2007

Testing a Maneuver Space-Based Collision Avoidance System: Experiment 1

William R. Knecht

TESTING A MANEUVER SPACE-BASED COLLISION AVOIDANCE SYSTEM: EXPERIMENT 1

William R. Knecht
Maneuver Space Technologies
Pocasset, OK

A new data visualization technology is demonstrated for en-route tactical aircraft separation maintenance. This technology displays *maneuver space* (MS), a nonveridical coordinate space based on the four key maneuver elements of heading, speed, altitude, and available maneuver time. In this first major test of a prototype 4D collision avoidance system (4CAS), eight licensed general aviation pilots each flew eight simulated free flight scenarios, with the goal of deviating as little as possible from a pre-assigned flight path, while still maintaining standard separation from traffic. Compared to a cockpit display of traffic information (CDTI), the CDTI+4CAS condition showed performance advantage for one dependent measure of maneuver efficiency, two measures of safety, and two measures of workload.

Introduction

Background

Right now, en-route commercial aircraft are routed along high-altitude jetways. Because jetways are usually not direct routes from departure to destination, aircraft flight paths are longer than necessary. The U.S. Federal Aviation Administration estimates that direct routing would reduce aircraft fuel use by about 6% (ORA, 1998). At 2005 prices, this would save U.S. airlines over $2B USD in fuel costs alone (derived from BTS [2006] statistics). However, air traffic controllers and pilots will face greater airspace complexity, since direct routing increases both the total number of routes and their geometric complexity. Advanced technology will be required to manage this complexity (Wickens, Mavor, Parasuraman, & McGee, 1998).

Currently, a large body of ingenious air traffic display technology stands poised to address this need. But, serious challenges remain. One issue is how best to display maneuver solutions for aircraft predicted to occupy the same space at the same time.

The crux of the problem is that solutions involving speed cannot be represented veridically. “Veridical” displays represent objects the way we see them in the real world, more or less like a photograph. And most air traffic information is displayed veridically.

So, the underlying, intractable issue lies with the very way speed information is displayed in systems that portray the world as a photograph. Such systems cannot easily display the full range of any aircraft maneuver solutions that involve speed.

In response to this impasse, Knecht and Smith (2001) proposed the idea of *maneuver space* (MS). MS is a *non*veridical information representation space. It does not portray air traffic like a photograph. Instead, MS portrays the four key components of aircraft *maneuver*—heading, speed, altitude, and available maneuver time. In theory, given a fixed lookahead time, a MS-based display *can* represent all possible solutions to all probable traffic conflicts.

The MS as a “hypothesis-tester”

The key to understanding MS is to understand what it represents and how it does it. MS does not represent physical space; it represents our autopilot and the settings we could dial into it. MS is going to tell us whether each of those settings is safe or not. It is a “space of possible maneuvers”—a maneuver space.

In formal terms, MS is a state space representing hypothetical states of the autopilot. It is a 3D Cartesian coordinate space with a left-right x-axis for heading, an in-out y-axis for speed (“throttle-in/throttle-out”), and an up-down z-axis for altitude. Figure 1 illustrates.

The middle of this state space (Cartesian x,y,z = 0,0,0) = current autopilot settings = “current maneuver.” Hypothetical maneuver is thus scaled relative to current maneuver. For example, movement of -1, -5, 1000 in MS corresponds to a 1° left turn, a 5 kt slowdown, and a 1000 ft climb.

In operation, every few seconds, an onboard mathematical conflict probe uses real-time aircraft positions and velocities to judge whether or not we would be safe for the next few minutes *flying straight and level at each hypothetical autopilot setting*.

If a given hypothetical maneuver would be safe, that region of MS is left uncolored. Therefore, safe maneuvers appear black in MS because black is the "color" of empty space.

However, if the conflict probe declares a maneuver unsafe, we call that a *conflict region* (CR) and color-code it by the amount of time left before separation failure would occur. This is its available maneuver time. As Figure 1 shows, green CRs have 6 minutes available maneuver time, yellow, 3 minutes, magenta, 0 minutes. Intermediate times are simply blends of the two nearest colors.

If any single maneuver is made unsafe by multiple aircraft, that CR is color-coded by the shortest available maneuver time.

