring the Effect of Sector Complexity During Merging Scenarios

S. M. B. Abdul Rahman^{*}, M.M. van Paassen[†], M. Mulder[‡]
Delft University of Technology, Delft, The Netherlands

When designing Air Traffic Control (ATC) sectors and procedures, traffic complexity and workload are important issues. For predicting ATC workload, metrics based on the Solution Space Diagram (SSD) have been proposed. This paper studies the effect of sector design on workload and SSD metrics. When considering the SSD in evaluation of a sector, each aircraft within the sector introduces a zone of conflict, the Forbidden Beam Zone (FBZ), on the SSD. The properties of these FBZ are systematically studied to increase understanding of the SSD usability in assessing workload and sector complexity. The effects of sector design variables on Air Traffic Controller (ATCo) workload and also SSD properties were evaluated. Example of sector properties are, number of streams to be merged, the merge angle, the proximity of incoming aircraft and the variability of traffic mix of small and large aircraft. Based on the findings, each sector design variable leads to different effect on both workload and SSD properties. Apart from that, correlation between the workload and the SSD properties were found to be in a higher level than of the number of aircraft within the sector, proving that the SSD-based analysis to be a good workload indicator. These correlations were studied based on two different groups of subjects with ranging experience in order to demonstrate the robustness of the method.

Nomenclature

A_{total}	=	Total area
A_{mean}	=	Mean area
n_{Sect}	=	Number of aircraft in the sector
n_{All}	=	Number of aircraft showed to the controller
$N_{Command}$	=	Number of command
$N_{Conflict}$	=	Number of conflict

I. Introduction

The growth of air traffic is limited by various factors. The Air Traffic Controller (ATCo) Workload is an important one.¹⁻³ Many studies have been performed to understand and distinguish between ATCo taskload, workload and sector complexity, and their effects on the ATCo performance. Hilburn and Jorna,⁴ in an effort to distinguish between taskload and workload stated that system factors such as airspace demands, interface demands and other task demands contribute to taskload, while operator factors like skill, strategy, experience and so on determine workload. Stassen, et al.⁵ have also discussed the taskload by using the term task demand load in considering an objective measure of the difficulty of the task to be performed by the controller. To accommodate a further growth of air traffic, a better understanding of what factors contribute to ATCo mental workload is needed. Thus, when a better understanding on how certain design properties affect workload is acquired, design can be improved from the workload perspective.

The mental workload can be assessed using few methods such as performance-based workload assessment

^{*} PhD Student, Control and Simulation Division, Faculty of Aerospace Engineering, Delft University of Technology. Kluyverweg 1, 2629 HS Delft, The Netherlands. E-mail: s.m.b.abdulrahman@tudelft.nl

[†] Associate professor, Control and Simulation Division, Faculty of Aerospace Engineering, Delft University of Technology. Kluyverweg 1, 2629 HS Delft, The Netherlands. Member AIAA. E-mail: m.m.vanpaassen@tudelft.nl

[‡] Professor, Control and Simulation Division, Faculty of Aerospace Engineering, Delft University of Technology. Kluyverweg 1, 2629 HS Delft, The Netherlands. Senior member AIAA. E-mail: m.mulder@tudelft.nl

through primary and secondary task performance, subjective workload assessment using continuous and discontinuous workload ratings and also using physiological measures of the mental workload. In order to understand workload in relation to sector design a method to quantify ATCo's workload is needed. One example is by using the sector complexity as an objective measurement indicator.

Aircraft count or aircraft density is one method of determining sector complexity by looking at the number of aircraft within a sector. It is defined as the number of aircraft per unit of sector volume. Experiments showed that of all the individual sector characteristics, aircraft density shows the highest correlation with ATCo subjective workload ratings.^{6,7} However, aircraft density has significant shortcomings in its ability to accurately measure and predict sector level complexity^{6,8} where it is unable to capture the dynamic behavior of aircraft in the sector. For example, five aircraft flying in the same direction do not exhibit the same complexity as the same number of aircraft flying in various directions.

Another measurement of sector complexity is dynamic density, which incorporates the dynamic behavior of aircraft in the sector. Research on dynamic density by Laudeman et al.⁹ and Sridhar et al.¹⁰ has indicated a number of variables for calculating the dynamic density and each factor is given a subjective weight. The dynamic density is a summation of these variables and its corresponding subjective weight. However, the summation of dynamic density that is based on subjective weights gathered from regression methods on samples of traffic data means that the method only be used on scenarios that do not deviate too much from the baseline scenario. Therefore the dynamic density metric is not generally applicable to just any situation.¹¹

In order to better assess airspace complexity, new methods such as 'complexity maps' and the 'solution space' are proposed by Lee et al.¹² and Hermes et al.¹¹, respectively. Both solutions act as an airspace complexity measure method, where a complexity map details the operator's control activity as a function of the parameters describing the disturbances, and the solution space details those two-dimensional speed and heading possibilities of one controlled aircraft that will not induce separation violation.

A. Sector Complexity and Workload

The sector complexity construct is described based on the important factors associated with a sector's structural and flow characteristics. The structural characteristics are fixed for a sector such as terrain, number of airways, airway crossings and navigation aids. The flow characteristics, however, vary as a function of time and depend on characteristics like number of aircraft, mix of aircraft, weather, separations between aircraft, aircraft speeds and flow restrictions. A combination of these structural and flow parameters influences the airspace demand, thus the controller's workload.¹⁰

Figure 1 shows the relationship between taskload and workload as described by Hilburn and Jorna⁴, where we adapted the position of sector complexity within the diagram. The function of the Solution Space Diagram (SSD) is included as a workload measure and alleviator, possibly also affecting the sector planning.

Given that sector complexity having relation with the taskload and thus workload as experienced by the controller, a sound research has to be performed in order to have a better understanding of the sector complexity measure and its effect to workload, performance and safety. Majumdar and Ochieng¹³, in their research has defined a total of over 50 complexity-shaping factors (CSF). Having said that there are numerous sector complexity variables that can influence controller's workload, a good sector design has to be achieved. In order

accomplish that, the impact of the Air Traffic Control (ATC) complexity variable on the controller's workload has to be assessed systematically. Several methods and metrics have been proposed to quantify the complexity of a sector, and among all is the SSD. The SSD has been shown to be useful as a good measure for workload, and the presentation of the diagram to ATCos even alleviated them in previous researches by D'Engelbronner et al.¹⁴ and Mercado-Velasco et al.¹⁵, respectively. In this paper, the investigations of sector design variables were included.

Figure 1: Taskload and Workload (adapted from Hilburn and Jorna, 2001).

B. The Solution Space Analysis Method

The SSD is used as a basis for obtaining a better metric for complexity, capable of predicting the task demand load of air traffic controllers. In previous research, the solution space-based metric proved to be a more objective and scenario-independent metric than a weighted combination of scenario properties, in which the weights are highly dependent on the baseline scenarios considered.^{11,14,15}

Initial work by Van Dam et al.¹⁷ has introduced the application of Solution Space in aircraft separation problems from the pilot's perspective. Hermes et al.¹¹, D'Engelbronner et al.¹⁴ and Mercado-Velasco et al.¹⁵ have continued the idea of using solution space in aircraft separation problem but for the Air Traffic Control (ATC) problem. Based on research conducted by Hermes et al.¹¹ and D'Engelbronner et al.¹⁴, a high correlation was shown to exist between the Solution Space and ATCo's workload. Mercado-Velasco et al.¹⁵ study the workload from a different perspective, looking at possibility of using the SSD as an interface to reduce controller's workload. Based on his studies,¹⁵ he indicates that the diagram could indeed reduce controller's workload in a situation of increased traffic level. Results also show the possibility of the existence of a traffic threshold level, up to which the SSD interface is effective in reducing workload.

This paper follows the proposed method by Hermes et al.¹¹, and adds an exploration of the effect of sector design variables on the controller's workload and the metrics based on the SSD. To this end, an experiment was conducted in which the traffic pattern was consistently varied. The SSD was used as an offline evaluation method of sector complexity and workload, and metrics based on the SSD were compared to workload ratings given in the experiment. Other performance related measurement such as the number of commands, number of conflicts, extra distance ratio, smallest aircraft separation et cetera are also evaluated in order to see the correlation between the performances of the subjects to the workload indicated. A questionnaire regarding the controller perceived workload and also comments regarding the experiment were also taken at the end of every simulation run. Based on the initial quantitative study conducted, it is concluded that different sector variable gives different workload pattern.¹⁶ In this experiment, it is expected that the differences between expert and non-expert subjects' behavior are uncovered and most importantly the workload behavior towards sector complexity variable is identified.

C. Construction of the Solution Space Diagram

The Forbidden Beam Zone (FBZ) of Van Dam et al.¹⁷ has been at the basis for presenting the SSD. It is based on analyzing conflicts between aircraft in the relative velocity plane. Figure 2(a) shows two aircraft, the controlled aircraft (A_{con}) and the observed (or intruder) aircraft (A_{obs}). In this diagram, the protected zone (PZ) of the observed aircraft is shown, a circle with radius of 5 NM (the common separation distance) centered on the observed aircraft. Intrusion of this zone is called a 'conflict', or, 'loss of separation'. Two tangent lines to the left and right sides of the PZ of the observed aircraft are drawn towards the controlled aircraft. The area inside these tangent lines is called the FBZ. This potential conflict can be presented on a SSD. Figure 2(b) shows the FBZ in the SSD of the controlled aircraft. The inner and outer circles represent the velocity limits of the controlled aircraft. Now, if the relative velocity lies inside the beam-shaped area, it means that the aircraft is headed towards the PZ, will eventually enter it, and separation will be lost.

Figure 2: Two Aircraft Condition (a) Plan View of Conflict and the Forbidden Beam Zone Definition. (b) Basic SSD for the Controlled Aircraft (adapted from Mercado-Velasco et al.¹⁵).

When considering the SSD for an aircraft, each neighboring aircraft introduces a zone of conflict, the FBZ in the SSD. The changes in these FBZ have been systematically studied to increase our understanding of the SSD usability in reducing workload and managing sector complexity in the previous case study. Based on the findings gathered in a theoretical study¹⁶ of the effects of traffic situations on the SSD, a number of hypotheses regarding complexity metric have been stated.

