Visualizing complexities: the human limits of air traffic control

Nicolas Durand, Jean-Baptiste Gotteland, Nadine Matton

Noname manuscript No.
(will be inserted by the editor)

Visualizing Complexities: the Human Limits of Air Traffic Control

Nicolas Durand · **Jean-Baptiste Gotteland** ·
Nadine Matton

Received: date / Accepted: date

Abstract Air traffic management is organized into filters in order to prevent tactical controllers from dealing with complex conflicting situations. In this article, we describe an experiment showing that a dynamic conflict display could improve human performance on complex conflict situations. Specifically, we designed a display tool that represents the conflicting portions of aircraft trajectories and the evolution of the conflict zone when the user adds a maneuver to an aircraft. The tool allows the user to dynamically check the potential conflicting zones with the computer mouse before making a maneuver decision. We tested its utility on a population of forty students: twenty Air Traffic Controller (ATC) students at the end of their initial training and twenty engineering students with the same background but no ATC training. They had to solve conflicts involving two to five aircraft with a basic display and with the dynamic visualization tool. Results show that in easy situations (2 aircraft), performance is similar with both displays. However, as the complexity of the situations grows (from 3 to 5 aircraft), the dynamic visualization tool enables users to solve the conflicts more efficiently. Using the tool leads to fewer unsolved conflicts and shorter delays. No significant differences are found between the two test groups except for

N. Durand
ENAC, 7 av Ed Belin 31055 Toulouse
Tel.: +33-562259565
E-mail: nicolas.durand@enac.fr

J.B. Gotteland
ENAC, 7 av Ed Belin 31055 Toulouse
Tel.: +33-562259584
E-mail: jean-baptiste.gotteland@enac.fr

N. Matton
ENAC, 7 av Ed Belin 31055 Toulouse
Tel.: +33-562259618
E-mail: nadine.matton@enac.fr

delays: ATC students give maneuvers that generate less delays than engineering students. These results suggest that humans are better able to manage complex situations with the help of our conflict visualization tool.

Keywords Air Traffic Control · Conflict Detection · Conflict Resolution · Visualization Tool · Complexity

1 Introduction

Air Traffic Controllers deal with a variety of complex tasks. The latest progress in new technologies offers more opportunities to develop tools that can improve their efficiency. Previous research (Karikawa et al, 2013) has illustrated the cognitive strategies of air traffic controllers in real work situations. Karikawa et al (2014) have developed a visualization tool of en-route Air Traffic Control tasks to analyze the performance characteristics of controllers' strategies. More recently Edwards et al (2017) studied the interaction of situation awareness and workload for different levels of automation in a real environment as well.

In this article we deliberately focus on mental resources required by the combinatorial aspect of aircraft conflict resolution through a simplified experiment comparing complex conflict situations involving 2, 3, 4 or 5 aircraft with or without a new dynamic conflict display tool. We conduct the experiment on a population of 40 students: a cohort of 20 Air Traffic Control (ATC) students at the end of their initial training and a cohort of 20 engineering students without air traffic control training. All students are currently at the ENAC (Ecole Nationale de l'Aviation Civile).

The purpose of the proposed experiments is to show that humans are able to deal with conflicts involving two or three aircraft but reach their limits in solving more complex situations without any aid. This phenomenon could have strong connections with the idea of limited-capacity working memory. Indeed, several studies support the view that human working memory is composed at least of a visuospatial component and a central executive component (Baddeley and Logie, 1999). Both components are likeky limited in capacity. Solving an air traffic control situation involves temporary storage of sequences of aircraft locations as well as the processing of this information (mental projection of future positions). As a consequence, as the situation becomes more complex, one can suppose that the working memory capacity limit would impair performance on conflict solving. Moreover, findings in cognitive psychology have highlighted that visuospatial memory and reasoning are intrinsically related (Tabachneck-Schijf et al, 1997). Therefore, the use of an external visualization tool may help people to complement their internal mental representations and find a solution to the conflict situation (Tversky, 2005). We did not assess mental capacity of the participants separately from the ATC experiment because the combinatorial characteristic of the problem cannot be addressed with human mental storage capacities when the problem involves 4 aircraft or more. Indeed, we show later in the introduction that for only 3 aircraft, a human would have to test up to 8 different trajectory combinations, but for 4 aircraft the number of combinations reaches 64 and 1024 for 5 aircraft, which is far more than process a human brain. The experiments show that the initial two year training of ATC students has a low impact on their capacity to manage complex conflicts: when comparing the results with a cohort of engineering students who have the same age and scientific background without the ATC training experience, we do not see any significant difference, except for delays: Air Traffic Controller students tend to pay more attention to delays than engineering students.

To explain this result, we suggest that the complexity of conflicting situations involving 4 or more aircraft is too high for humans to deal with. We show that a simple tool could help simplify complex situations. We call it the Controller Assistance Tool (CAT). It is meant to help the controller manage complex situations without giving the optimal solution. We use a simplified environment in order to isolate the intrinsic complexity of conflict resolution. We use a single flight level (altitude) because controllers work on 2D displays. We only explore horizontal maneuver options on randomly generated clusters of different sizes. We check how a tool capable of showing parts of potentially conflicting trajectories can help humans separate aircraft for different levels of complexity.

1.1 Background

Air traffic controllers' en-route tools have not evolved much over the last few decades despite the exponential growth of computing power. Whereas on board systems fully rely on automation, en-route air traffic controllers are still dealing with complex situations with simple tools. Conflict detection and resolution require much of their mental resources. In order to provide a tolerable level of complexity to air traffic controllers, the current Air Traffic Management system is divided into layers or filters, each with a decreasing time horizon. Each layer is meant to reduce the complexity of the next one. There are four major layers:

1. Strategic (several months), ASM (Air Space Management): design of routes, sectors and procedures
2. (Pre-)Tactical (a few days to a few hours), ATFM (Air Traffic Flow Management): control centers open schedules and define hourly capacities of each open sectors (or groups of sectors). To respect these capacity constraints, the NMOC (Network Manager Operations Center) computes and updates flow regulations and reroutings according to the posted flight plans and resulting workload excess.
3. Real time (5/10 minutes), tactical control: surveillance, coordination with adjacent centers, conflict resolution by various simple maneuvers (heading, flight level, speed) transmitted to the pilots.
4. Emergency (less than 5 minutes), safety nets: ground-based (Short Term Conflict Alert, Minimum Safety Altitude Warning) and airborne (Traffic Alert and Collision Avoidance System, Ground Proximity Warning System).

In France, air traffic controllers are trained to detect conflicts using a 2D horizontal visualization of the traffic. Aircraft are represented by plots. Past positions of the aircraft are represented by a comet. The speed vector is materialized by a line segment representing x minutes of flight. This line segment helps the air traffic controller project future positions of the aircraft to detect potential conflicts. Controllers can also manually measure distances between points to check minimum separations, but this information is not automatically shown on the screen.

There have been many attempts to organize the controller's work or assist the conflict detection task. Uncertainty management plays a key role in air traffic control. Corver and Grote (2016) recently proposed a field study exploring sources of uncertainties and management strategies adopted by controllers. Some of the first research experiments on controller assistance tools were carried out in the 1990s. In Europe, HIPS (Meckiff and Gibbs, 1994; Price and Meckiff, 1997) the Highly Interactive Problem Solver was issued from ARC2000 (Dean et al, 1995). HIPS offered a representation of the conflict zones, called no-go-zones, in an interactive way for a chosen aircraft, knowing the intent of the other aircraft and taking into account uncertainties. It was tested on diverse scenarios with real controllers. It was tested in a Free Flight environment (Duong et al, 1997) and on Oceanic traffic management (Price and Meckiff, 1997). It did not offer any automatic solution using combined maneuvers but was meant to dynamically show the evolution of conflict zones. The major drawback of the no-go-zones is that they did not give any useful information on how to return to the original route and did not check conflicts on this part of the trajectory. Uncertainty was taken into account and an optimal maneuver time could thus be defined, but the uncertainty modeling was not detailed in the articles. Bakker and Blom (2000) compare different conflict prediction models taking into account uncertainties. They explain that uncertainty model adopted by HIPS uses geometric approach: Aircraft are modeled by ellipses and the conflict predictor compares the distance between ellipses and the separation standard. The size of the ellipse grows with time in the speed direction. This models uncertainties on aircraft speeds. Bakker and Blom compare on different scenarios the result of the geometric conflict predictor and the probabilistic model used by Erzberger (1997) and later adopted by Arthur and McLaughlin (1998) in the American project URET (User Request Evaluation Tool). They also introduce a third conflict predictor based on a collision risk approach. The challenge of any conflict predictor is to detect every conflict without overestimating potential conflict that will not lead to effective separation loss.

In the US, Erzberger (1997) introduced a conflict predictor in the 90s that was then used by Arthur and McLaughlin (1998) in URET. The model is much more technically advanced than the geometric approach. It can display conflict probabilities in complex situations. The conflict probe was used by Prevot et al (2005) as a tool to assist controllers in the conflict detection task. Prevot et al (2008) describe complex experiments done in 2008 to check how new displays of conflicts and an interactive conflict solver can help controllers deal with 3 times the current traffic. In (Prevot et al, 2011), bad weather conditions and time constraints are added to check the robustness of the automated solver tool. More recently, Borst et al (2017) proposed a solution space diagram for conflict detection and resolution in the horizontal plane. It shows the speed vector and different forbidden zones generated by other aircraft and offers interaction to the user.