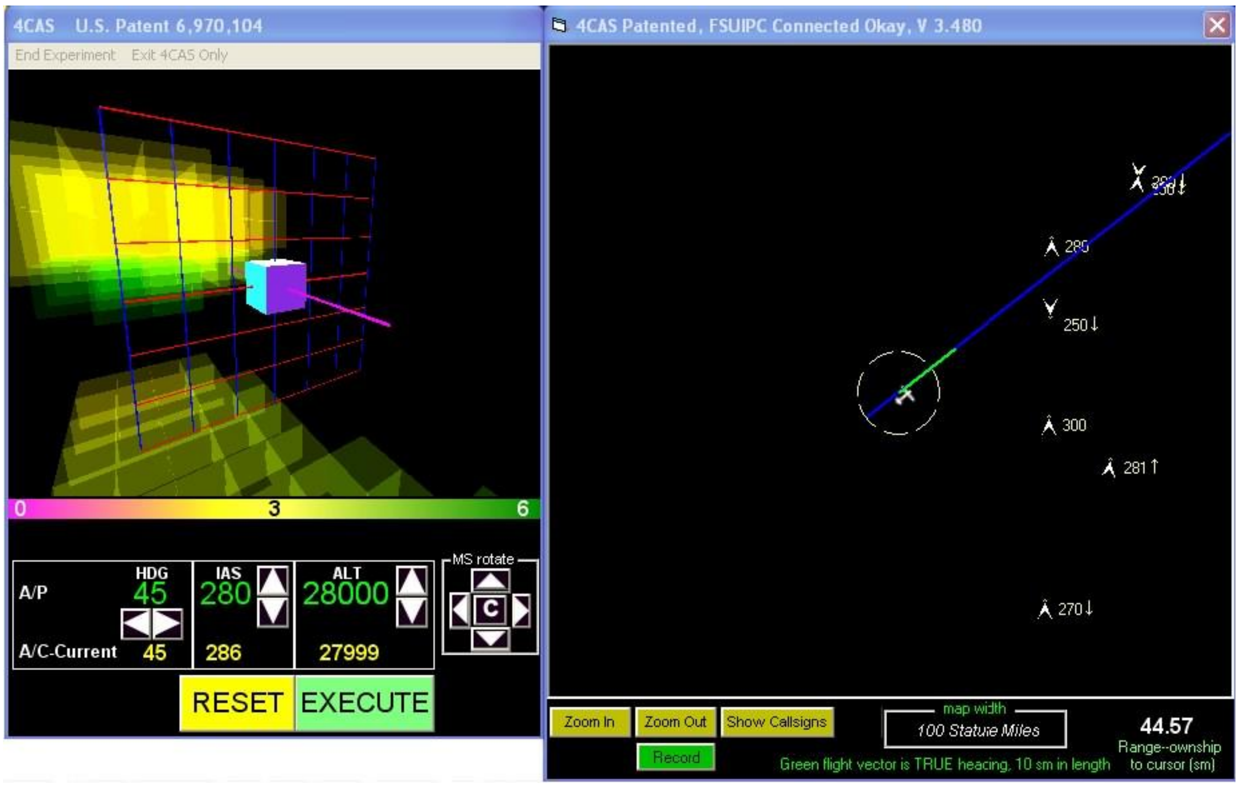

Figure 1. (Right) CDTI, showing scenario L045. (Left) The resulting maneuver space and conflict regions. In the rotatable 4CAS display, the *Heading* axis extends left-right, the *Altitude* axis extends up-down, and the *Speed* axis extends in-out of the virtual plane. To avoid any conflict, no matter how complex, we simply use the control arrows to move the 3D planning cursor (seen here at MS center) to a safe (black) region of MS, and hit "Execute" to reset our autopilot.

All CRs are drawn translucent. This lets us see behind those in front to anything lying farther away in MS.

Once an appropriate range of hypothetical maneuvers has been tested, what we have is one unified display showing *all predicted outcomes of all practical maneuvers*. As Figure 1 clearly shows, unsafe maneuvers visually "pop out" against the black background.

Essentially, the MS is a maneuver hypothesis-tester. If any CR should occur exactly at display center (our current autopilot settings), avoidance is necessary—but easy. We simply move a 3D planning cursor to any safe region of MS, hit the "Execute" button, and our aircraft resets its autopilot to that safe setting. No matter how complex the traffic situation, all this can be done in seconds.

In other words, while MS may be intellectually abstract, in practical terms, it may be quite easy and effective to use.

Purpose of this research

To date, no MS-based collision avoidance system has been successfully tested in a formal setting. The goal of the present research is to build and begin testing such a device for use in en-route tactical airspace.