II. Hypotheses

The experiment was intended to study the effect of sector complexity variable during a merging scenario to the controller's workload. It is also the intention of the experiment to study and confirm the relation between SSD properties and workload. There are several sector complexity variables that are investigated in this experiment namely intercept angle, number of streams, proximity of incoming aircraft and traffic mix. It is hypothesized that the Solution Space Diagram properties are related to ATCo workload. It is also hypothesized that a certain configuration of these complexity variables leads to different effects on workload and also SSD properties. Some of the hypotheses are based on the findings gathered in the theoretical study¹⁶ of the effects of traffic situations on the SSD. The hypotheses concerning complexity metric are as follow:

1. **Intercept angle:** Larger intercept angle results in different behavior between incoming aircraft and present aircraft. This can be observed through Figure 3 where a static aircraft at 35NM distance from the intercept point observing aircraft at variable distance from the intercept point at variable angle. From the figure it is observed that *larger intercept angle for incoming aircraft* (aircraft with distance more than 35NM) results in a less dense area inside the SSD and therefore lower complexity metrics or workload. However, *larger intercept angle for currently present aircraft* (aircraft with distance less than 35NM) results in a more dense area inside the SSD and therefore higher complexity metrics or workload.¹⁶
2. **Proximity:** Higher *proximity between incoming aircraft position* results in a more dense area inside the SSD and therefore higher complexity metrics or workload.¹⁶
3. **Traffic Mixes:** Having *mix of traffic* (small and large aircraft) imply that different speed limits are present. Different speed limits impose different workload on the ATCo.
4. **Number of Stream:** Having different number of incoming aircraft streams impose different workload on the ATCo.

Figure 3: Intercept Angle of Variable Incoming Distance

III. Experiment setup

An experiment, in which subjects controlled aircraft in a 2 Dimensional ATC simulation with no flight level differences, was conducted using a standalone simulator originally developed using MATLAB® by Mercado-Velasco et al.¹⁵ However, changes have been implemented to the interface in the sector properties in order to match the experiment goal. Figure 4 shows the interface that is used in the experiment. Please note that the color scheme and the font size of the interface have been changed so that the image can be printed out with clarity.

The interface consists of three parts, the Plan View Display (PVD) area, the control panel area and the time information area. The left part is the PVD area that shows the route, the sector under control, the surrounding area of the sector and the aircraft within the area. The right part consists of the speed and heading command area and the respective control buttons that can be used in order to merge the controlled aircraft to the assigned route. In the bottom right part of the simulator, the time information was shown where the Estimated Time of Arrival (ETA) to the sector's exit point is present for all aircraft dedicated to this point.

Figure 4: Experiment Simulator Screenshot.

A. The Plan View Display

The PVD in the right side of the simulator presents the sector under control and the area around the controlled sector. The subject can only give commands when the aircraft is inside the controlled area. The route on which the aircraft have to merge was also present on the PVD. The route points were shown, namely HOOKS, RIVET, OPTIC and WATLE as the last point in this experiment, served as the sector's exit point. The subjects were instructed to have all aircraft exit the sector at this point with a speed of 180 knots.

There are two types of aircraft present during the experiment, which is heavy and light with speed ranges of 160-250 knots, and 120-250 knots, respectively which can be differentiated by the size of the aircraft icon. Different colors were used for the aircraft symbols. When in conflict, the aircraft involved would turn red. After an aircraft was merged to the route either by giving a "Direct To Route Point" or "Intercept Route" command, the color of the data tag would turn from yellow to green. A yellow tag would mean an aircraft flying "free" (just straight ahead, and waiting for a command). The selected aircraft would also show a circle around it, representing the 5 NM separation radius as assumed for this research. Aircraft symbols were shown with a tag, and the information in the tag included:

- First line: aircraft callsign,
- Second line: current speed > set speed (in knots),
- Third line: current heading > set heading (in degree).

B. Control Panel Section

The control panel section consists of a speed and heading command area and also five groups of buttons consisting of one "Intercept Route" button and another four "Direct to Route Point" buttons. The intercept button was used to instruct an aircraft to maintain current heading and speed until the route is crossed and then to continue along the route. The direct to buttons were used to direct an aircraft to fly to a particular route point and from then on to stay on the route.

The speed and heading command area was used to give speed and heading instructions to the aircraft. The outer and inner circles correspond to the maximum and minimum speed of the selected aircraft, respectively. The middle circle represents the 180 knots speed limit, which is the speed that the aircraft were supposed to exit the sector. A set of heading lines to every route is also presented with the name of the point route at the end of the line on the speed and heading command area in order to have a guidance to fly the aircraft towards one of the four route points. A thick black line represents the current speed and heading vector of the aircraft.

To instruct an aircraft to change its speed and/or heading, the controller can use one of three methods. The first method is by positioning the mouse to the desired speed and heading within the maximum and minimum speed circles. This will instruct the aircraft to change both its speed and heading. The second method is by scrolling the mouse up and down in order to increase or reduce the speed vector while maintaining the current heading. The last method is by giving a right click on the speed and heading command area and it is considered as a potential heading

change with constant speed. The desired speed and/or heading instruction were illustrated as a thicker and lighter colored line on top of the current vector line or curved along the speed line according to the desired command. A left mouse click would make the command effective.

C. Time Information Section

Information regarding the remaining time it would take for an aircraft that is instructed to fly on the route to reach the exit point, i.e., WATLE, was shown in the bottom right of the simulator. The time information is only visible when the aircraft is assigned to the route or the sector's exit point and it is labeled with the aircraft callsign.

IV. The Experiment

The experiment was intended to test the impact of several air traffic sector design factors to the workload imposed on the controller during a merging scenario. It is the aim of the experiment to study the effects of the number of streams to be merged, the angle, which the incoming streams originate from, the proximity of incoming aircraft and the variability of the traffic mix of small and large aircraft to the workload imposed on the ATCo and also to study the correlation between the SSD properties to the workload and other performance criterias such as the number of aircraft, number of command et cetera. Each sector design variable was hypothesized to have a different impact on the ATCo's workload.

A. Subjects

Ten male subjects have participated in the experiment. There are two groups of test subjects, representing two population groups: five Graduate Students of the Faculty of Aerospace Engineering and five subjects that participated in an extensive ATC instruction course. These two population groups are named in this research as Students and Experts, respectively. The Experts' ages ranged between 26 and 48 ($\mu = 35.80$, $\sigma = 9.50$), and the Students' ages ranged between 24 and 33 ($\mu = 27.40$, $\sigma = 3.51$). Every subject had a scenario sequence assigned to him. The application of two groups of subjects was to investigate possible differences in control strategies between both groups and also to study the robustness of the method applied.

B. Procedure

Subjects were briefed on the nature of the experiment, the goals to be achieved and the simulator that was used for the experiment. Then, one training session was conducted in order to get the test subject familiarized with the simulator. After the test subject indicated that they understood the task to be performed, a total of 6 full experiment scenarios were presented to the test subjects according to the scenario sequence planned. At the end of each scenario, a questionnaire regarding the experienced workload in merging scenarios was handed out.

C. Scenarios

There are in total six scenarios excluding one training scenario of 20 minutes at the beginning of the experiment (I to VI). Each scenario has been set up to have a distinct character. The scenario settings can be seen in Table 1. In an attempt to gather more information of the dynamics of events, to avoid under-achievement from the controllers, and following the procedure of previous experiment,^{11,14,15} a fast-time simulation was run at 4 times real-time. A negative side effect could be that the subject would experience the task to be very different then in real-time, which would be an undesired effect influencing the experiment results. However, based on previous experiments,^{11,14,15} no such effect was expected. With this simulation speed every scenario was run for 20 minutes real time, with some breaks between scenarios.

Table 1: Scenario Settings

SCENARIO	NO. OF STREAMS	STREAM SETTING	PROXIMITY	PERCENT OF SMALL AIRCRAFT
I	3	Main route, +45°, +100°	High	43 %
II	3	Main route, -45°, +100°	Low	39 %
III	4	Main route, +45°, -45°, +100°	-	41 %
IV	4	Main route, +45°, +100°, -100°	-	42 %
V	3	Main route, +45°, +100°	High	32 %
VI	3	Main route, -45°, +100°	Low	33 %

*+/- angle from the main route

C.1. Incoming Aircraft

There are basically six incoming aircraft sequences. Each sequence is created as if the aircraft entered the sector randomly from the assigned incoming streams. The six incoming aircraft sequence is randomized with a predetermined setting to limit the dependent variable present in the experiment (S1 to S6). From the six scenarios, there are four incoming aircraft patterns during the experiment, two patterns with three streams and two with four streams. The incoming aircraft patterns for three and four streams scenario are presented in Figure 5.

Figure 5: Incoming Aircraft Pattern. (a) Three Streams (Scenario I and V) (b) Three Streams (Scenario II and VI) (c) Four Streams (Scenario III) (d) Four Streams (Scenario IV).

C.2. Traffic Mix and Conditions

The traffic level in every scenario was established with the arrival of new aircraft at fixed intervals of time. Training scenarios had one aircraft every 240 seconds simulated time, while full scenarios had one aircraft every 190 seconds simulated time. In the initial setup of the experiment, there will be eight aircraft during the training session and 11 during full session.

There are six traffic mix situations (T1 to T6), with two conditions having small aircraft ranges between 32% to 33% of the total aircraft (Scenario V and VI) and four conditions having small aircraft between 39% to 43% of the total aircraft (Scenario I to IV). The traffic mix is set to get the required percentage of small and large aircraft mixture (between 30% to 43% of small aircraft) during each scenario.

There are two types of aircraft present in the simulation; heavy and light aircraft with speed ranges of 160-250 knots and 120-250 knot, respectively. The different aircraft types are displayed using different aircraft icon size. The motion of aircraft was controlled only by means of simplified kinematic equations. A constant acceleration (or deceleration) of 3 m/s^2 (5.83 knots/s) and a constant heading change rate of 3 deg/s were used to compute trajectories either for heavy or light aircraft.

C.3. Scenario Combination

Out of the six scenarios, with a certain combination, four sector complexity variable were investigated namely, the proximity of incoming aircraft streams (low and high proximity), the angle in relation to the main route in which the aircraft is coming from (45° or 100° angle in respect to the main route), the aircraft mix (different percentage of

small and heavy aircraft) and the number of incoming aircraft streams (3 streams or 4 streams). More combinations can be found, however, only the respected six cases shown in Table 2 were investigated.