None of the cited approaches have analyzed the combinatorial aspect of conflict resolution. NASA has built complex experiments that show that new technologies can increase the manageable traffic density. However, in all these experiments it is hard to isolate which factor contributes to major improvements.

Much research has been done on automatic conflict resolution (Durand et al, 1996; Oh et al, 1997; Frazzoli et al, 2001; Pallottino et al, 2001, 2002; Christodoulou and Kontogeorgou, 2008; Alaeddini et al, 2006; Gariel and Feron, 2009; Allignol et al, 2013). Complex conflict situations can now be handled by automatic solvers. Hypotheses used by researchers to model the trajectories and uncertainties are generally not realistic enough to imagine a usable application, but even in the most realistic models using simple maneuvers (Durand et al, 1996; Allignol et al, 2013), the solutions found by automatic solvers on complex situations are hard to understand and cannot be easily used as an aid for air traffic controllers in real time situations. Alaeddini et al (2006) proposed a conflict solver that is close to Air Traffic Controllers practices but his algorithm uses a sequential resolution model. Sequential resolutions models are most of the time less efficient on complex situations than global approaches.

1.2 Conflict Resolution Complexity

We suggest humans cannot deal with complex conflicts without assistance because the problem is combinatorial. Durand (1996) explained how difficult the mathematical problem to solve a cluster of conflicting aircraft is (Durand et al, 1996). Even if trajectory minimization can be modeled by a convex function, the search space in the horizontal plane is divided into many components which would each require a local search for a solution. In the horizontal plane, for a conflict involving two aircraft, if trajectories do not loop, the set of conflict free trajectories has two separated components. For a conflict involving n aircraft there may be $2^{\frac{n(n-1)}{2}}$ different components in the free trajectory space which strongly suggests that any method that requires exploring every connected component is NP-difficult. It is important to note that this complexity is independent from the modeling chosen. The problem is that even for small values of n, the number of potential different solutions in the horizontal plane grows exponentially. For $n = 4$ aircraft there are already 64 potential options, and for $n = 5$ the number rises to 1024. Durand and Granger (2003) statistically studied the complexity of actual air traffic data with cluster analysis and showed that 2 and 3 aircraft clusters are the majority, but 5 and more aircraft clusters are not rare in current densities of traffic if no filter is applied. This means that as traffic becomes denser, air traffic controllers will likely face the issue of solving more complex clusters. The number of conflict combinations for n aircraft clusters is the number of graphic sequences and it also grows exponentially with the number of aircraft involved. For example, with 2 aircraft, there is only one possible graphic sequence, representing a conflict between the two aircraft whereas for 5 aircraft, there are 20 different graphic sequences (see figure1) representing 20 different possible structures of conflicts inside the cluster. Every vertex represent an aircraft and each edge is a conflict. The graphical sequence also gives the minimum number of maneuvers necessary to solve a conflict.

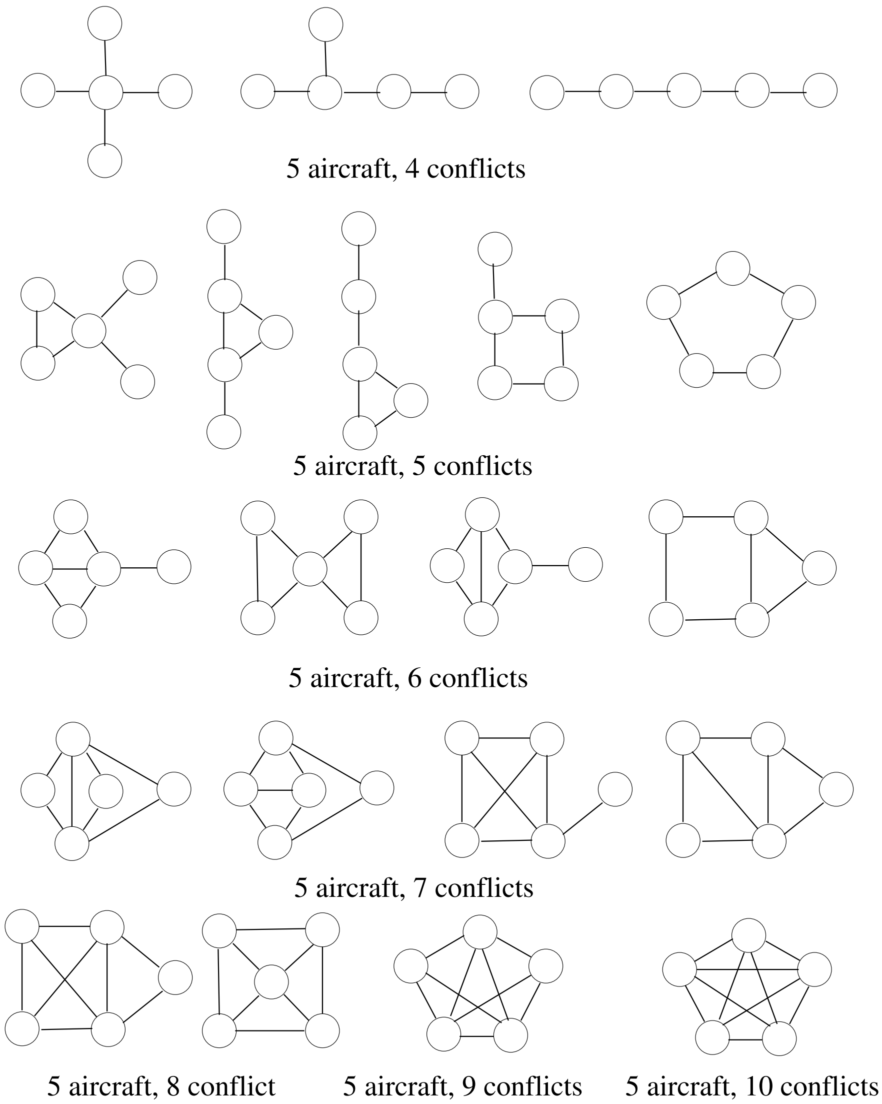

Fig. 1 Different structures for 5 aircraft clusters. Each vertice represents an aircraft, each edge is a conflict between two aircraft.

1.3 Outline

In part 2, we describe the tool used in our experiments to ease conflict resolution. We explain how it was designed, and how it works and detail the experiment. Part 3 analyses the results. We conclude and suggest further avenues of inquiry.

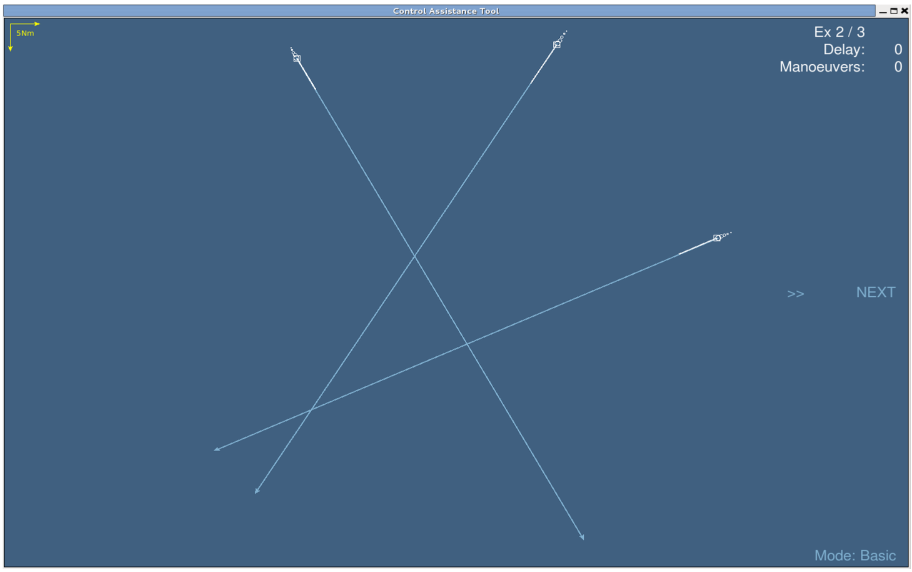

Fig. 2 3 aircraft conflict in the *Basic* mode.

2 Experiment Description

2.1 Tool description

The tool is rather simple and intuitive. We use a 2D-display of the traffic. Aircraft are represented by plots. A comet shows the past position and the speed vector is represented by a segment showing the future position of the aircraft in one minute. The five nautical miles separation standard is displayed on the top left of the screen.

The trajectory track is represented as a light line on which the plot representing the aircraft moves. In the experiments, we compare two modes of conflict display. In the *Basic* mode, no information of the conflict location is given on the trajectory track. Figure 2 gives an example of a 2 aircraft conflict in the *Basic* display mode. In the *Dynamic* display mode, the part of the trajectory of each aircraft in conflict is represented by black segments. Figure 3 shows the same 2 aircraft example in the *Dynamic* mode.

For the experiments, the double white arrow allows the user to move forward in time and the *NEXT* button moves to the next exercise.

In both modes, the user can modify the trajectory with the cursor. When the user positions the cursor over a trajectory line, it turns blue. The user can select a point of the trajectory by holding the left click. While holding the left click the user can move the trajectory line and add a maneuver when releasing it. A right click during the move cancels the maneuver. Moving the mouse wheel advances the time just like clicking the double arrow. The number of maneuver actions and the delay caused by maneuvers are displayed on the top right of the screen.