Method

Apparatus

A part-task flight simulator was assembled, using a Compaq Presario V2000 laptop with an ATI dual-head video card. The 4-Dimensional Collision Avoidance System (4CAS) and its companion CDTI were displayed on the laptop's color monitor, while Microsoft Flight Simulator 2004 (MFS2004) was displayed on an outboard 15" color monitor.

MFS2004 was set up to fly its Boeing 737-400 model. Its native Artificial Intelligence (AI) Traffic was used to create simulated enroute air traffic. A shareware program, Traffic Tools V2.02, allowed partial control of this traffic. An interprocess communication program, FSUIPC V3.48, allowed 4CAS and MFS2004 to talk back and forth.

Custom software displayed a CDTI and 4CAS. The CDTI showed a veridical, top-down, moving-map view of physical space for the pilot's own ship (ownship). A blue path vector remained static relative to the absolute position of the virtual earth beneath. This path ended with a red dot depicting the "destination," signaling the end of each scenario.

Normally, the CDTI refreshed and wrote data to file every 2000 ms. Incursion of traffic within 5 nm laterally and 1000 ft vertically was considered a *pilot deviation* (PD). This triggered a burst timer, boosting the data-sampling rate to 40 ms, capturing point-of-closest-approach (PCA) to within ±30 ft.

4CAS displayed the MS and CRs corresponding to the current traffic situation. For this first experiment, CRs were calculated by a deterministic conflict probe, with assumption of straight-line trajectories. CRs were translucent, allowing pilots to see the 3D planning cursor, other CRs, and safe MS, even when obscured by CRs in front. A color reference bar was drawn just under the MS, to allow rapid understanding of available maneuver time for CRs. A scaling grid was drawn in the virtual *xz*-plane, with a grid scale of 2° in heading and 1000 ft in altitude.

Figure 1 shows a screenshot of what pilots saw 30 seconds into scenario L045 (an approach from the left @ 45°). A color video of the device in operation is posted on www.maneuverspacetechnologies.org.

In Figure 1, an upper-left set of CRs can be seen. This is actually two sets, the higher corresponding to northbound traffic approaching the path from the right at FL300, and the lower set corresponding to traffic, also northbound, approaching the path from the right at FL281, climbing at >100 fpm. A lower-left set of CRs depicts southbound traffic at FL250 (not visible, beyond the top of the CDTI). 4CAS registers no conflict right now because no CR is present at MS 0,0,0. However, increased speed + a 2000 ft climb would *produce* a conflict in about 3 minutes, by causing us to run into the ship at FL300. Similarly, turning left would produce a conflict in about 4.5 minutes, as would a 4000' descent in about 4 minutes.

Clearly, what the CRs show us is maneuvers we do not want to make. Yet, they (presumably) do it with minimal perceptual and cognitive effort. That is an empirical issue about to be tested.

Task

Tests scenarios involved free flight traffic. All scenarios began en-route in mid-flight, at a flight level of 28000 ft (FL 280) and indicated airspeed of 280 kt. Aircraft were not restricted to normal odd-or-even flight levels by thousands (i.e., no "East-West Rule"). The pilots' overall task was simply to stay on course—path + initial altitude + speed—deviating for traffic as necessary, then returning to course when clear of traffic. Upon reaching the geographical "destination" after about 10 minutes, trials were ended manually.

For this first experiment, traffic was programmed to fly straight and level. The number of traffic aircraft in each scenario varied dynamically, but maintained a light-to-moderate density of 4-10 per 200x200 sm^2 maximum area on the CDTI. Each scenario represented crossing a bi-directional, vertical (north- or southbound) traffic stream from one of four approach angles, 0, 45, 135, 225, or 315° (aero coordinates, north=0, increasing clockwise). Xu and Rantanen (in press) suggest that these represent approach angles with relatively high conflict potential. One scenario (L045) contained no conflict. This was designed to test the false-alarm rate. The remaining three scenarios contained one primary conflict each.

Experimental design

The four 10-minute scenarios were used as repeated measures. Half the pilots started in a CDTI-only condition, running the four scenarios, followed by a short break, followed by the CDTI+4CAS condition, using the same scenarios and presentation order. The remaining pilots ran similarly, but with the CDTI+ 4CAS first. Presentation order was counterbalanced according to a 4x4 Latin square. Pilots were not told they would be repeating scenarios. This overall technique was found in prior research to minimize both scenario-specific learning and asymmetrical transfer (Knecht & Hancock, 1999).