Table 2: Scenario Combination.

Case	Scenario Combination	
	A	B
Proximity of Incoming Aircraft Case 1	Scenario I: High Proximity	Scenario II: Low Proximity
Proximity of Incoming Aircraft Case 2	Scenario V: High Proximity	Scenario VI: Low Proximity
Angle of Incoming Aircraft	Scenario III: 45° Angle	Scenario IV: 100° Angle
Traffic Mix of Incoming Aircraft	Scenario I: 43% Small Aircraft	Scenario V: 32% Small Aircraft
Number of Incoming Stream Case 1	Scenario II: 3 Streams	Scenario III: 4 Streams
Number of Incoming Stream Case 2	Scenario VI: 3 Streams	Scenario III: 4 Streams

To counterbalance the experiment learning curves, the combinations of the scenario sequence are randomized. There are in total six scenarios (I to VI) with six incoming aircraft sequence (S1 to S6) and six traffic mix setting (T1 to T6). The combinations of independent variables in the full scenarios are shown in Table A.1 in the Appendix section.

D. Workload Measure

There are two main workload measures used in this experiment, which are the subjective instantaneous workload rating provided by the test subjects during the experiment and an overall rating at the end of every session. Correlation between the workload measures was conducted by means of several statistical analyses. Each workload measure will be explained in more detail in the next sections.

D.1. Instantaneous Self-Assessment

The workload experienced by the controller varies in time due to the dynamic behavior of the airspace, and for that reason it is crucial that the workload rating is taken at different instances in time. From the many different developed techniques for subjective workload determination, the ISA method is one of the simplest tools with which an estimate of perceived workload can be obtained during real-time simulations or actual tasks.¹⁸ This method requires the operator to give a rating either verbally or by means of a keyboard, of the workload he/she perceives. Due to the ease of implementation and its low intrusiveness, this method was considered ideal for the current experiment.

During the experiment, subjects were asked to give a subjective workload rating every 60 seconds on a scale from 1 (very low) to 7 (very high) by means of a red blinking message on the PVD in an area not taken up by the visualization of traffic. This scale was provided with a regular computer keyboard.

D.2. Questionnaires and Comments

There are two types of questions that were presented to the test subjects; one was in the form of a subjective rating scale and the other one in the form of comments. The questionnaire, in which quantitative and qualitative data were gathered, was provided right after every simulation run was finished. Subjective ratings on a scale from 1 (very low) to 7 (very high) were to be given to the following eight questions in Table 3, constructed to find out how difficult the subjects experienced solving the separation assurance in merging scenarios and why they assessed it as such.

Table 3: Questions Answered by Subjects After Every Experimental Run.

Subjective rating questions
1. How difficult was the separation assurance problem during this scenario?
2. Did you feel the level of traffic you had to handle influenced the difficulty of solving the scenario?
3. Did you feel that time pressure influenced the difficulty of solving the scenario?
4. Did you feel that the incoming aircraft angle influenced the difficulty of solving the scenario?
5. Did you feel that the number of incoming aircraft stream influenced the difficulty of solving the scenario?
6. Did you feel that the combination of incoming aircraft stream influence the difficulty of solving the scenario?
7. Did you feel that the aircraft mix (heavy and light) influenced the difficulty of solving the scenario?
8. Did you feel that aircraft limits influenced the difficulty of solving the scenario?

Comment questions

1. What do you think was hard in this scenario?
 2. What do you think made this scenario easy?
-
-

D.3. Controller's Performance Metrics

During the experimental runs, besides the ISA ratings of workload performed every minute and the questionnaire at the end of every session, there are ten other dependent variables that were being monitored in order to get more insight into the subjects' performance in relation to the workload rating given:

1. **Number of Commands.** Every click in either the Control Panel interface, or one of the "Direct to Route Point" buttons, or the "Intercept Route" button was counted as one command. Despite the fact that a click on the Control Panel interfaces can have the effect of giving two commands (heading and speed changes). However, during this experiment, it is assumed that one click cannot represent a cognitive effort worth of two commands. The contribution of all these types of clicking commands to the total summation of the number of commands was also calculated in terms of a percentage.
2. **Number of Conflicts** that occurred during every simulated scenario. A conflict was defined as a distance between aircraft smaller than 5 NM.
3. **Number of Aircraft Delivered.** The total number of aircraft delivered per subject per scenario.
4. **RMS Aircraft Count.** The root-mean square (RMS) value of the number of aircraft inside the controlled sector, measured every second (simulated time).
5. **RMS Exit Speed.** The fact that subjects were instructed to ensure that aircraft leave the sector at 180 knots, the RMS value of the exit speed of all delivered flights in every scenario was measured as a performance indicator.
6. **RMS Exit Heading.** The RMS value of the exit heading is used to assess the heading of aircraft leaving the sector. This is to monitor the differences in the exit heading of the subject that can be seen as differences in the way information is used by each group or subject.
7. **RMS Exit Separation.** The minimum separation between aircraft while leaving the sector at point WATLE.
8. **RMS Extra Distance Ratio:** For each aircraft, the most efficient trajectory would be a straight line from the point at which the aircraft enters the sector to WATLE. Any other trajectory adds to this distance. The extra distance ratio for every delivered flight was calculated as in Mercado-Velasco et al.¹⁵ in Equation (1). The RMS value of the extra distance ratio of all delivered flights per subject and scenario constitutes this dependent measure.

$$\eta_{ijk} = \frac{d_{ijk} - d_{ijk_{\min}}}{d_{ijk_{\min}}}, \quad (1)$$

where d_{ijk} is the actual flown distance of flight k , controlled by subject i in scenario j , and $d_{ijk_{\min}}$ is the minimum possible distance for that flight.

9. **RMS Smallest Aircraft Separation.** The RMS value of the smallest separation of all delivered flights during their passing through the sector.
10. **RMS Sector Time.** The RMS values of the time aircraft were kept in the sector before being delivered.

D.4. Solution Space Diagram Area Analysis

To relate the workload with area, a continuous calculation of area is being carried out. From previous work by Hermes et al.¹¹ and D'Engelbronner et al.¹⁴ it was concluded that the safe area in the solution space that offers separation violation solutions has a strong (inverse) correlation with ATCo workload level. In this experiment, there are two area calculation methods used, which are the total unsafe area and mean unsafe area for all the aircraft within the sector. Both area calculations used in order to investigate the relation between effected area as a whole or within the performance limit of the aircraft and the workload imposed to the ATCo.

During the experiment, the sector and route geometry were kept constant throughout all scenarios. Only the traffic mix, sequence of entry and point within the airspace where aircraft enter will differ according to the scenario. Previous research of the assessment of the solution space properties concluded that in a four times fast-time simulation, operators would try to plan ahead the development of events approximately 10 minutes in real time.¹⁴ With these reasons in mind, the solution space was calculated to provide information relevant to trajectories of known flight plan 10 minutes ahead of real time. With this setting, the area properties were calculated.

The total area (A_{total}) affected that is covered by the FBZ was calculated as a percentage of the total area between the minimum and the maximum velocity circles in the SSD, based on currently controlled aircraft. The mean area (A_{mean}) affected is calculated based on the A_{total} affected for all aircraft in the sector divided by the number of

aircraft. This will give an overview of the complexity metric for the whole sector. The calculations for the A_{total} and the A_{mean} affected can be seen in Equations (2) and (3), respectively; where A is the area and n is the number of aircraft.

$$A_{total} = \sum A_{affected} \quad (2)$$

$$A_{mean} = \frac{\sum_{i=1}^n A_{total_i}}{n} \quad (3)$$

To thoroughly assess the area properties, both A_{total} and A_{mean} were calculated based on the *whole area presented to the subject* (where n equals to the total number of aircraft present to the controller) and also the *sector under controlled area* (where n equals to the number of aircraft within the sector under controlled). Based on the two situations of sector and whole condition, the area is then calculated based on the *Instance Area Data* gathered every one minute and also the *Mean Area Data* throughout the previous one minute. This brings us to a total of eight area properties and can be seen in Figure 6.

Figure 6: Area Properties Calculation

V. Results and Discussion

A. Outlier Analysis

A total of 18 outliers were detected for the 1,200-workload measurements that have been taken during the experiment clustered based on scenario. All of the outliers are from 6 different subjects in two different scenarios and all coming from both the student and expert groups. Closer inspection of these outliers revealed that the high ISA scores given by the subjects are consistent with a separation violation or a possible separation violation situation and low ISA scores are consistent with low number of aircraft present at the moment. This suggests that the workload was indeed high or low at the moment. Because no unexplainable outliers were found, no outliers were excluded from the analysis, all were accepted to be a part of the subjects' behavioral pattern.

B. Learning Curve

In order to establish whether a significant training effect was present in the experiment, the known estimators for ATCo taskload, which is the Number of Commands, were explored to observe if there are differences between simulation runs. Based on mean and median trend of the number of commands per minute data based on the sequence of simulation runs as seen in Figure 7, the mean data do not differ more than one command per minute between scenarios. This resulted in the conclusion that no learning effect was present in the dataset.

C. Data Analyses and Transformation

All workload measures were tested for normality and homogeneity of variance in order to opt for either parametric test or for Friedman's non-parametric ANOVA. Dependent measures that satisfied normality with the Shapiro-Wilk test and homogeneity of variance with Levene's test were RMS of Sector Time, RMS of Sector Aircraft, RMS of Extra Distance Ratio, RMS of Smallest Separation and Number of Commands.

The subjective ratings performed in the ISA workload ratings performed during the simulation and in the questionnaire were all measured in an ordinal scale. In other words, the scale used was ordered because it goes from

1 (very low workload) to 7 (very high workload), but the increase in workload necessary to increase the rating from 1 to 2 is not necessarily the same as the one required for increasing it from 6 to 7. Therefore in the initial assessment of the all the scenarios, the Ordinal Logistic Model were used in order to preserved its ordinality.

For subsequent data exploration, Z scores of the subjective ratings were used in order to correct for intersubject differences. This correction was performed by calculating the Z scores for every test subject. The Z score for subject i on scenario j was calculated using

$$S_{z_{i,j}} = \frac{S_{i,j} - \bar{S}_i}{\sigma_i}, \quad (4)$$

where S_i represents the subject's mean score along all the scenarios, and σ_i represents the subject's standard deviation in the scoring along all scenarios.