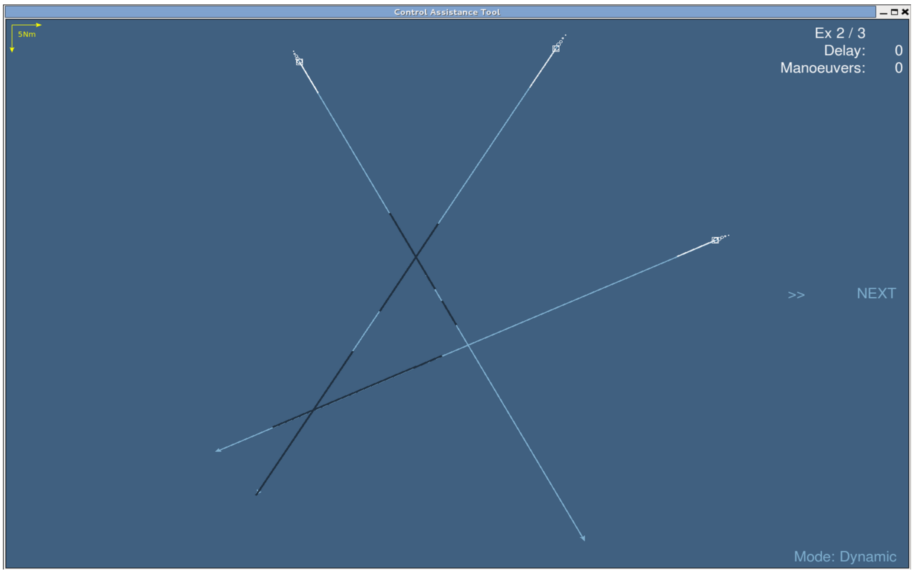

Fig. 3 3 aircraft conflict in the *Dynamic* mode.

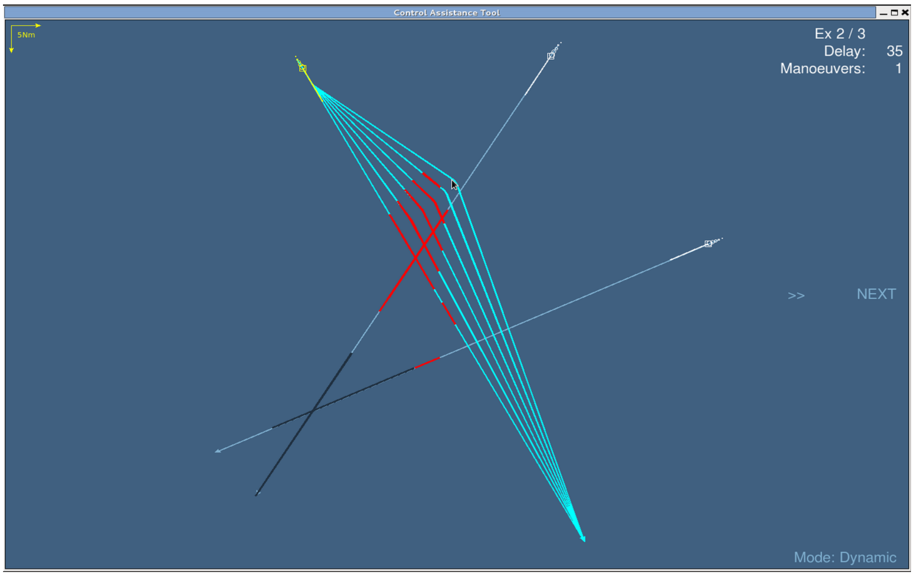

Fig. 4 Mouse interactions in the *Dynamic* mode.

In the *Dynamic* mode, conflicting parts of trajectories are represented in black. Because of uncertainties, these parts can shrink or vanish with time. When the cursor selects a trajectory (blue), the conflicting parts involving this trajectory turn red. When selecting a point of the trajectory and moving it, the user can dynamically see the evolution of the conflict zone (see figure 4).

The experiment starts with a short text explaining the display and the mouse functions. It explains that aircraft are represented by plots, comets show the past positions and the speed vectors give the expected positions within one minute. It reminds that the separation standard is on the top left of the screen. The mouse functions are detailed as follows:

- By clicking and holding the left button, the user can catch and move a point of the trajectory and create a maneuver when releasing the button;
- While moving a point, a right click cancels the move;
- Moving the mouse wheel advances the time;
- The Next button moves to the next exercise.

Before the set of *Basic* mode exercises we give one page of guidelines to inform the participants that conflict zones are not represented. We remind the priorities: they must first solve conflicts, second minimize the number of maneuvers, and third limit the delay. The first exercise is a training exercise.

Before the set of *Dynamic* mode exercises we give one page of guidelines to inform the participants that conflicting parts of trajectories are represented in black. We inform them that these part can shrink or vanish with time because of uncertainty. When the cursor selects a trajectory, it turns blue and the conflicting parts involving this trajectory turn red. We also remind the priorities: they must first solve conflicts, second minimize the number of maneuvers, and third limit the delay. The first exercise is a training exercise.

2.2 Conflict Detection Calculation

Each trajectory is divided in line segments on which the aircraft has a constant speed. We use a standard rate (3 degrees per second) to model turns and approximate turns with segments. The angle between two consecutive segments cannot exceed 10 degrees in order to keep the trajectory display smooth. To measure the separation between two trajectories, we add points on both trajectories in order to synchronize the lists of segments. Once segments are synchronized, the problem is reduced to checking the distance between aircraft 1 and aircraft 2 flying at constant speed on segment A_1B_1 and segment A_2B_2 (see figure 5). The distance is the length of vector $\mathbf{X_r} = t\,\mathbf{A_r} + (1-t)\,\mathbf{B_r}\,(t \in [0,\,1])$ and must be compared to the separation standard d. We must solve the equation $\mathbf{X_r}^2 - d^2 = 0$. This is an easy quadratic equation with either no solution, a double solution t_{double} or two solutions $t_A < t_B$:

- If the equation has no solution in $[0,1]$ then either the whole segments are in conflict or the whole segments are separated, depending on the length of A_1A_2;
- If the equation has a double solution, the whole segments are separated;
- If $t_A < 0$ and $t_B > 1$ then the whole segments are in conflict;
- If $t_A \geq 1$ or $t_B \leq 0$ then the whole segments are separated;
- If $0 < t_A < 1 < t_B$ then X_{A1} and X_{A2} are the positions of aircraft at time t_A. A_1X_{A1} and A_2X_{A2} are separated and $X_{A1}B_1$ and $X_{A2}B_2$ are in conflict;
- If $t_A < 0 < t_B < 1$ then X_{B1} and X_{B2} are the positions of aircraft at time t_B. A_1X_{B1} and A_2X_{B2} are in conflict and $X_{B1}B_1$ and $X_{B2}B_2$ are separated;

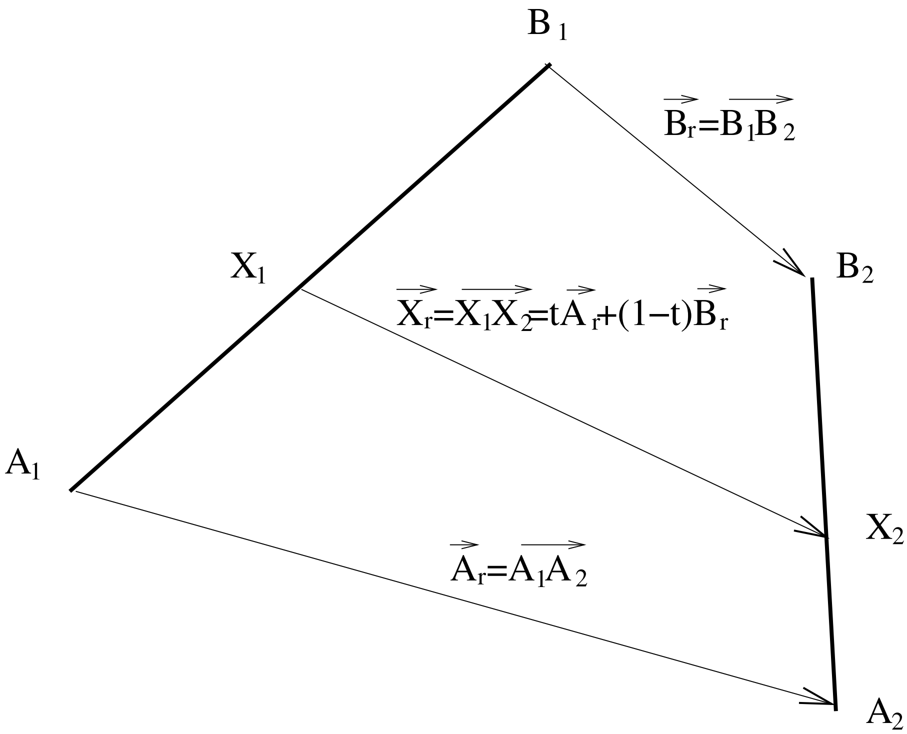

Fig. 5 Segment separation calculation.

– If $0 < t_A < t_B < 1$ then X_{A1} and X_{A2} are the positions of aircraft at time t_A and X_{B1} and X_{B2} are the positions of aircraft at time t_B. A_1X_{B1} and A_2X_{B2} are separated, $X_{A1}X_{B1}$ and $X_{A2}X_{B2}$ are in conflict and $X_{B1}B_1$ and $X_{B2}B_2$ are separated.