Dependent measures

Measures 1-8 were recorded in real time, either every 2000 ms, or immediately after an event occurred, whichever was appropriate. Measures 9-12 came from Likert scales (range 1-6) on a written debrief form administered after the experiment.

1. *Maneuvers made:* Total number of maneuvers made during each scenario.
2. *Maneuver types made:* Number of maneuver *types* used per scenario (max = 3, the types being heading, speed, and/or altitude maneuvers).
3. *Path length:* Total distance traveled, lateral+altitude (sm).
4. *3D max. deviation from path:* 3D normalized maximum deviation from nominal flight path.
5. *Rmin:* Minimum 3D normalized range (Eq. 1) to closest traffic during that scenario.
6. *Maneuver onset time:* Elapsed time from start of a scenario to first maneuver (seconds).
7. *PD duration:* Duration of each pilot deviation (seconds).
8. *PDs:* Experiment-wide number of pilot deviations.
9. *Task ease:* How easy was it for you to avoid traffic?
10. *Time sufficiency:* Was there sufficient time to avoid traffic?
11. *Enjoyability:* How enjoyable was it to use the system?
12. *Training requirements:* How many hours of training would you prefer before handling real traffic such as that experienced during the experiment?

Paired *t*-tests were used for continuous data, where a *z*-test of skew and kurtosis showed distributions to be arguably normal (e.g., $z_{skew} = skew/SE_{skew}$). Standard errors can be found in Fisher (1925/1970). For *Pilot deviations*, a score of 1 was assigned to scenarios containing a PD, 0 otherwise; scores were then analyzed by the nonparametric McNemar test, since low expected cell values violated the assumptions of ^{2}. Remaining scores were analyzed with the nonparametric Wilcoxon test for paired-score ranks.

Angle-of-approach could not cleanly be tested for effect. Scenario L045 was confounded with being the only non-conflictual scenario. Plus, angle was not independent of amount or positioning of traffic since MFS2004 AI Traffic did not allow sufficient precision. However, since scenario order was counterbalanced and all pilots experienced all scenarios, this was expected to exert little experiment-wide effect.

Pilot deviations were defined similar to operational error in en-route air traffic control (ATC), namely, simultaneous approach of traffic to less than 5 nautical miles lateral (*xy*) distance *and* less than 1000 feet vertical (*z*) distance of the ownship in physical space.

Path length was merely the sum of raw linear distances traveled in physical space from one data sample to the next, whether laterally, or in altitude. During maneuver, deviation from path increased the total path length. The more extreme the maneuver, the greater its effect on *Path length*.

3D Maximum deviation from path captured the length of a 3D vector drawn orthogonally from the actual path to the nominal path at the point of farthest deviation-from-path in normalized physical space. Normalization consisted of dividing lateral distances by 5 and vertical distances by 1000 to transform them into "standard ATC separation units" in physical space.

Rmin (Eq. 1) was based on the same normalization just described. It was the minimum distance (the PCA) in normalized physical space from the ownship to the closest traffic ship during each scenario,

$$R_{\min} = \min \sqrt{\left(\frac{xy}{5}\right)^2 + \left(\frac{z}{1000}\right)^2} \qquad (1)$$

xy being lateral separation (nm) and *z* being horizontal separation (ft) in physical space.

Rmin was first described in Knecht & Hancock (1999). Here, its use was bimodal. During PDs, *Rmin* was used as a safety measure, greater *Rmin*s being considered safer. During non-PDs, *Rmin* was used as an efficiency measure, lower *Rmin*s being considered more efficient. While the latter was certainly not always true, it was assumed for the purpose of testing the metric.

Maneuver onset time is particularly important in conflict resolution, for two reasons. First, aircraft take time to maneuver. Second, angular solutions are easily achieved when far from an obstacle, but become increasingly difficult as time grows short. The clearance angle required is = $\tan^{-1}(y/x)$, x being distance to the obstacle (a variable) and y being the obstacle's half-width (a fixed number). As $x \rightarrow 0$, increases rapidly exactly at the time available maneuver time is shrinking fast.

Participants and training

Eight male general aviation pilots received $50 USD each to participate in this first test. Median age was 23 (range 22-34, mean 24.8, SD 3.9). Median civilian flight hours was 870 (range 250-1600, mean 891, SD 439). All participants held at least a private pilot's license, six held instrument ratings, six were double-certified as both Certified Flight Instructor (CFI) and Certified Flight Instructor-Instrument (CFII). Seven held Commercial ratings.