Figure 7: The Number of Commands given per minute

D. ISA Workload Rating

With 20 ratings per simulated scenario, 10 subjects and 6 scenarios per subject, 1,200 subjective ratings of workload were measured during the entire experiment. The subjective ISA ratings performed during the simulation were measured using Ordinal Logistic Regression and the Z score of the subjective ratings using Friedman's ANOVA. Results of the Ordinal Logistic Regression analysis and Friedman's ANOVA are shown in Table 4 and Table 5 respectively.

Based on the results of the Ordinal Logistic regression analysis, it showed that all scenarios except for Scenarios I and II showed high significant values. Group category also showed high significance value, thus proving the importance of analyzing different scenario combination and group behavior. In order to examine the cases and group behavior, Friedman's ANOVA test was performed.

As seen in Table 5, the traffic mix showed a significant value only in the student group with $\chi^2(1) = 18.050$ and $p < 0.001$. It is may be due to the fact that due to the randomization of the traffic, the total difference between scenario 1 and 5 for student is 13% (43% of small aircraft for scenario I and 30% for scenario V) and only 8% (42% of small aircraft for scenario I and 34% for scenario V) for expert. It is also indicated in the questionnaire comment at the end of every scenario runs that they notice the different types of aircraft more because they use the minimum velocity to separate the aircraft in order to get a good merging sequence. Based on the traffic mixed mean ranks value, it showed that a higher mean rank is seen when the traffic have less small aircraft present, thus conforming the comments given by the subjects.

As for the number of incoming aircraft streams, the expert showed highly significant effect ($p < 0.001$) in both cases and only in case 1 ($\chi^2(1) = 4.945$, $p = 0.026$) for student. However, both groups show the same pattern of mean ranks where higher mean rank is seen when there is more streams of incoming aircraft are available.

Table 4: Parameter Estimates and Ordinal Logistic Regression Analysis Result

Parameter	Estim.	Std. Error	Wald	df	Sig.
[Rating = 1]	-1.264	0.216	34.153	1	<0.001
[Rating = 2]	0.663	0.203	10.697	1	0.001
[Rating = 3]	1.988	0.210	89.545	1	<0.001
[Rating = 4]	2.954	0.219	181.759	1	<0.001
[Rating = 5]	4.562	0.253	325.410	1	<0.001
[Rating = 6]	6.255	0.381	269.897	1	<0.001
[Scenario=I]	0.172	0.180	0.913	1	0.339
[Scenario=II]	0.113	0.180	0.395	1	0.530
[Scenario=III]	0.821	0.180	20.757	1	<0.001
[Scenario=IV]	0.919	0.180	25.914	1	<0.001
[Scenario=V]	0.581	0.180	10.455	1	0.001
[Scenario=VI]	0 ^a	.	.	0	.
Group	0.622	0.105	35.474	1	<0.001

Table 5: Friedman's ANOVA Result on ISA Workload Rating

Group	Proximity of Incoming Aircraft							
	Mean Ranks				Mean Ranks			
	$\chi^2(1)$	p	Scen. I	Scen. II	$\chi^2(1)$	p	Scen. V	Scen. VI
Student	3.947	0.047	1.43	1.58	7.385	0.007	1.62	1.38
Expert	2.449	0.118	1.57	1.44	9.281	0.002	1.62	1.39

Group	Number of Incoming Aircraft Streams							
	Mean Ranks				Mean Ranks			
	$\chi^2(1)$	p	Scen. II	Scen. III	$\chi^2(1)$	p	Scen. VI	Scen. III
Student	4.945	0.026	1.41	1.60	2.922	0.087	1.43	1.58
Expert	15.338	<0.001	1.34	1.67	28.595	<0.001	1.27	1.73

Group	Traffic Mix of Incoming Aircraft				Angle of Incoming Aircraft			
	$\chi^2(1)$		Mean Ranks		$\chi^2(1)$		Mean Ranks	
	Scen. I	Scen. V	Scen. I	Scen. V	Scen. III	Scen. IV	Scen. III	Scen. IV
Student	18.05	<0.001	1.31	1.69	0.941	0.332	1.46	1.54
Expert	0.373	0.541	1.53	1.48	0.053	0.819	1.49	1.51

Table 6: Wilcoxon's Signed-Rank Results on ISA Workload Rating

Case	Expert					
	Proximity of Incoming Aircraft		Angle of Incoming Aircraft	Traffic Mix	Number of Stream	
	Case 1	Case 2			Case 1	Case 2
Z	-1.910 ^a	-2.473 ^a	-0.818 ^b	-1.077 ^a	-4.076 ^b	-4.973 ^b
p	0.056	0.013	0.414	0.282	<0.001	<0.001

Case	Student					
	Proximity of Incoming Aircraft		Angle of Incoming Aircraft	Traffic Mix	Number of Stream	
	Case 1	Case 2			Case 1	Case 2
Z	-0.600 ^b	-2.917 ^a	-0.850 ^b	-4.560 ^b	-2.604 ^b	-1.768 ^b
p	0.548	0.004	0.395	<0.001	0.009	0.077

a. Based on negative ranks.

b. Based on positive ranks.

In the case of proximity of incoming aircraft case 2, based on Table 5, significant values were present for both groups ($\chi^2(1) = 9.281$, $p = 0.002$ for expert and $\chi^2(1) = 7.385$, $p = 0.007$ for student) and both groups show the same mean ranks trend. However, it is not the case for proximity of incoming aircraft case 1, where only the student group appears to show statistically significant values and also project the opposite trend in mean ranks value from case 2. This may be due to the differences in strategy while handling incoming aircraft resulted in the final mean separation of aircraft in the sector between scenario I and II for student to be approximately the same for both scenarios. While in case 2 for comparison between scenario V and VI showed that indeed scenario VI have larger mean separation between aircraft in the sector for both groups. Also, the mean ranks showed a similar trend with indicators of scenario VI having lower ranking. This concludes that higher proximity did result in higher workload level as observed from the proximity between incoming aircraft in case 2. The probability plots of all the cases can be seen in Figure 8 (a) to (f) where it confirms that in some cases the differences in ranking given by the subjects is large, showing that the effect is significant and does not invalidate the initial hypotheses.

Figure 8: ISA Workload Rating Probability Distribution (a) Proximity of Incoming Aircraft Case 1 (b) Proximity of Incoming Aircraft Case 2 (c) Angle of Incoming Aircraft (d) Traffic Mix of Incoming Aircraft (e) Number of Incoming Aircraft Streams Case 1 (f) Number of Incoming Aircraft Streams Case 2.

Post Hoc analyses of the ISA workload rating are conducted using Wilcoxon's Signed-Rank analysis with Bonferroni correction having significance cutoff at 0.025 (based on 2 comparative scenarios). Both tests showed almost the same pattern excluding one case where using Friedman's ANOVA, the significance level is $p = 0.047$ for proximity of incoming aircraft for case 1 whereas analysis using Wilcoxon's Signed Ranks analysis showed $p = 0.548$. The results of Wilcoxon's Signed-Rank analysis are shown in Table 6. Overall, it is concluded that apart from the case mention above, other cases were confirmed to be statistically significant as in Friedman's ANOVA.

E. Questionnaire Workload Rating

The effects of the independent variables on the eight items included in the questionnaire were analyzed using Friedman's ANOVA. A total of 480 subjective ratings (8 questions x 6 scenarios x 10 subjects) were measured during the experiment. Based on the analysis, only three questions during the case of number of incoming aircraft streams in case 1 showed statistically significant values and it is shown in Table 7. Significant values were found for the question that regards to the level of traffic, the time pressure effect on subject and the effect of incoming aircraft angle. Based on the mean rank values, higher mean rank was seen in scenario where four streams available. This also means that the difficulty experienced by the subjects in relation to the level of traffic, the time pressure and the angle of incoming aircraft is higher in cases where more streams of aircraft are present in the sector.

As for comments gathered at the end of every simulation run, according to the respective scenario the subjects mention the number of streams, the angle or from which direction where the incoming aircraft originate and the differences in speed of incoming aircraft did contribute to the difficulty of performing the task required. This is because the subjects have to line the aircraft with different speed to a single route. The angle of incoming aircraft was also mentioned by a few subjects for the conditions where aircraft were coming in only from the north, which causes limited airspace to work with, and also for the condition with aircraft coming from both north and south at the same time, creating possible future collisions. This shows that combination of incoming aircraft angle and also limited airspace is seen as some of the contributing factors of sector complexity.

F. Performance Metric

Apart from the ISA workload ratings and Questionnaire rating, there are ten other performance metrics related analysis that has been carried out in order to investigate the performances of the subject in different case studies. The performance data that satisfies the assumptions of normality with the Shapiro-Wilk test and homogeneity with Levene's test of variance are analyzed using Factorial ANOVA (Table 8) and performances metric that does not satisfy both assumptions are analyzed with Friedman's non-parametric ANOVA (Table A.2).

Table 8 shows results of the performance metrics analyzed with Factorial ANOVA, statistically significant values are underlined. Based on the performance data, only one performance that is statistically significant which is the RMS Smallest Separation ($F = 5.975$, $p = 0.026$) in case of different incoming aircraft angle. Other performances did not show any significant value in all the cases. However, both groups behaved differently and this is shown in the analysis data where in 16 out of 30 performance metric data (5 performance metric x 6 case studies), the group factor showed a statistically significant value (p between 0.001 to 0.030). In the other 14 performance metrics, the significance value lied between $p = 0.055$ (RMS Extra Distance Ratio) to $p = 0.981$ (RMS Smallest Separation in case of different incoming angle). Based on the results in Table 8, there are no factorial performance metric and group that results in significant value. This concludes that the interaction between cases and groups is not statistically significant, thus conclude that both groups behave differently during the experiment and also performed differently.

Performance metrics in Table 8 showed that the group in proximity of incoming aircraft case 1 showed no significant value ($p < 0.05$) except in number of command. However, in proximity of incoming aircraft case 2, significant value were present in all performance metrics. The difference between these two cases that is case 1 has the approximately the same traffic mix values (42% and 43% small aircraft for High Proximity case and 39% small aircraft for Low Proximity case) and case 2 has different traffic mix values (34% and 30% small aircraft for High Proximity case and 31% and 35% small aircraft for Low Proximity case).

Table 7: (a) Friedman's ANOVA Analysis Result on Questionnaires Workload.