We add some uncertainty in the model by increasing the separation standard linearly with time. We replace d with $d_0 + t\,s_d$ in the previous equation. We keep a quadratic equation but in some rare cases the aircraft are separated in the segment between the roots. This depends on the sign of $a = \|\mathbf{A_r} - \mathbf{B_r}\|^2 - s_d^2$. When $a >= 0$ then the equation can be treated like previously whereas when $a < 0$ the segments are in conflict outside the roots. In the experiments, the initial separation standard is 5 nautical miles and it increases by 5% of the mean of the aircraft speeds. This means that the uncertainty model can handle a 5% error on the speed and a 2.8 degree error on the aircraft track.

Using such a growing norm is very convenient for quickly calculating the conflicting zone, but it can only model an isotropic growth of uncertainty. This is the main problem of such a model. We are currently working on improving the uncertainty model to take into account real uncertainties, such as the maneuver execution time, which depends on the pilot reaction, the speed uncertainty, or heading change uncertainty.

In real life, controllers often wait until the conflict is certain before maneuvering an aircraft. This simple isotropic model is not completely realistic but it is able

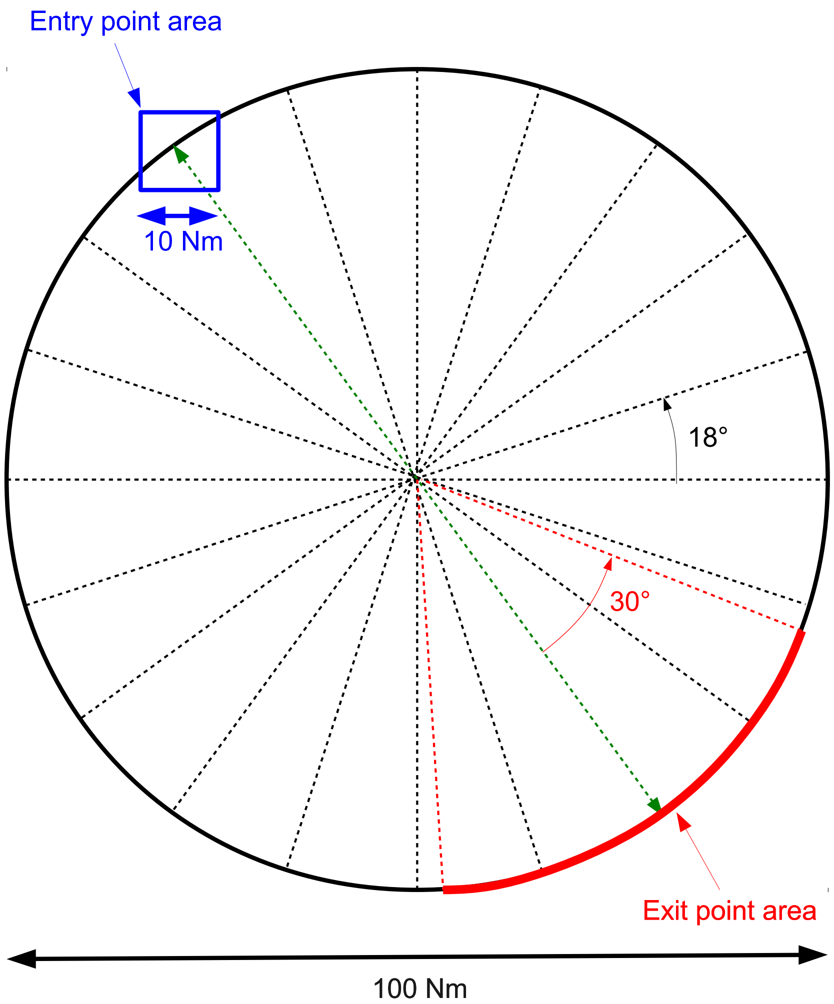

Fig. 6 Generation of random traffic situations.

to model the interest of waiting before acting by showing conflicting zones that can either shrink with time if the conflict disappears or remain if not. This aspect is essential to build experiments where predicted conflicts integrate uncertainties on the trajectory prediction.

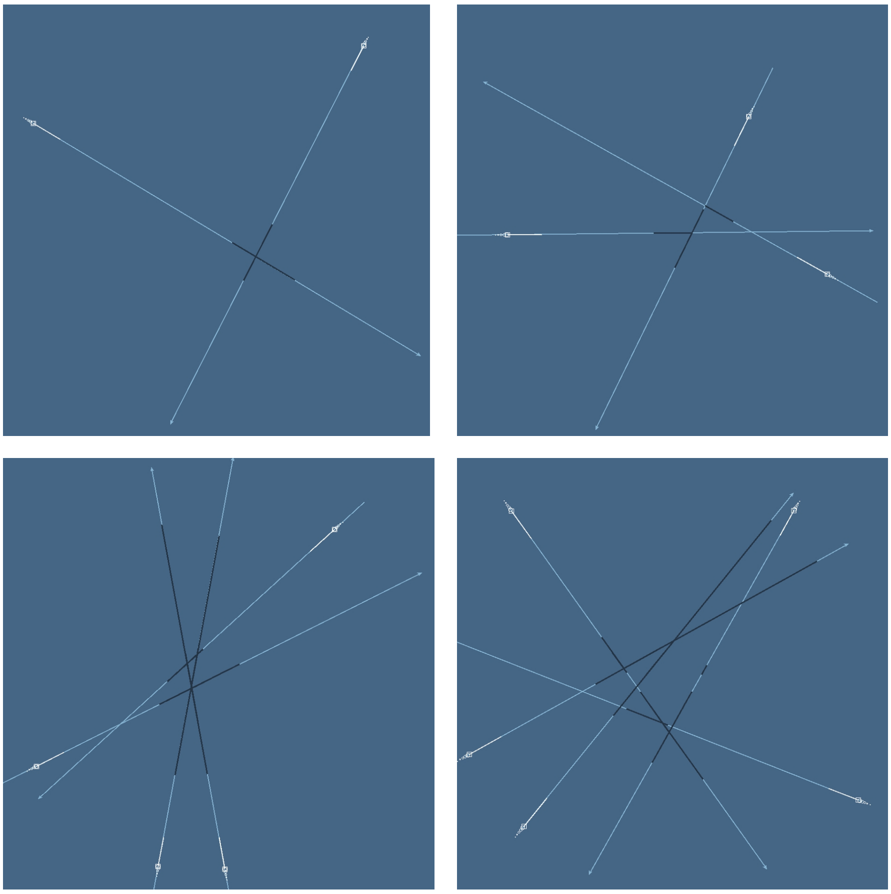

Fig. 7 Examples of traffic situations, with 2, 3, 4 and 5 aircraft.

2.3 Exercise generation

In order to generate some random traffic situations with different types of conflicts between aircraft, we consider a circular sector with a diameter of 100 nautical miles (about 15 minutes of flying time for an aircraft), with 20 possible entry points regularly positioned on its circumference (see figure 6). With these orders of magnitude, the distance between two neighboring entry points is over 15 nautical miles (which is three times greater than the minimal separation distance between aircraft).

The number of aircraft in the traffic situations vary from 2 to 5. Each aircraft is randomly assigned:

- a nominal speed, between 370 and 550 knots;
- its own entry point, in a rectangular area of 10 nautical miles around one of the sector's entry points;

- an exit point on the opposite side of the sector, in a slice extending by plus or minus 30 degrees around the opposite point on the circle.

Initially, each aircraft flies directly from its entry point to its exit point.

In order to avoid unmanageable traffic situations, the following constraints were required additionally (situations not respecting these constraints were discarded):

- A minimal duration of 3 minutes was required before the first conflict happens.
- A conflict solver using a genetic algorithm, as described by Durand et al (1996), was run to check that a solution exists using some simple maneuvers defined by three values per aircraft (t_0, t_1, α): the aircraft turns α degrees at t_0 and resumes its course at t_1, with the following ranges:
  - $\alpha \in [-40°, 40°]$
  - $t_0 \in [0, 10]$ (minutes)
  - $t_1 \in [t_0, 10]$ (minutes)

Figure 7 gives some examples of such traffic situations, displayed in the *Dynamic* mode.

Using this process, two series of 16 exercises were generated for the experiment: each series is intended to be run in a different mode (*Basic* or *Dynamic*) by each group of participants, and contains in the following order:

1 training situation	with 2 aircraft
3 situations	with 2 aircraft
4 situations	with 3 aircraft
4 situations	with 4 aircraft
4 situations	with 5 aircraft

2.4 Participants

We conducted the experiment with two categories of ENAC students:

- 20 Air traffic control (ATC) students (ICNA), that were at the end their training;
- 20 engineering students (IENAC), in their second school year (of three) with light air traffic control background.

All the participants were between 21 and 24 year old. There were 14 males and 6 females in the ATC student cohort and 15 males and 5 females in the engineering student cohort.