Pre-training was purposely kept brief, in order to explore inherent ease-of-use. Pilots were first given a one-page instruction sheet describing their general task (i.e., to navigate safely through a bidirectional stream of traffic, generally staying on a blue path line at FL 280, deviating for traffic as necessary, then returning to course as soon as possible). They were then shown a one-page description of the CDTI and its operation. Those in the 4CAS+CDTI condition were then shown an additional one-page description of 4CAS and its operation. Pilots were next allowed to practice on two training scenarios, being "walked through" the process of using the CDTI (and 4CAS, depending on to which half they were randomly assigned). They were offered the opportunity to practice as much as they wanted before starting data collection. Most elected to start after about 25-30 minutes of practice. After completion of four data-collection scenarios, pilots were given a short break, after which they retrained for the second half of the experiment. Those who began with only the CDTI were given the instruction sheet for 4CAS. Everyone, no matter what their treatment order, was given the two practice scenarios again, before beginning data collection, and allowed to re-run those practice scenarios until ready to start data collection. Again, the typical re-training session lasted about 25-30 minutes.

Results

Table 1 summarizes the relative performance of CDTI-alone trials versus 4CAS+CDTI trials for the 8 participants x 4 trial-pairs = 64 total trials.

	Table 1. Experiment 1: Relative performance of CDTI-alone v. CDTI + 4CAS trials.							
	Dependent measure (1)	Mean 1 (CDTI-only)	Mean 2 (CDTI + 4CAS)	p(skew normality) *distr1 , distr2*	p(kurtosis normality) *distr1 , distr2*	p (2-tail) *m1* v.*m2*	test	Direction-ality favors 4CAS?
	Efficiency Measures							
1	Path length	78.01	77.39	.009 , .169	.121 , .200	**.006**	Wilcoxon	Yes
2	3D max. deviation from path	1.57	1.37	.002 , .304	<.001 , .147	.280	Wilcoxon	Yes
3	*Rmin* (non-Pilot Deviations only)	1.82	1.79	.362 , .364	.429 , .369	.693	t-test	Yes
4	False alarms (counts) (2)	N=7	N=6			1.0	McNemar	Yes
	Safety Measures							
5	*Rmin* (PDs only)	1.21	1.34	.276 , .075	.190 , .374	.481	t-test	Yes
6	Maneuver onset time (sec)	171.55	119.45	.301 , .014	.029 , .209	**.045**	Wilcoxon	Yes
7	Pilot Deviations, duration (sec)	27.57	11.91	.003 , .023	.006 , .497	**.041**	Wilcoxon	Yes
8	Pilot Deviations (counts)	N=8	N=5			.508	McNemar	Yes
	Workload Measures							
9	N. maneuvers made	6.09	4.63			.070	Wilcoxon	Yes
10	N. maneuver types made (3)	1.78	1.44			**.012**	Wilcoxon	Yes
11	Ease of avoiding traffic	3.9	4.9			.054	Wilcoxon	Yes
12	Had sufficient time to avoid traffic	4.5	5.4			.102	Wilcoxon	Yes
13	Enjoyability of use	3.6	5.3			**.033**	Wilcoxon	Yes
14	Amount of training required (hr)	8.4	10.9			.458	Wilcoxon	No
(1)	Measures 1-3, 5-7, 9, 10 compare trial-pairs within-pilots M 4, 8 are experiment-wide totals. M 11-13 are experiment-wide means on a Likert scale of 1-6, with higher numbers indicating superiority. M 14 is estimated number of hours training needed to achieve competency.							
(2)	Based on-conflict trials only							
(3)	Maximum score of 3 per scenario (heading, speed, and/or altitude).							

Table 1 divides results into three categories of *efficiency*, *safety*, and *workload*. In theory, the safest aircraft are maximally separated, while the most efficient stick to their flight paths. However, safety and efficiency are theoretically antithetical when traffic forces deviations from flight path. The best we can do is make necessary deviations as small as possible, given the safety standard (5 nm/1000 ft). High workload is theoretically antithetical to both safety and efficiency. When operators are stressed, we expect mistakes to be made.

In summary, path length, maneuver onset time, and the duration of pilot deviations were all significantly shorter with 4CAS present. Pilots also reported enjoying the 4CAS trials significantly more than the CDTI-only trials. The remaining measures were non-significant, but all showed directionality in favor of 4CAS, with the exception of estimated hours of training needed to master the system. Unsurprisingly, that showed reverse directionality, since there were two systems to learn instead of just one.