Question 2: Level of Traffic				
Group	$\chi^2 (1)$	p	Mean Ranks	
			Scenario II	Scenario III
Student	1	0.3173	1.3000	1.7000
Expert	4	<u>0.0455</u>	1.1000	1.9000

Question 3: Time Pressure				
Group	$\chi^2 (1)$	p	Mean Ranks	
			Scenario II	Scenario III
Student	0	1	1.5000	1.5000
Expert	4	<u>0.0455</u>	1.1000	1.9000

Question 4: Incoming Aircraft Angle				
Group	$\chi^2 (1)$	p	Mean Ranks	
			Scenario II	Scenario III
Student	4	<u>0.0455</u>	1.1000	1.9000
Expert	1	0.3173	1.3000	1.7000

Table 8: Factorial ANOVA Results for Performance Metrics

Factor	RMS Sector Time		RMS Sector Aircraft		RMS Extra Distance Ratio		RMS Smallest Separation		Number of Command	
	F	p	F	P	F	p	F	p	F	p
Proximity 1	0.709	0.412	1.175	0.294	0.284	0.602	1.438	0.248	0.001	0.973
Group	3.169	0.094	3.329	0.087	3.706	0.072	0.826	0.377	5.652	<u>0.030</u>
Proximity 1*Group	0.263	0.615	0.52	0.481	0.231	0.627	3.150	0.095	0.134	0.719
Proximity 2	0.327	0.575	0.146	0.707	0.541	0.473	1.705	0.210	0.181	0.676
Group	11.624	<u>0.004</u>	12.787	<u>0.003</u>	15.263	<u>0.001</u>	6.226	<u>0.024</u>	9.434	<u>0.007</u>
Proximity 2*Group	0.002	0.969	0.018	0.896	0.003	0.959	0.707	0.413	0.112	0.742
Angle	0.241	0.630	0.396	0.538	1.352	0.262	5.975	<u>0.026</u>	0.542	0.472
Group	2.361	0.144	3.202	0.092	2.002	0.176	0.001	0.981	8.176	<u>0.011</u>
Angle*Group	0.158	0.696	0.018	0.895	2.014	0.175	0.208	0.654	0.001	0.977
Traffic	0.032	0.861	0.033	0.858	0.098	0.758	0.146	0.708	0.146	0.707
Group	6.529	<u>0.021</u>	7.216	<u>0.016</u>	8.349	<u>0.011</u>	5.705	<u>0.030</u>	5.666	<u>0.030</u>
Traffic*Group	0.266	0.613	0.253	0.621	0.138	0.715	0.286	0.600	0.138	0.715
Num. of Stream 1	0.293	0.596	0.407	0.533	0.005	0.946	0.620	0.442	0.086	0.773
Group	2.271	0.151	2.791	0.114	0.689	0.419	0.551	0.469	6.566	<u>0.021</u>
Num. of Stream 1*Group	0.001	0.976	0.153	0.701	0.697	0.416	0.101	0.755	0.067	0.799
Num. of Stream 2	0.785	0.389	0.615	0.444	0.988	0.335	2.669	0.122	0.103	0.752
Group	7.018	<u>0.017</u>	8.454	<u>0.010</u>	4.281	0.055	2.649	0.123	9.393	<u>0.007</u>
Num. of Stream 2*Group	1.549	0.231	0.868	0.365	4.305	0.055	3.975	0.064	0.413	0.530

Figure 9: Error Bars for Performance Metrics analysed with Factorial ANOVA in the case Proximity between Incoming Aircraft case 2 (a) RMS Sector Aircraft (b) RMS Extra Distance Ratio (c) RMS Sector Time (d) RMS Smallest Separation (e) Number of Commands.

The same pattern goes in the number of incoming aircraft stream cases where RMS Sector Time and RMS Sector Aircraft showed significant value in case 2 but not in case 1. For the number of incoming aircraft stream, the traffic mixes is approximately the same in case 1 (39% for three streams case and 41% small aircraft for four streams case) and case 2 have different traffic mix values (31% and 35% small aircraft to 41% small aircraft for three streams and four streams case respectively). It is observed that the combination of two sector design variables (proximity and number of incoming aircraft streams with traffic mixes) results in the performance metrics showing significant values in case 2 and this cannot be observed in case 1. This also means that the performances of the subjects vary between cases with only one or two sector design variable.

The influence of group on the number of commands is significant ($p < 0.05$) for all cases, indicating that the two groups behaved differently. This can already be seen in Figure 9. The elaboration of differences in number of command behaviour is explained in the subsequent section. The different traffic mix showed statistically significant value on all of the performance metrics ($p < 0.05$) whereas for the case of different angle of incoming aircraft only the number of command showed significant value ($p = 0.011$). Figure 9 shows the means and 95% confidence intervals between the two means of performance metrics for both the student and expert groups, for the low and high proximity conditions. The figures confirm the findings that indeed the performance metrics did not show large differences between cases, but larger differences between group for certain cases.

Figure 10: Error Bars for Performance Metrics analysed with Friedman's Analysis in the case Angle of Incoming Aircraft (a) Number of Aircraft Delivered (b) Number of Conflict (c) RMS Exit Speed (d) RMS Exit Heading (e) RMS Exit Separation.

For the performance metrics that did not satisfy the normality and homogeneity assumption, a non-parametric Friedman's ANOVA was carried out. These were number of aircraft delivered, number of conflicts, the RMS exit speed and RMS exit separation. Results are shown in Table A.2 in the Appendix section. For the expert group, the angle of incoming aircraft displayed a significant difference in number of aircraft delivered ($\chi^2(1) = 4.000$, $p = 0.046$) and RMS Exit Separation ($\chi^2(1) = 5.000$, $p = 0.025$) with both having mean rank showing better performance (more aircraft delivered and smaller RMS exit separation) in bigger incoming aircraft angle. However, it is assumed that both performance metrics are influenced by the fact that in bigger incoming angle case, the aircraft is closer to the exit point, results in relatively more aircraft that can be delivered and smaller separation during exit.

The student are affected by the proximity of incoming aircraft in case 1 where significant value is present for the RMS exit speed ($\chi^2(1) = 4.000$, $p = 0.046$) with both cases showing a large deviance between scenarios from 180 knots assignment of exit speed and in number of incoming aircraft streams showing significant value in RMS Exit Heading ($\chi^2(1) = 5.000$, $p = 0.025$) where the exit heading in scenario with more streams showing a large deviance from the route angle having a mean of 7.54° . However a more interesting aspect to see if the differences between RMS Exit Heading in student and expert where the expert showed a larger deviance from the route angle than student with expert having median of 14.50° compared to 0.50° for student. This showed that both groups have different control strategy. Figure 10 shows the means and 95% confidence intervals between the two means of performance metrics for both the student and expert groups, for the different angle on incoming aircraft.

In post hoc analysis using Wilcoxon's Signed-Rank test analysis only two data in the case of the angle of incoming aircraft, namely the RMS exit separation and number of aircraft delivered both for expert group ($p = 0.043$ and 0.046 respectively), which is statistically significant. However, evaluating the results with Bonferroni correction having significance cutoff at 0.025 (based on 2 comparative scenario), no significant value was found.

C.1. Commands Percentages

The number of command for both experts and students is different in the entire scenario. Nevertheless, it is more interesting to look into the distribution of the commands where it is divided into percentage of intercept clicks, percentage of direct to's click and percentage of speed and heading instruction. Table 9 shows the result of the Factorial ANOVA done on the commands percentage. Based on the result, only the group difference is statistically significant in almost all of the cases. Closer inspection to Figure 11 confirms that different group indeed behaved differently in certain cases.

Figure 11: Error Bar of Commands Percentage based on Proximity between Incoming Aircraft Case 1 (a) Intercept Click Percentage (b) Direct To's Percentage (c) Speed and Heading Instruction Percentage.

Intercept percentage shows statistically significant value for group in all other case except for proximity between incoming aircraft case 2 and traffic mix. This means that different intercept click pattern applies for almost all cases between groups. Thus relates to differences in the control strategy between expert and student.

However, it is interesting to see the differences between command pattern in proximity between aircraft case 1 where it shows significance value in intercept, direct to's and speed and heading clicks percentage and for case 2 where no significance value occur at all. The difference in traffic mixes between these two cases was discussed in earlier section where it is presumed that the influence of traffic mixes differences might have changed the way commands are given that results in proximity cases where the clicking patterns is significant (case 1) being insignificant (case 2). The case on traffic mix itself did not reveal statistical significance differences in the way the commands are given. This can be observed from the direct to patterns between cases 1 and 2 in Figure 12 (a) and (b). The same mixed result in the number of stream case were found with case 1 being significant and case 2 being insignificant for the direct to's percentage. Again, the differences in the traffic pattern between both cases are presumed to influence the way the commands are given. This can be observed from Figure 12 (d) and (e).

Table 9: Factorial ANOVA Analysis for Command Percentage

Factor	Intercept Click Percentage		Direct To's Click Percentage		Speed and Heading Click Percentage		Airspace Click Percentage	
	F	p	F	p	F	p	F	p
Proximity 1	0.235	0.635	0.016	0.902	0.225	0.642	0.004	0.953
Group	5.075	<u>0.039</u>	8.231	<u>0.011</u>	6.179	<u>0.024</u>	0.641	0.435
Proximity 1*Group	0.471	0.502	0.037	0.85	0.029	0.867	0.132	0.721
Proximity 2	0.002	0.965	0.161	0.693	0.296	0.594	0.078	0.783
Group	0.001	0.981	1.544	0.232	2.421	0.139	0.341	0.567
Proximity 2*Group	0.582	0.457	0.446	0.514	0.068	0.798	0.666	0.426
Angle	0.01	0.923	0.265	0.614	0.542	0.472	0.654	0.431
Group	7.928	<u>0.012</u>	5.067	<u>0.039</u>	1.177	0.294	3.505	0.080
Angle*Group	1.031	0.325	0.03	0.864	0.195	0.665	0.026	0.873
Traffic	0.172	0.684	0.236	0.634	1.008	0.33	0.125	0.728
Group	0.101	0.755	2.261	0.152	3.766	0.70	2.197	0.158
Traffic*Group	1.089	0.312	1.068	0.317	0.535	0.475	0.001	0.971
Num. of Stream 1	0.776	0.391	0.284	0.601	1.395	0.255	0.088	0.771
Group	14.19	<u>0.002</u>	8.931	<u>0.009</u>	2.284	0.150	1.164	0.297
Num. of Stream 1*Group	0.278	0.605	0.21	0.66	0.755	0.398	0.458	0.508
Num. of Stream 2	0.541	0.473	0.303	0.59	1.191	0.291	0.021	0.886
Group	5.631	<u>0.031</u>	4.363	0.053	1.164	0.297	0.709	0.412
Num. of Stream 2*Group	2.219	0.156	0.107	0.747	0.245	0.627	1.088	0.312

Figure 12: Frequency Plots of Direct To's. (a) Proximity of Incoming Aircraft Case 1 (b) Proximity of Incoming Aircraft Case 2 (c) Angle of Incoming Aircraft (d) Traffic Mix of Incoming Aircraft (e) Number of Incoming Aircraft Streams Case 1 (f) Number of Incoming Aircraft Streams Case 2.