Each category of students was divided into two groups, which ran the experiment at the same time in the same computer room (see figure 8):

- The first group did the first series of exercises in the *Basic* mode and the second series in the *Dynamic* mode
- The second group did the first series of exercises in the *Dynamic* mode and the second series in the *Basic* mode

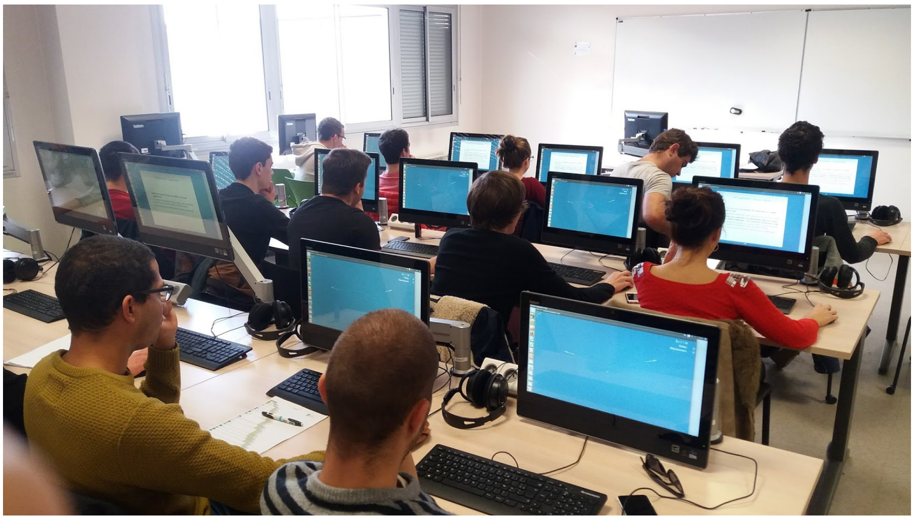

Fig. 8 Experiment with the ATC students.

Exercises were randomized for the same number of aircraft but we kept an increasing number of aircraft. There may be a learning effect over time, but it would affect both series of exercises and does not prevent us from concluding on complexity issues.

There was no time constraint. The whole exercise lasted from 30 to 45 minutes. Before the experiment, each participant was asked to read and sign an agreement form, allowing us to use the results collected anonymously.

2.5 Measures

During the experiment, each participant tries to solve the different traffic situations one by one, by modifying some aircraft trajectories. For each exercise, the following information is recorded:

- All the mouse actions and the resulting aircraft trajectories, with two time indications:
  - the relative time at which it happened in real life (counted from the beginning of the exercise);
  - and the corresponding relative time in the simulation (which is often different as the participant moves forward in time during the exercise).
- The number of aircraft pairs for which some conflicts have not been solved at the end of the exercise (i.e. for which the minimal separation distance of 5 nautical miles was not ensured during some periods).
- The number of modified aircraft trajectories.
- The cumulative delay (in seconds) that has been generated by the different maneuvers.
- The time spent handling each exercise (in real life).

3 Results

3.1 Presentation of results

In the following sections, we analyze five parameters:

1. The number of unsolved conflicts;
2. The number of maneuvers used to solve conflicts;
3. The delays of maneuvered aircraft.

These three parameters give a measure of the user performance in the different modes.

4. The number of mouse actions;
5. The time spent on the exercises.

These two parameters give information about the way users interact with the tool.

The results are gathered by number of aircraft and by experiment mode (*Basic* or *Dynamic*): each value is the average across participants and across exercises, for a given number of aircraft and a given experiment mode. On the figures, the error bars show the 95% confidence intervals of the means.

Regarding the delay, and the number of mouse actions, the figures always include all the exercises, even the ones that were not solved. We include them because we observed that we obtain the same trends and same order of magnitude when we only take into account the solved exercises.

For each of the five parameters, we assessed the potential effects of the group of students, the difficulty of the exercise, i.e., the number of aircraft, the experiment mode and their interaction. The residuals of a linear model of these factors were not normally distributed for any of the five parameters ($p < .001$ for all Shapiro tests). Therefore, for each of the five parameters, we could not analyze the data with an ANOVA (analysis of variance) and fitted the data to a generalized least square (GLS) model, using the *rms* package of R (version 3.1.1). This GLS model analyzed the effect of following variables and their interactions: Group (ATCO vs. Engineering students), Number of aircraft (2, 3, 4 or 5) and Mode (Basic vs. Dynamic).

3.2 Unsolved conflicts

The first goal of the experiment was to solve the conflicts: figure 9 shows the number of unsolved conflicts, measured by the number of aircraft pairs that still have a conflict in the proposed solution.

After fitting a GLS model, we can state that the Group had no effect ($t = 0.9$, $p = .38$) on the number of unsolved conflicts. However, the Mode ($t = 2.8$, $p = .005$), the Number of aircraft ($t = 10.9$, $p < .001$) and the interaction Mode x Number of aircraft ($t = 5.3$, $p < .001$) significantly contributed to explaining variance in unsolved conflicts. More precisely:

– with two aircraft, there is no difference between both modes, indeed the predicted 95% confidence intervals of the number of unsolved conflicts were $[-0.08; 0.15]$ for the *Basic* and $[-0.13; 0.11]$ for the *Dynamic* mode.

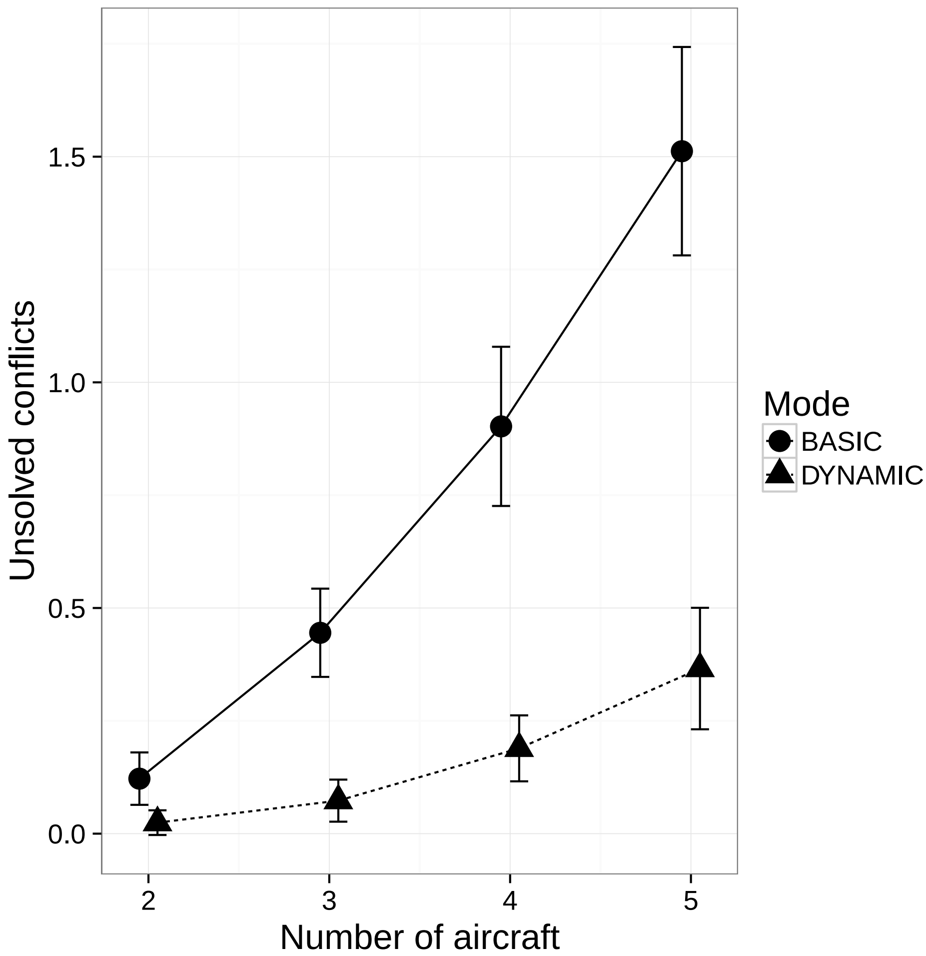

Fig. 9 Unsolved conflicts for different cluster sizes. Error bars represent 95 confidence intervals.

– from three to five aircraft, there are significantly more unsolved conflicts in the *Basic* mode than the *Dynamic* mode and the unsolved parameter increases linearly for both conditions as a function of the number of aircraft. The largest difference is then observed for five aircraft which is the largest amount tested: estimated 95% confidence intervals are $[1.34; 1.55]$ for the *Basic* mode and $[0.22; 0.44]$ for the *Dynamic* mode.

These results provide the main contribution of this experiment: for both air traffic control and engineering students, we observe far fewer unsolved conflicts in the *Dynamic* mode and the difference between the two modes increases with the number of aircraft involved in the clusters. Most of the exercises with two aircraft were solved but in the *Basic* mode, the number of unsolved conflicts rapidly increased with the number of aircraft regardless of the user profile.

Table 1 Means and standard deviations of the dependent variables according to the groups (ATCO vs. Engineering students), mode (Basic vs. Dynamic) and number of aircraft (from 2 to 5).