Regression of maneuver onset time onto scenario presentation order showed no apparent experimental effects merely due to the passage of time (e.g. practice, learning, or fatigue effects).

Discussion and Conclusions

This experiment represents the first successful formal test of a new method of representing vehicular maneuver. Eight licensed general aviation pilots each flew eight simulated free flight scenarios, with the goal of deviating as little as possible from a preassigned flight path while still maintaining standard en-route separation from traffic. Compared to a CDTI, the CDTI+4CAS condition showed performance advantage for one dependent measure of maneuver efficiency, two measures of maneuver safety, and two measures of user workload.

It needs to be clear that this was a preliminary experiment only, and should be considered only a very modest step in validation-of-concept. First, it was a "straw man" experiment, which pitted a device with both conflict detection and resolution capability against a device which had neither. As such, it was a logical place to begin, but deserves little more than historical note. Second, 4CAS performance was not perfect; both pilot deviations and false alarms did occur. This is typical of manually initiated CAS (MICAS), but must be stated nonetheless.

A number of important tactical and strategic heuristic issues were duly revealed:

1. Testing more participants is needed to boost statistical power.
2. More challenging scenarios are needed to elicit larger treatment effects.
3. Potentially trivial solutions (e.g., simply "diving beneath the entire traffic stream") need to be controlled by the use of blocking aircraft.
4. The currently high 4CAS false alarm rate could be reduced by adding some kind of alert to confirm when maneuver is necessary in "close calls."
5. Automatic program shutoff is needed to increase the accuracy of path length measures.
6. One pilot misunderstood the task. Better instructions are needed.
7. A better traffic creation system would enhance measurement of treatment effect.
8. A small fraction of pilots may have failed to understand what 4CAS is, and how it works.

Lessons 1-6 need no further explanation. Lesson 7 involved MFS2004 AI Traffic's lack of full control over traffic position and velocity. We could generate "traffic streams," but only the ownship could be fully controlled. The AI Traffic flew straight quite well. But, vertical speed was sometimes as absurdly high as 6000 fpm. This led to disconcerting altitude "porpoising" (oscillating overshoot/undershoot) and occasional conflict probe false alarms. Clearly, a better method of generating traffic is needed.

Lesson 8 came, both from noting PDs and reversals of directionality in the recorded data (i.e., where performance was better in the CDTI-only condition). Both trains of logic led one to question whether all participants fully understood 4CAS. The issue may just be one of training. But, a second, equally plausible, hypothesis is that the CDTI provides such a compelling picture of possible conflict that it cognitively overrides the "signal" from 4CAS. Finally, a third possibility is that some fraction of the general population may inherently have trouble with certain varieties of conceptual thought.

These issues are central to determining validity-of-concept, and require further experimentation. Once technological and methodological details have been worked out, we should start getting a better feel for cause and effect here.

Acknowledgments

The author wishes to thank the Wolf Aviation Fund for funding pilot salaries. Dr. Kevin Williams is acknowledged for his contribution of computer code that formed the initial core of the CDTI.

References

Bureau of Transportation Statistics. Airline fuel cost and consumption (1977-2006). Retrieved December 18, 2006 from www.bts.gov/xml/fuel/report/src/index.xml.

Fisher, R.A. (1970) *Statistical methods for research workers.* Edinburgh: Tweeddale Court. (Original work published 1925).

Knecht, W.R. & Hancock, P.A.. (1999). Separation maintenance in high-stress free flight using a time-to-contact-based cockpit display of traffic information. *Proceedings of the 43rd Annual Meeting of the Human Factors and Ergonomics Society*, 16-20.

Knecht, W.R., & Smith, K. (2001) The manoeuvre space. In D. Harris (Ed.), *Engineering psychology and cognitive ergonomics*. Aldershot: Ashgate.

Operations Research and Analysis. (1998). *The impact of national airspace systems modernization on aircraft emissions.* (Technical Report DOT/FAA/SD-400-98/1). Washington, DC: Federal Aviation Administration.

Wickens, C.D., Mavor, A.S., Parasuraman, R., & McGee, J.P. (Eds.). (1998). *The future of air traffic control: Human operators and automation*. Washington, DC: National Academy Press.

Xu, X., & Rantanen, E.M. (in press). Effects of air traffic geometry on pilots' conflict detection with cockpit display of traffic information. *Human Factors.*