Overall, the way the group behave in giving direct to's command is different between one another. In all cases, the expert group showed increasing command to the final route point, whereas for student group, the middle two route points are the main directed route points. This shows different styles of separation assurance between groups where experts uses both vector and speed to ensure separation and student mostly use speed commands to ensure separation.

The additional time tool provided to the controller during the experiment may have also resulted the differences in strategy between the groups to be bigger than expected. The expert group found that the tool is quite useful and have used the tool to direct the aircraft straight to the last waypoint, whereas the student group did not use the tool as frequent and have used the direct to tools to merge the aircraft to the route points before exiting the sector.

G. Solution Space Diagram Area Analysis

The SSD area values do not satisfy the normality and homogeneity assumption based on the Shapiro-Wilks and Levene's test. Thus, the analysis of the SSD area was extended by performing Friedman's ANOVA on both A_{total} and A_{mean} of the SSD for all 8 area properties. Based on the analysis, all of the area properties that were significant showed the same trends in the mean ranks values. Area properties for the whole area shown to the controller using the mean area data of the last one minute have the most number of cases having significant value. Thus, to better investigate the effects of sector variable to SSD area properties, the results of whole area using mean area data were used and shown in Table 10. The values that were statistically significant were underlined.

Table 10: Friedman's ANOVA for A_{total} and A_{mean} for Whole Projected, Mean Area Data

Group	A_{total}								A_{mean}							
	Proximity of Incoming Aircraft								Proximity of Incoming Aircraft							
	Mean Ranks				Mean Ranks				Mean Ranks				Mean Ranks			
	$\chi^2(1)$	p	Scen. I	Scen. II	$\chi^2(1)$	p	Scen. V	Scen. VI	$\chi^2(1)$	p	Scen. I	Scen. II	$\chi^2(1)$	p	Scen. V	Scen. VI
Student	14.440	<u><0.001</u>	1.66	1.31	4.840	<u>0.028</u>	1.61	1.39	5.760	<u>0.016</u>	1.62	1.38	4.000	<u>0.046</u>	1.60	1.40
Expert	10.240	<u>0.001</u>	1.69	1.34	2.560	0.110	1.58	1.42	12.960	<u><0.001</u>	1.68	1.32	1.440	0.230	1.56	1.44
Group	Number of Incoming Aircraft Streams								Number of Incoming Aircraft Streams							
	Mean Ranks				Mean Ranks				Mean Ranks				Mean Ranks			
	$\chi^2(1)$	p	Scen. II	Scen. III	$\chi^2(1)$	p	Scen. VI	Scen. III	$\chi^2(1)$	p	Scen. II	Scen. III	$\chi^2(1)$	p	Scen. VI	Scen. III
	$\chi^2(1)$	p	Scen. II	Scen. III	$\chi^2(1)$	p	Scen. VI	Scen. III	$\chi^2(1)$	p	Scen. II	Scen. III	$\chi^2(1)$	p	Scen. VI	Scen. III
Student	0.040	0.841	1.51	1.49	1.000	0.317	1.45	1.55	6.760	<u>0.009</u>	1.63	1.37	6.760	<u>0.009</u>	1.63	1.37
Expert	4.000	<u>0.046</u>	1.40	1.60	0.040	0.841	1.49	1.51	1.440	0.230	1.44	1.56	0.360	0.549	1.47	1.53
Group	Traffic Mix of Incoming Aircraft				Angle of Incoming Aircraft				Traffic Mix of Incoming Aircraft				Angle of Incoming Aircraft			
	Mean Ranks				Mean Ranks				Mean Ranks				Mean Ranks			
	$\chi^2(1)$	p	Scen. I	Scen. V	$\chi^2(1)$	p	Scen. III	Scen. IV	$\chi^2(1)$	p	Scen. I	Scen. V	$\chi^2(1)$	p	Scen. III	Scen. IV
	$\chi^2(1)$	p	Scen. I	Scen. V	$\chi^2(1)$	p	Scen. III	Scen. IV	$\chi^2(1)$	p	Scen. I	Scen. V	$\chi^2(1)$	p	Scen. III	Scen. IV
Student	4.840	<u>0.028</u>	1.61	1.39	5.760	<u>0.016</u>	1.38	1.62	0.040	0.841	1.51	1.49	31.360	<u><0.001</u>	1.22	1.78
Expert	0.640	0.424	1.54	1.39	4.000	<u>0.046</u>	1.40	1.60	2.560	0.110	1.58	1.42	11.560	<u>0.001</u>	1.33	1.67

In proximity between incoming aircraft case 1, the results showed significant values in both A_{total} and A_{mean} properties for both groups with $p < 0.001$ (student) and $p = 0.001$ (expert) for A_{total} and $p = 0.016$ (student) and $p < 0.001$ (expert) for A_{mean} . The mean ranks values also showed the same pattern with the ISA workload rating discussed in the previous sections with incoming aircraft having higher proximity have higher mean ranks than incoming aircraft with lower proximity. The same pattern goes for proximity between incoming aircraft case 2 where significant values were present in student group with $\chi^2(1) = 4.840$ and $p = 0.028$ for A_{total} and $\chi^2(1) = 4.000$ and $p = 0.046$ for A_{mean} . This also concludes that higher proximity relates to higher workload.

In the case of the number of incoming aircraft streams, it is observed that only in case 1 for experts where the A_{total} properties showed a statistically significant value ($\chi^2(1) = 4.000$, $p = 0.046$). The others showed no significant values. However, for the A_{mean} , the number of incoming aircraft showed a statistically significant value in both case 1 and 2 for students with both cases showing $\chi^2(1) = 6.760$, $p = 0.009$. In A_{total} scenario with more streams showed a higher mean ranks thus agreeing with the pattern found in the ISA workload rating. Whereas in A_{mean} the opposite pattern occurs where higher mean ranks were found in cases where smaller number of streams. This might be due to the fact that with less number of streams, the aircraft are more concentrated in certain areas of the sector leaving the A_{mean} properties having higher percentage values. This also means that the concentration of aircraft has a bigger influence on the mean area properties than the number of streams.

Table 11: Wilcoxon Signed-Rank Results on A_{total} and A_{mean} for Whole Projected, Mean Area Data

Group		A_{total}					
		Proximity of Incoming Aircraft		Angle of Incoming Aircraft	Traffic Mix	Number of Stream	
		Case 1	Case 2			Case 1	Case 2
Expert	Z	-3.961 ^a	-1.540 ^a	-1.829 ^b	-1.073 ^a	-1.943 ^b	-0.921 ^b
	p	<u><0.001</u>	0.123	0.067	0.283	0.052	0.357
Student	Z	-2.562 ^a	-2.242 ^a	-1.747 ^b	-2.826 ^a	-0.536 ^a	-1.159 ^b
	p	<u>0.010</u>	<u>0.025</u>	0.081	<u>0.005</u>	0.592	0.247
		A_{mean}					
		Proximity of Incoming Aircraft		Angle of Incoming Aircraft	Traffic Mix	Number of Stream	
		Case 1	Case 2			Case 1	Case 2
Expert	Z	-4.380 ^a	-1.272 ^a	-3.686 ^b	-2.194 ^a	-1.516 ^b	-0.849 ^b
	p	<u><0.001</u>	0.203	<u><0.001</u>	<u>0.028</u>	0.129	0.396
Student	Z	-2.854 ^a	-1.681 ^a	-5.979 ^b	-0.739 ^a	-1.846 ^a	-2.864 ^a
	p	<u>0.004</u>	0.093	<u><0.001</u>	0.460	0.065	<u>0.004</u>

a. Based on positive ranks.

b. Based on negative ranks.

Figure 13: Error Bar for A_{mean} Percentage. (a) Proximity of Incoming Aircraft Case 1 (b) Proximity of Incoming Aircraft Case 2 (c) Angle of Incoming Aircraft (d) Traffic Mix of Incoming Aircraft (e) Number of Incoming Aircraft Streams Case 1 (f) Number of Incoming Aircraft Streams Case 2.

Traffic mix only showed significant values in A_{total} ($\chi^2(1) = 4.840$, $p = 0.028$) for student where it is found that when more small aircraft is present, the higher mean ranks were found based on the area properties. However, the pattern is opposite from what is found from the ISA workload rating analysis. Nonetheless, significant values in both ISA and area properties analysis indicates that the traffic mixes indeed effect the level of sector complexity.

Angle of incoming aircraft showed significant values in both A_{total} and A_{mean} for both groups with $p = 0.016$ and $p = 0.046$ for A_{total} (student and expert, respectively) and $p < 0.001$ and $p = 0.001$ for A_{mean} (student and expert, respectively). The mean ranks of angle of incoming aircraft cases in the area properties studies showed that bigger intercept angle results in higher mean rank values. This corresponds to the hypothesis that larger intercept angle of currently present aircraft results in higher workload to the controller.

Post hoc analysis was performed using Wilcoxon's Signed-Ranked Analysis with Bonferroni correction having significance cutoff at 0.025 (based on 2 comparative scenario) and the results are shown in Table 11. Based on the analysis, indeed the values that are highly significant in Friedman's ANOVA were confirmed significant except in the case of angle of incoming aircraft where $p = 0.081$ (where $p = 0.016$ is obtained using Friedman's ANOVA) for A_{total} and in the case of number of incoming aircraft stream where $p = 0.065$ (where $p = 0.009$ is obtained using Friedman's ANOVA) for A_{mean} . Figure 13 showed the box plot bar of A_{mean} percentage where the significant results of certain cases can be seen.