Variables			Unsolved conflicts		Maneuvers		Delay		Actions		Time	
Mode	Group	Aircraft	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Basic	ATCO	2	0,10	0,30	1,25	0,57	33,35	29,63	1,85	1,36	24,03	11,12
Dynamic	ATCO	2	0,03	0,18	0,95	0,77	24,33	31,55	1,63	2,05	22,48	13,00
Basic	Engineer	2	0,14	0,35	0,98	0,52	39,81	50,73	1,65	1,38	20,37	10,70
Dynamic	Engineer	2	0,02	0,13	1,02	0,61	36,76	48,91	1,63	1,73	19,97	11,69
Basic	ATCO	3	0,36	0,62	2,08	0,78	87,08	99,00	3,08	2,02	37,90	23,85
Dynamic	ATCO	3	0,08	0,31	2,28	0,69	74,59	66,90	4,46	3,53	49,15	24,59
Basic	Engineer	3	0,52	0,65	1,89	0,84	108,94	118,79	2,70	1,66	29,51	12,60
Dynamic	Engineer	3	0,07	0,30	2,11	0,69	122,21	110,61	4,00	3,26	37,63	26,39
Basic	ATCO	4	0,81	1,07	2,85	1,06	129,68	108,58	4,39	2,43	49,74	23,18
Dynamic	ATCO	4	0,21	0,54	2,76	1,21	98,89	100,14	5,79	5,19	60,69	42,06
Basic	Engineer	4	0,99	1,23	2,73	0,91	155,07	136,65	3,96	2,27	42,32	20,53
Dynamic	Engineer	4	0,17	0,41	2,64	1,15	168,67	185,22	4,37	3,54	40,48	34,61
Basic	ATCO	5	1,54	1,53	3,58	0,92	193,50	113,05	5,86	2,32	57,04	27,09
Dynamic	ATCO	5	0,49	1,01	3,36	1,23	141,14	103,39	7,69	6,55	75,58	52,58
Basic	Engineer	5	1,49	1,50	3,23	1,10	226,61	183,43	4,77	2,36	49,73	26,22
Dynamic	Engineer	5	0,25	0,73	3,70	1,15	295,95	270,11	7,14	4,50	56,10	34,56

These results confirm that making the conflicting portions of trajectories dynamically visible while modifying one of them provides significant assistance in solving the conflicts, even though it has to be done step by step.

These results also confirm the difficulty for humans to find solutions for conflicts involving many aircraft, even on problems for which the genetic algorithm solver finds instantly a simple solution involving no more than one maneuver per aircraft.

3.3 Number of maneuvers

The guidelines of the experiment also asked participants to minimize the number of maneuvers (after solving the conflicts).

In the recorded data, one way to assess the number of maneuvers is to count the number of aircraft trajectories that were modified, while ignoring the number of mouse interactions that were actually needed, as shown in figure 10.

From the GLS model, we can state that the number of modified aircraft trajectories increases with the number of aircraft involved in the situation ($t = 15.4$ and $p < .001$). The other variables of the model did not contribute significantly to explain the variance of the number of modified aircraft. About $n-1$ aircraft trajectories are modified on average for a situation involving n aircraft. Indeed, in the *Basic* mode the estimated 95 confidence intervals were [1.1;1.3], [1.9;2.0], [2.6;2.8] and [3.3;3.6] for situations involving two, three, four and five aircraft, respectively. Confidence intervals for the *Dynamic* mode, were non significantly different from these. This can be explained by the fact that participants solve complex conflicts by dealing with one aircraft pair at a time. For each conflicting aircraft pair, participants typically moved only one of the two aircraft in order to avoid the other (and if possible other already maneuvered aircraft). At the end of this process, it looks like participants

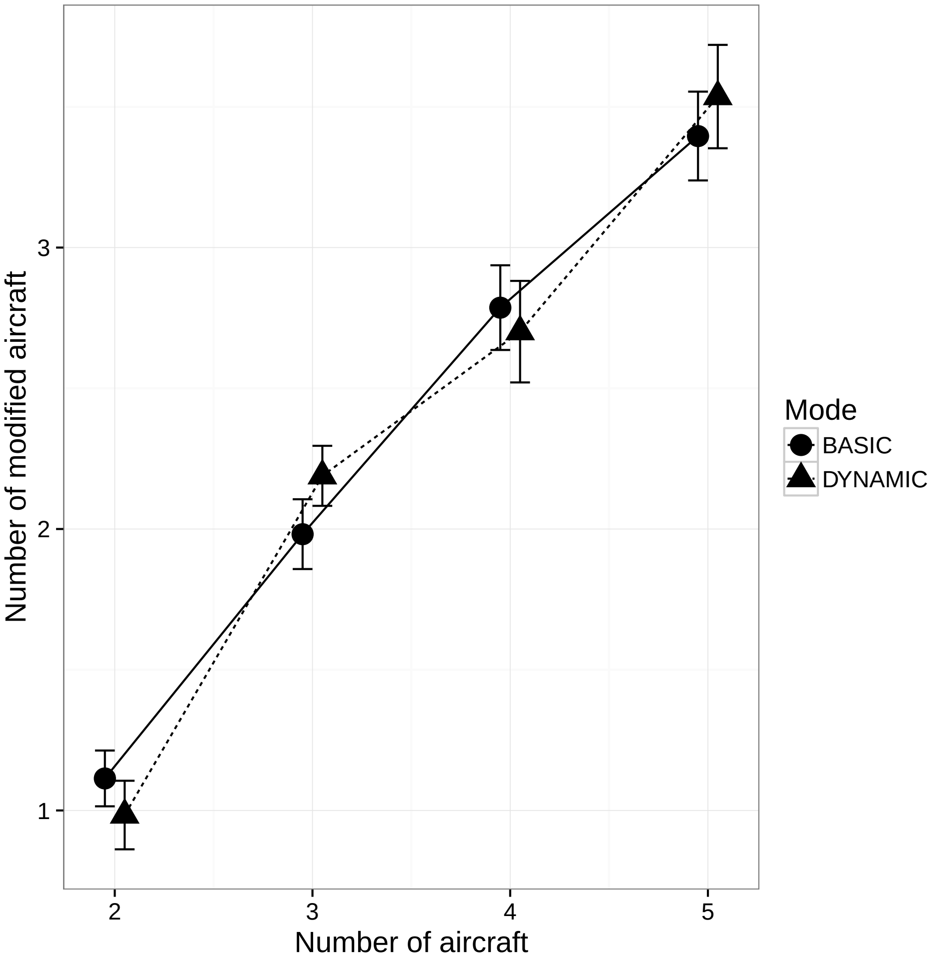

Fig. 10 Average number of modified aircraft trajectories. Error bars represent 95 confidence intervals.

sorted conflicts by priority, and in most cases, they left one aircraft trajectory completely unchanged (as if this aircraft had the highest priority).

3.4 Delay

The third guideline given to the participants was to limit the delay (after solving the conflicts and minimizing the number of maneuvers). Figure 11 shows how the total delay per exercise increases with the number of aircraft, in the two modes.

From the GLS model we can state that only two variables contributed significantly to explaining the variance of the delay: Number of aircraft ($t = 7.6$, $p < .001$) and the interaction Group x Mode x Number of aircraft ($t = 2.8$, $p = 0.006$). Indeed, the delay

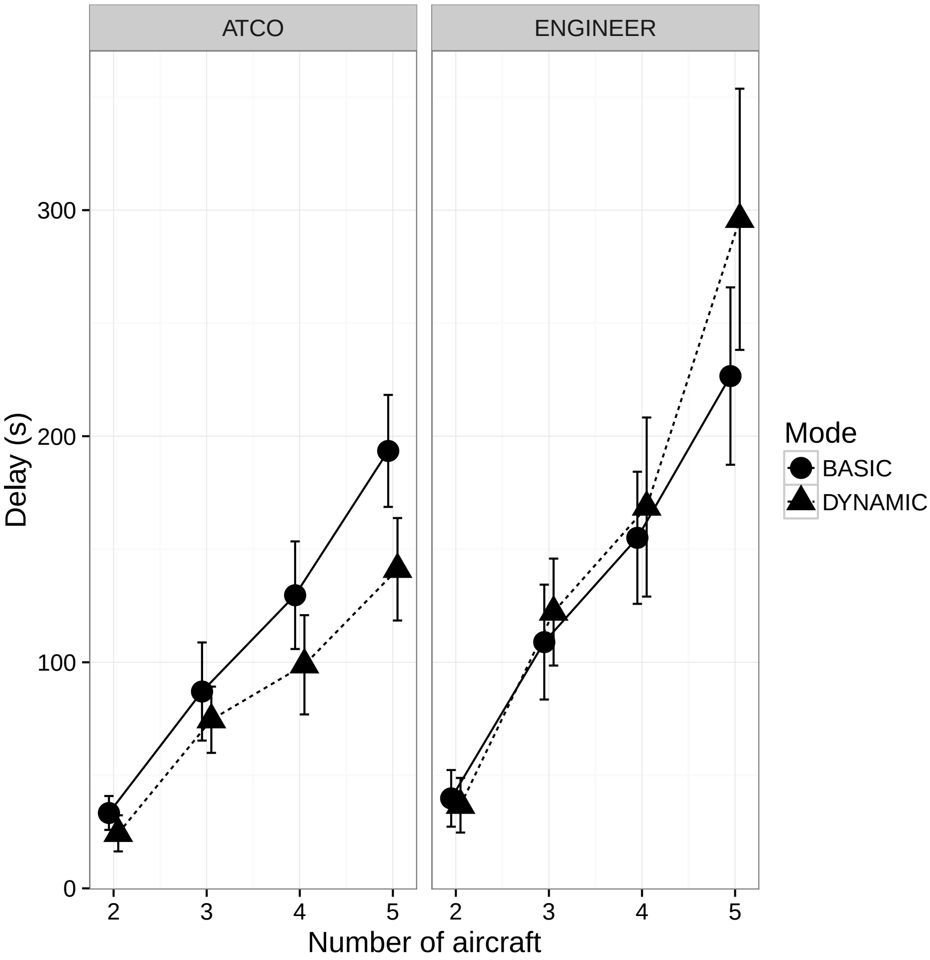

Fig. 11 Average delay (seconds). Error bars represent 95 confidence intervals.

globally increases with the complexity of the exercises. However, for Engineering students the delay is largest with five aircraft in the Dynamic mode ([256.3;303.2]) while on the contrary, for ATCO students the delay is largest with five aircraft in the Basic mode ([200.0;246.9]).