H. Correlation Analyses

Correlation analyses were conducted to determine how well the SSD area properties correlate with the subjects' ISA Z score workload ratings, the number of aircraft presented to the controller (n_{All}), the number of aircraft within the controlled sector (n_{Sect}), the number of commands ($N_{Command}$) and the number of conflicts ($N_{Conflict}$). The correlation study were conducted based on the overall situation where all subjects were investigated as one and also based on the grouping of expert and student, to investigate the possible effect of experienced and inexperienced subject.

The nonparametric Kendall's tau correlation coefficient was used instead of its parametric counterpart, the Pearson correlation, since this coefficient does not require data to be normally distributed or to be interval scale.¹⁹ However, this method gives correlations that cannot be compared to correlations obtained via Pearson's method due to the differences between the methods.

H.1. Subject Correlations

Initial individual subject correlation was carried out. This is to ensure consistent correlation of each subject with the workload related measures. Based on the analysis, nine out of ten subjects showed the same consistency of correlation effect. Only one of the experts had odd correlation values, however it is only present in the case where the correlation in concern is based on the subject's rating score. Other values are considered comparable. Thus, the subject is not removed from the analysis.

H.2. Cross correlation

Cross correlation analysis performed between several workload related measures. Based on the analysis, without considering the group difference, all combination showed significance value with $p < 0.001$. Looking at group category, all combination showed significance value at the 0.01 level except for one combination having significant value at higher than 0.01 level, which is in analysis between the $N_{Command}$ to A_{total} for student group in cases using mean area data case.

Table 12 showed the comparison between the SSD area properties correlation with other sector complexity measures such as n_{Sect} , n_{All} , $N_{Command}$ and $N_{Conflict}$ to the ISA Z score gathered from the ISA workload rating given by the subject. Based on the table, the SSD area properties, namely A_{mean} showed the largest correlation using whole projected area with $R = 0.228$ (instance area data) and $R = 0.227$ (mean area data) compared to other sector complexity measure where $R = 0.181$ (n_{Sect}), $R = 0.200$ (n_{All}), $R = 0.189$ ($N_{Command}$) and $R = 0.222$ ($N_{Conflict}$). The A_{mean} using whole projected area and sector area data showed the highest correlation with the ISA Z score compared to n_{Sect} and n_{All} for both groups. This concludes that the A_{mean} give the best correlation thus far to the workload as indicated by the subject, especially in case where the whole area projected to the controller is taken into consideration. However, when comparing the area properties correlation with the ISA Z score with the $N_{Command}$ and

Table 12 (a): ISA Z score to Area Properties.

	Whole Area		Sector Area	
	A_{total}	A_{mean}	A_{total}	A_{mean}
Using Instance Area Data				
All	0.202	<u>0.228</u>	0.209	0.222
Expert	0.235	<u>0.275</u>	0.232	0.262
Student	0.164	0.194	0.185	<u>0.198</u>
Using Mean Area Data				
All	0.122	<u>0.227</u>	0.117	0.222
Expert	0.128	<u>0.278</u>	0.120	0.270
Student	0.111	<u>0.191</u>	0.118	0.189

Table 12 (b): ISA Z score to Number of Aircraft, Command and Conflict Properties.

	n_{Sect}	n_{All}	$N_{Command}$	$N_{Conflict}$
All	0.181	0.200	0.189	<u>0.222</u>
Expert	0.233	0.246	<u>0.282</u>	0.281
Student	0.145	0.173	<u>0.206</u>	0.182

$N_{Conflict}$ within group, the A_{mean} showed a slightly lower correlation level compared to the $N_{Command}$ for both groups and a lower correlation than $N_{Conflict}$ for expert group.

The correlation between the SSD area properties and the $N_{Command}$ and $N_{Conflict}$ were also investigated. Table 13 showed the results of the correlation analyses. Based on the results, using the mean area data, the A_{mean} using whole projected area showed the largest correlation with both $N_{Command}$ and $N_{Conflict}$ ($R = 0.215$ and $R = 0.259$, respectively). However, using instance area data, mixed results were gathered where A_{mean} using whole projected area showed the largest correlation compared to other SSD area properties with the $N_{Command}$ ($R = 0.212$) whereas A_{total} using whole projected area showed the largest correlation with the $N_{Conflict}$ ($R = 0.275$). However, in both cases the significant level is at 0.01. Overall, the A_{mean} using the whole area showed the highest correlation in the most case.

Table 13: SSD Area Properties to Number of Command and Conflict Properties.

	$N_{Command}$				$N_{Conflict}$			
	Whole Area		Sector Area		Whole Area		Sector Area	
	A_{total}	A_{mean}	A_{total}	A_{mean}	A_{total}	A_{mean}	A_{total}	A_{mean}
Using Instance Area Data								
All	0.151	<u>0.212</u>	0.167	0.206	<u>0.275</u>	0.254	0.249	0.226
Expert	0.171	0.138	<u>0.197</u>	0.119	0.238	<u>0.256</u>	0.220	0.221
Student	0.112	0.168	0.144	<u>0.193</u>	<u>0.304</u>	0.225	0.274	0.203
Using Mean Area Data								
All	0.081	<u>0.215</u>	0.088	0.201	0.147	<u>0.259</u>	0.126	0.229
Expert	0.111	0.121	<u>0.129</u>	0.104	0.169	<u>0.266</u>	0.168	0.232
Student	0.051*	0.193	0.059*	<u>0.195</u>	0.132	<u>0.224</u>	0.095	0.194

* Correlation at higher than 0.01 level

Correlation analysis between the area properties and number of aircraft properties were also carried out and the results are laid out in Table 14. Based on the figures in Table 14, A_{mean} have higher correlation with the number of aircraft for both n_{Sect} and n_{All} . Overall, the A_{mean} gathered from data collected based on aircraft within the sector area have highest correlation with n_{Sect} ($R = 0.440$ and $R = 0.476$ using instance area data and mean area data, respectively), and A_{mean} gathered from data collected based on all aircraft presented to the controller have highest correlation with n_{All} ($R = 0.512$ and $R = 0.513$ using instance area data and mean area data, respectively).

Table 14: SSD Area Properties to Number of Aircraft Properties.

	n_{Sect}				n_{All}			
	Whole Area		Sector Area		Whole Area		Sector Area	
	A_{total}	A_{mean}	A_{total}	A_{mean}	A_{total}	A_{mean}	A_{total}	A_{mean}
Using Instance Area Data								
All	0.251	0.420	0.226	<u>0.440</u>	0.260	<u>0.512</u>	0.241	0.505
Expert	0.233	0.421	0.238	<u>0.511</u>	0.270	0.544	0.278	<u>0.600</u>
Student	0.259	<u>0.362</u>	0.223	0.323	0.247	<u>0.432</u>	0.222	0.370
Using Mean Area Data								
All	0.149	0.447	0.151	<u>0.476</u>	0.144	<u>0.513</u>	0.123	0.481
Expert	0.152	0.458	0.173	<u>0.554</u>	0.160	0.558	0.152	<u>0.570</u>
Student	0.153	<u>0.378</u>	0.150	0.356	0.137	<u>0.420</u>	0.115	0.346

Figure 14 (a) to (d) show the plots of workload in regards to the number of aircraft and also area properties taken every minute in all six scenarios (6 scenarios x 20 ratings per scenario) for one subject. In Figure 14 (e) and (f), the plots between A_{mean} and the number of aircraft were illustrated. All of the plots illustrate the level of correlation between the parameters as discussed in the earlier section.

VI. Conclusions and Recommendations

In experiments with a simplified Air Traffic Control simulation, the workload level as reported by the controller and also from the SSD area properties, confirms the findings on the theoretical study conducted previously¹⁶ on the effects of traffic situations on the SSD. First, in cases where horizontal proximity of the incoming aircraft is high, the workload is reported to be higher. This is confirmed in both ISA workload rating and SSD area properties. Secondly, a larger incoming aircraft angle produces a higher level of mean ranks, thus implicating higher workload apparent only through the study of SSD properties. Even though it is indicated in an earlier study that a larger intercept or merging angle of incoming aircraft produces a less dense Solution Space, the effect of currently present aircraft is more prevalent than that of incoming aircraft. Thus, the workload as perceived by the controller is higher as the intercept or merging angle is bigger.

Figure 14: Plots of Parameter Comparison for One Subject. (a) Workload to the Number of Aircraft in the Area (b) Workload to the Number of Aircraft in Controlled Sector (c) Workload to A_{total} (d) Workload to A_{mean} (e) A_{mean} to All Aircraft (f) A_{mean} to Sector Aircraft.

The third sector variable showed that based on the ISA workload rating, the traffic mix condition is only significant for students where they give lower workload rating with larger variation of traffic. In spite of this, it is already indicated by the student group subjects that differences in speed range for large and small aircraft gives them more room to control and space the aircraft in the sequence. However, the effect of traffic mix alone in the result of

the SSD area properties is only apparent in A_{total} for student group but in the opposite manner from the ISA rating analysis where larger variation of traffic give higher mean ranks in SSD area properties. However, the fact that differences in traffic mixes level, showed significant value in area properties showed that indeed the traffic mixes did have influence on the sector complexity.

Lastly is the number of streams, where in A_{total} more streams showed a higher mean ranks thus agreeing with the pattern found in the ISA workload rating however, for A_{mean} , the opposite criteria were discovered between workload study and SSD properties in, where more streams indicated lower mean ranks using the SSD area properties. This might be due to the fact that with less number of streams, the aircraft are more concentrated in certain areas of the sector leaving the area properties having higher percentage values. This also means that the concentration of aircraft has a bigger influence on the area properties for A_{mean} than the number of streams.

Correlation study also showed promising results with overall correlation between the ISA workload rating and the SSD properties, namely the A_{mean} properties were found to be in a higher level than with the n_{Sect} , n_{All} , $N_{Command}$ and $N_{Conflict}$. Despite the fact that both groups performed differently and had a different control strategy, the SSD area properties were found to be in a higher correlation in both groups with the ISA Z score than with number of aircraft, proving that the SSD-based analysis to be a good workload indicator. The A_{mean} of the instance area data were also found to have the highest correlation between all SSD area properties with all aircraft properties were taken into consideration.

Despite the good correlation values, for future, it is recommended that a further study of SSD area properties is conducted. Differences in the behavior between the SSD area properties and ISA workload rating especially in the traffic mixes and the number of streams cases showed that even when the SSD area properties showing the highest correlation with the ISA workload rating, more area needed to be explore from the area properties in order to enable SSD area properties being used not only as a workload indicator but also as sector planning tool. It is also recommended in the future to study the effect of the time tool towards the controller strategy.

For the exploration of sector design variable, it is recommended for future studies that the effect between sector design variable and the combination of sector design variable that leads to higher workload to be studied. Professionals ATCo were also expected to participate in future studies, to investigate differences between experts and professional ATCo, and also ensuring that the same level of correlation can be achieved.

Appendix

Table A.1: Scenario in Combination with Traffic and Incoming Aircraft Type Sequence.

		SEQUENCE					
SUBJECT		1	2	3	4	5	6
1	EXPERT	II S2T2	III S4T4	I S6T3	IV S3T1	VI S1T5	V S5T6
2	STUDENT	III S3T4	IV S4T1	II S6T2	V S5T5	I S2T3	VI S1T6
3	STUDENT	IV S3T1	V S5T6	III S4T3	VI S1T5	II S6T4	I S2T2
4	EXPERT	V S5T5	VI S2T6	IV S4T4	I S6T3	III S3T2	II S1T1
5	STUDENT	VI S2T5	I S1T3	V S5T6	II S6T1	IV S3T4	III S4T2
6	EXPERT	IV S4T1	V S1T5	I S6T2	VI S2T6	III S3T3	II S5T4
7	STUDENT	I S2T1	IV S3T3	III S4T2	V S1T6	II S5T4	VI S6T5
8	STUDENT	V S6T6	VI S5T5	IV S3T2	II S1T4	I S2T1	III S4T3
9	EXPERT	II S6T4	III S4T2	VI S5T5	I S1T3	V S2T6	IV S3T1
10	EXPERT	III S3T2	I S5T1	II S2T4	IV S4T3	VI S6T6	V S1T5

Table A.2: Friedman's ANOVA Results for Performance Metric

	Proximity of Incoming Aircraft								Angle of Incoming Aircraft			
	Mean				Mean				Mean			
	χ^2 (1)	p	Scen I	Scen II	χ^2 (1)	p	Scen V	Scen VI	χ^2 (1)	p	Scen III	Scen IV
Number of Aircraft Delivered												
Student	2.000	0.157	25.20	26.4	0.333	0.564	25.40	25.20	0.333	0.564	25.80	26.40
Expert	0.333	0.564	27.60	27.2	0.333	0.564	27.20	27.00	4.000	<u>0.046</u>	27.00	27.80

	Number of Conflict											
Student	0.000	1.000	2.20	1.80	0.000	1.000	1.20	1.60	1.000	0.317	2.20	3.00
Expert	1.000	0.317	1.00	0.60	0.200	0.655	2.60	1.60	1.000	0.317	1.00	2.40
	RMS Exit Speed											
Student	4.000	<u>0.046</u>	179.69	180.23	1.000	0.317	180.28	180.00	2.000	0.157	180.58	180.28
Expert	0.000	1.000	180.78	181.37	1.000	0.317	180.27	181.75	0.333	0.564	180.89	180.94
	RMS Exit Heading											
Student	0.000	1.000	3.86	4.27	1.800	0.180	3.51	0.66	1.800	0.180	7.54	0.44
Expert	1.800	0.180	10.71	12.82	0.200	0.655	13.00	12.04	0.200	0.655	12.17	11.31
	RMS Exit Separation											
Student	0.200	0.655	10.21	9.87	0.200	0.655	10.15	10.12	0.200	0.655	10.02	9.99
Expert	1.800	0.180	9.24	9.51	0.200	0.655	9.40	9.62	5.000	<u>0.025</u>	9.49	9.14
	Number of Streams								Traffic Mix			
	Mean				Mean				Mean			
	χ^2 (1)	p	Scen II	Scen III	χ^2 (1)	p	Scen VI	Scen III	χ^2 (1)	p	Scen I	Scen V
	Number of Aircraft Delivered											
Student	0.200	0.655	26.40	25.80	0.200	0.655	25.20	25.80	0.333	0.564	25.2	25.40
Expert	0.333	0.564	27.20	27.00	0.000	1.000	27.00	27.00	0.333	0.564	27.6	27.20
	Number of Conflict											
Student	1.000	0.317	1.80	2.20	0.000	1.000	1.60	2.20	0.000	1.000	2.20	1.20
Expert	0.000	1.000	0.60	1.00	0.333	0.564	1.60	1.00	1.000	0.137	1.00	2.60
	RMS Exit Speed											
Student	1.000	0.317	180.23	180.58	1.000	0.317	180.00	180.58	2.000	0.157	179.70	180.28
Expert	1.000	0.317	181.37	180.89	0.333	0.564	181.75	180.89	2.000	0.157	180.78	180.27
	RMS Exit Heading											
Student	0.200	0.655	4.27	7.54	5.000	<u>0.025</u>	0.66	7.54	1.800	0.180	3.86	3.51
Expert	1.800	0.180	12.82	12.17	0.200	0.655	12.04	12.17	1.800	0.180	10.71	13.00
	RMS Exit Separation											
Student	0.200	0.655	9.87	10.02	0.200	0.655	10.12	10.02	0.200	0.655	10.207	10.15
Expert	0.200	0.655	9.51	9.49	0.200	0.655	9.62	9.49	0.200	0.655	9.247	9.37

References

- ¹ B. G. Hilburn, "Cognitive Complexity in Air Traffic Control – A Literature Review," EEC Note No. 04/04, EUROCONTROL, Bretigny-sur-Orge, France, 2004.
- ² M. Janic, "A Model of Air Traffic Control Sector Capacity Based on Air Traffic Controller Workload," Transportation Planning and Technology, Vol. 20, 1997, pp. 311–335.
- ³ A. Koros, P. Della Rocco, G. Panjwani, V. Ingurgio and J.-F. D'Arcy, "Complexity in Air Traffic Control Towers: A Field Study," Technical Note DOT/FAA/CT-TN03/14, NTIS, Springfield, Virginia, 2004.
- ⁴ B. G. Hilburn and P. G. A. M. Jorna, "Workload and Air Traffic Control", In PA Hancock and PA Desmond (Eds.). Stress, Workload and Fatigue: Theory, Research and Practice. Hillsdale, New Jersey, USA: Erlbaum, 2001.
- ⁵ H. G. Stassen, G. Johannssen and N. Moray, "Internal Representation, Internal Model, Human Performance Model and Mental Workload", Automatica, Vol. 26, No. 14, 1989, pp 811-820.
- ⁶ P. Kopardekar and S. Magyarits, "Dynamic Density: Measuring and Predicting Sector Complexity", Proceeding of the 21st Digital Avionics System Conference, Inst of Electrical and Electronics Engineers, Piscataway, NJ, 27-31 Oct. 2002, pp. 1-9.
- ⁷ A. J. Masalonis, M. B. Calaham and C. R. Wanke, "Dynamic Density and Complexity Metrics for Realtime Traffic Flow Management", The MITRE Corp. McLean, VA, 2003.
- ⁸ G. B. Chatterji and B. Sridhar, "Measures for Air Traffic Controller Workload Prediction", Proceedings of the 1st AIAA Aircraft Technology, Integration, and Operations Forum, Los Angeles, CA. Oct. 16-18, 2001. No AIAA 2001-5242
- ⁹ I. V. Laudeman, S.G. Shelden, R. Branstron and C.L. Brasil, "Dynamic Density: An Air Traffic Management Metric". NASA-

- TM-1998-112226, April, 1998.
- ¹⁰ B. Sridhar, K.S. Sheth and S. Grabbe, "Airspace Complexity and its Application in Air Traffic Management". 2nd USA/Europe Air Traffic Management 1998 R&D Seminar, Orlando, Florida.
 - ¹¹ P. Hermes, M. Mulder, M. M. van Paassen, J. H. L. Boering and H. Huisman, "Solution Space Based Analysis of the Difficulty of Aircraft Merging Tasks". November-December 2009. *Journal of Aircraft*. Vol. 46, No.6.
 - ¹² K. Lee, E. Feron, and A. R. Pritchett, A. R., "Describing Airspace Complexity: Airspace Response to Disturbances," *Journal of Guidance, Control, and Dynamics*, Vol. 32, No. 1, 2009, pp. 210–222.
 - ¹³ A. Majumdar, W.Y. Ochieng. Air Traffic Control Complexity and Safety: A Framework for Sector Design Based Upon Controller Interviews of Complexity Factors. *Transportation Research Records*. pp. 70-80.2007
 - ¹⁴ J. G. D'Engelbronner, M. Mulder, M. M. van Paassen, S. de Stigter and H. Huisman, "The Use of the Dynamic Solution Space to Assess Air Traffic Controller Workload". *Proceeding of AIAA Guidance, Navigation, and Control Conference*. Toronto, Ontario Canada. 2 - 5 August 2010. No. AIAA 2010-7542.
 - ¹⁵ G. A. Mercado-Velasco, M. Mulder, M. M. van Paassen, "Analysis of Air Traffic Controller Workload Reduction Based on the Solution Space for the Merging Task". *Proceeding of AIAA Guidance, Navigation, and Control Conference*, Toronto, Ontario Canada. 2 - 5 August 2010. No. AIAA 2010-7541.
 - ¹⁶ S. M. B. Abdul Rahman, M. Mulder, M. M. van Paassen, "The Effects of Air Traffic Control Sector Design on the Solution Space Diagram", *Proceedings of the 28th European Conference on Cognitive Ergonomics*, Delft, The Netherlands 25-27 August 2010, ISBN 978-94-90818-04-3.
 - ¹⁷ S. B. J. Van Dam, M. Mulder, M. M. van Paassen, "Ecological Interface Design of a Tactical Airborne Separation Assistance Tool". *IEEE, System, Man and Cybernetics – Part A: Systems and Humans*, Vol. 38, No. 6, November 2008.
 - ¹⁸ A. Tattersall and P. Foord, "An Experimental Evaluation of Instantaneous Self-Assessment as a Measure of Workload," *Ergonomics*, Vol. 39, No. 5, 1996, pp. 740–748.
 - ¹⁹ A. Field, *Discovering Statistics Using SPSS*, 3rd ed., Sage, London, 2009.