– In situations involving only two aircraft, there is no significant difference between the delays obtained with both modes (95 confidence intervals were [19.4;45.1] for the *Basic* mode and [16.6;42.2] for the *Dynamic* mode.
– In situations involving three to five aircraft, the delay is lower with the *Dynamic* mode than with the *Basic* mode. The largest difference is observed for situations involving five aircraft : [128.9;152.4] for the *Dynamic* mode and [177.6;201.2] for the *Basic* mode.

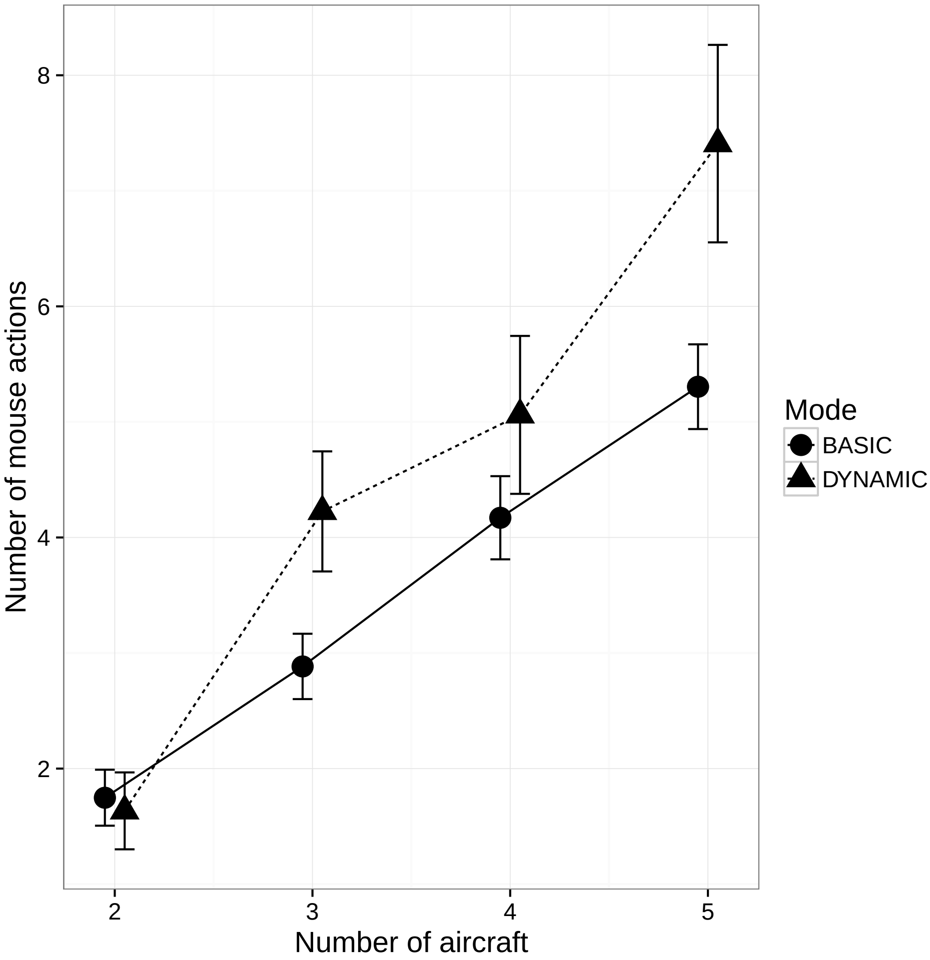

Fig. 12 Average number of mouse actions. Error bars represent 95 confidence intervals.

These results show that the *Dynamic* mode helps specifically trained users to find accurate maneuvers in the most complex situations.

3.5 Number of mouse actions

The number of mouse actions needed to handle the exercise gives an estimation of the number of maneuver adjustments. This value overestimates the effective number of maneuvers because a single and simple maneuver can be adjusted several times with different consecutive mouse actions. Furthermore, this value also provides interesting feedback on the complexity of the maneuver elaboration. We can observe in figure 12 that users tend to interact more with the mouse in the *Dynamic* mode.

When analyzing results of GLS modeling on this parameter, only two variables contributed significantly to explaining the variance: Number of aircraft ($t = 7.7, p < .001$) and the interaction Mode x Number of aircraft ($t = 2.3$, $p = .02$). We can state that:

- In situations involving only two aircraft, there is no significant difference between the number of mouse actions observed with both modes (95 confidence intervals were [1.3;2.2] for the *Basic* mode and [1.4;2.4] for the *Dynamic* mode.
- In situations involving three to five aircraft, the number of mouse actions is higher with the *Dynamic* mode than with the *Basic* mode. The largest difference is observed for situations involving five aircraft : [6.9;7.7] for the *Dynamic* mode and [4.9;5.7] for the *Basic* mode.

This result is at first surprising because the highlighting of the conflicting portions of the trajectories should theoretically help to define more efficiently all the needed maneuvers, resulting in fewer mouse actions. In fact, it appears that the participants were neither prompted nor trained to take advantage of this mode for this purpose (minimizing the number of mouse actions). On the contrary, the *Dynamic* mode encouraged them to test different options of maneuvers, in a ”what if” way of thinking.

Because the *Dynamic* mode gives instant feedback on the user’s action, the user may be more tempted to adjust the previous actions to reduce delays.

Future experiments with different guidelines and with participants trained in the *Dynamic* mode could help us confirm this first analysis.

3.6 Time spent handling the exercises

Similarly to trends for mouse actions, figure 13 shows that time spent to handle the exercises increased in the *Dynamic* mode, for both ATC and engineering students.

When analyzing results of GLS modeling on this parameter, only two variables contributed significantly to explaining the variance: Number of aircraft ($t = 7.4$, $p < .001$) and the interaction Mode x Number of aircraft ($t = 2.8$, $p = .006$). We can state that:

- In situations involving only two aircraft, there is no significant difference between the time spent to handle the exercises in both modes (95 confidence intervals were [19.0;27.1] for the *Basic* mode and [20.6;28.7] for the *Dynamic* mode.
- In situations involving three to five aircraft, the number of mouse actions is higher with the *Dynamic* mode than with the *Basic* mode. The largest difference is observed for situations involving five aircraft : [62.3;69.7] for the *Dynamic* mode and [50.9;58.3] for the *Basic* mode.

During the experiment, the participants did not have a time limit and had no reason to rapidly solve the exercises. They therefore took more time to explore the different solutions for each aircraft in the *Dynamic* mode, because they had more information about the transformations of the conflict areas.

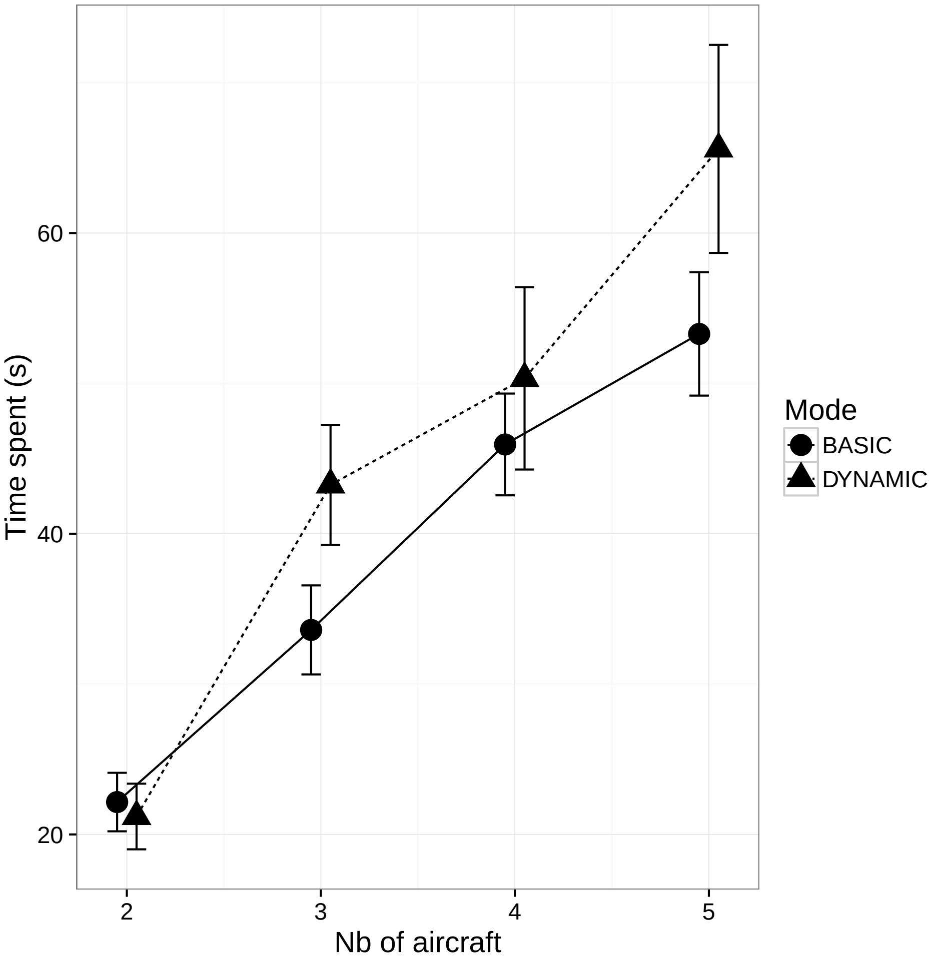

Fig. 13 Average time spent to handle the exercises (seconds). Error bars represent 95 confidence intervals.

4 Conclusion and Further Work

To conclude, we first confirmed with our experiments that the *Dynamic* mode helped ATC and engineering students deal with clusters involving more than two or three aircraft. In the *Basic* mode, they start having major issues when clusters involve 4 and 5 aircraft. We are aware that the lack of vertical maneuvers made the exercises more complicated for those cases. In the *Dynamic* mode, experiments show that the proposed tool helps the users deal with cluster complexity. There was no reduction of the number of maneuvers, but in terms of delay, we can also see the benefit of having an accurate tool to limit the size of maneuvers. It seems that the tool does not provide any clue to the user about how to reduce the number of actions. The reactive aspect of the tool led to more mouse manipulations. By analyzing one by one the answers, we found that the user tends to use the *Dynamic* mode like a “what if” mode to check options. This can also explain why the time spent solving the exercises is longer with the *Dynamic* mode. We did not ask participants to solve the problems as fast as possible and they did not get trained to use the tool.

There was no significant difference between ATC students and Engineering students except for delays: ATC students tend to give maneuvers which generate less delays than engineering students. It seems that most of the separation task skills are acquired once the ATC controllers are finished with their initial training at ENAC and when they move to their Air Traffic Control Centers.

This initial experiment was a first test to prove the potential usefulness of showing the conflicting parts of the trajectories to air traffic controllers. Other findings did not systematically highlight benefits of visualization tools in spatial tasks. For instance, providing participants with a computer visualization tool did not necessarily enhance performance of inferring and drawing cross sections of a three-dimensional object (Keehner et al, 2008). Thus, more research is needed to identify which component(s) of the visualization tool helped solve the most complex conflicts.

Furthermore, we can build on our results by taking the following steps:

– We first want to redo the test with qualified air traffic controllers in order to check if we find any difference with the present results.
– In a second phase, we need to add vertical maneuvers and interactions in the model. This requires adding altitude information to an aircraft tag and defining a simple interaction to maneuver the aircraft vertically. With this first improvement we can build scenarios including climbing and descending aircraft.
– We also need to improve the uncertainty model. First we want to make it compliant to realistic uncertainties. In addition to uncertainties due to speed prediction in both vertical and horizontal planes, we need to model the uncertainty caused by the pilot answer to orders and the track accuracy, especially when aircraft are following headings. We also want the controller to be able to adjust in real time the uncertainty parameters in order to comply to his/her own preferences.
– We finally want to test the tool presented on a real traffic environment with qualified controllers. To reach this goal we will connect our tool to a real simulator performing real data sets of traffic and perform experiments in a more realistic environment.

References

Alaeddini A, Erzberger H, Dunbar W (2006) Distributed logic-based conflict resolution of multiple aircraft in planar en-route flight. In: AIAA Guidance, Navigation and Control Conference and Exhibit

Allignol C, Barnier N, Durand N, Alliot JM (2013) A new framework for solving en-routes conflicts. In: 10th USA/Europe Air Traffic Management Research and Developpment Seminar

Arthur WC, McLaughlin MP (1998) User request evaluation tool (uret). interfacility conflict probe performance assessment. In: Procceedings of the 2nd USA/Europe ATM R and D Seminar

Baddeley AD, Logie RH (1999) Working memory: The multiple-component model. In: Miyake A, Shah P (eds) Models of Working Memory, Cambridge University Press, pp 28–61, DOI 10.1017/cbo9781139174909.005, URL https://doi.org/10.1017%2Fcbo9781139174909.005

Bakker G, Blom H (2000) Wp1: Comparative analysis of probabilistic conflict prediction approaches in atm. Tech. rep., NLR Contract Report

Borst C, Bijsterbosch VA, van Paassen MM, Mulder M (2017) Ecological interface design: supporting fault diagnosis of automated advice in a supervisory air traffic control task. Cognition, Technology & Work 19(4):545–560, DOI 10.1007/s10111-017-0438-y, URL https://doi.org/10.1007/s10111-017-0438-y

Christodoulou MA, Kontogeorgou C (2008) Collision avoidance in commercial aircraft free flight via neural networks and non-linear programming. Int J Neural Syst 18(5):371–387

Corver S, Grote G (2016) Uncertainty management in enroute air traffic control: a field study exploring controller strategies and requirements for automation. Cognition, Technology & Work 18(3):541–565, DOI 10.1007/s10111-016-0373-3, URL https://doi.org/10.1007/s10111-016-0373-3

Dean G, Fron X, Miller W, Nicolaon J (1995) Arc2000 : An investigation into the feasibility of automatic conflict. Tech. rep., Centre Exprimental Eurocontrol

Duong VN, Hoffman E, Nicolaon JP (1997) Initial results of investigation into autonomous aircraft concept (freer-1). In: Procceedings of the 1st USA/Europe ATM R and D Seminar

Durand N (1996) Optimisation de trajectoires pour la résolution de conflits en route. PhD thesis, INPT

Durand N, Granger G (2003) A traffic complexity approach through cluster analysis. In: 5th ATM R&D Seminar

Durand N, Alliot JM, Noailles J (1996) Automatic aircraft conflict resolution using genetic algorithms. In: Proceedings of the Symposium on Applied Computing, Philadelphia, ACM

Edwards T, Homola J, Mercer J, Claudatos L (2017) Multifactor interactions and the air traffic controller: the interaction of situation awareness and workload in association with automation. Cognition, Technology & Work 19(4):687–698, DOI 10.1007/s10111-017-0445-z, URL https://doi.org/10.1007/s10111-017-0445-z

Erzberger H (1997) Conflict probing and resolution in the presence of errors. In: Procceedings of the 1st USA/Europe ATM R and D Seminar

Frazzoli E, Mao ZH, Oh JH, Feron E (2001) Resolution of conflicts involving many aircraft via semidefinite programming. AIAA Journal of Guidance, Control and Dynamics 24(1)

Gariel M, Feron E (2009) 3d conflict avoidance under uncertainties. In: Digital Avionics Systems Conference, 2009. DASC '09. IEEE/AIAA 28th, pp 4.E.3–1–4.E.3–8, DOI 10.1109/DASC.2009.5347480

Karikawa D, Aoyama H, Takahashi M, Furuta K, Wakabayashi T, Kitamura M (2013) A visualization tool of en route air traffic control tasks for describing controller's proactive management of traffic situations. Cognition, Technology & Work 15(2):207–218, DOI 10.1007/s10111-012-0222-y, URL https://doi.

org/10.1007/s10111-012-0222-y
Karikawa D, Aoyama H, Takahashi M, Furuta K, Ishibashi A, Kitamura M (2014) Analysis of the performance characteristics of controllers' strategies in en route air traffic control tasks. Cognition, Technology & Work 16(3):389–403, DOI 10.1007/s10111-013-0268-5, URL https://doi.org/10.1007/s10111-013-0268-5
Keehner M, Hegarty M, Cohen C, Khooshabeh P, Montello D (2008) Spatial reasoning with external visualizations: What matters is what you see, not whether you interact. Cognitive Science: A Multidisciplinary Journal 32(7):1099–1132, DOI 10.1080/03640210801898177, URL https://doi.org/10.1080%2F03640210801898177
Meckiff DC, Gibbs DP (1994) PHARE : Highly interactive problem solver. Tech. rep., Eurocontrol
Oh JH, Shewchun J, Feron E (1997) Design and analysis of conflict resolution algorithms via positive semidefinite programming [aircraft conflict resolution]. In: Decision and Control, 1997., Proceedings of the 36th IEEE Conference on, vol 5, pp 4179–4185 vol.5, DOI 10.1109/CDC.1997.649489
Pallottino L, Bicchi A, Feron E (2001) Mixed integer programming for aircraft conflict resolution. In: AIAA Guidance Navigation and Control Conference and Exhibit
Pallottino L, Feron E, Bicchi A (2002) Conflict resolution problems for air traffic management systems solved with mixed integer programming. Intelligent Transportation Systems, IEEE Transactions on 3(1):3–11, DOI 10.1109/6979.994791
Prevot T, Lee P, Smith N, Palmer E (2005) Atc technologies for controller-managed and autonomous flight operations. In: AIAA Guidance, Navigation and Control Conference and Exhibit
Prevot T, Homola J, Mercer J (2008) Initial study of controller/automation integration for nextgen separation assurance. In: AIAA Guidance, Navigation and Control Conference and Exhibit
Prevot T, Homola J, Martin L, Mercer J (2011) Automated air traffic control operations with weather and time-constraints. In: 9th ATM R&D Seminar
Price A, Meckiff C (1997) Hips and its application to oceanic control. In: 1st ATM R&D Seminar
Tabachneck-Schijf HJ, Leonardo AM, Simon HA (1997) CaMeRa: A computational model of multiple representations. Cognitive Science 21(3):305–350, DOI 10.1207/s15516709cog2103_3, URL https://doi.org/10.1207%2Fs15516709cog2103_3
Tversky B (2005) Functional significance of visuospatial representations. In: Shah P, Miyake A (eds) The Cambridge Handbook of Visuospatial Thinking, Cambridge University Press, pp 1–34, DOI 10.1017/cbo9780511610448.002, URL https://doi.org/10.1017%2Fcbo9780511610448.002