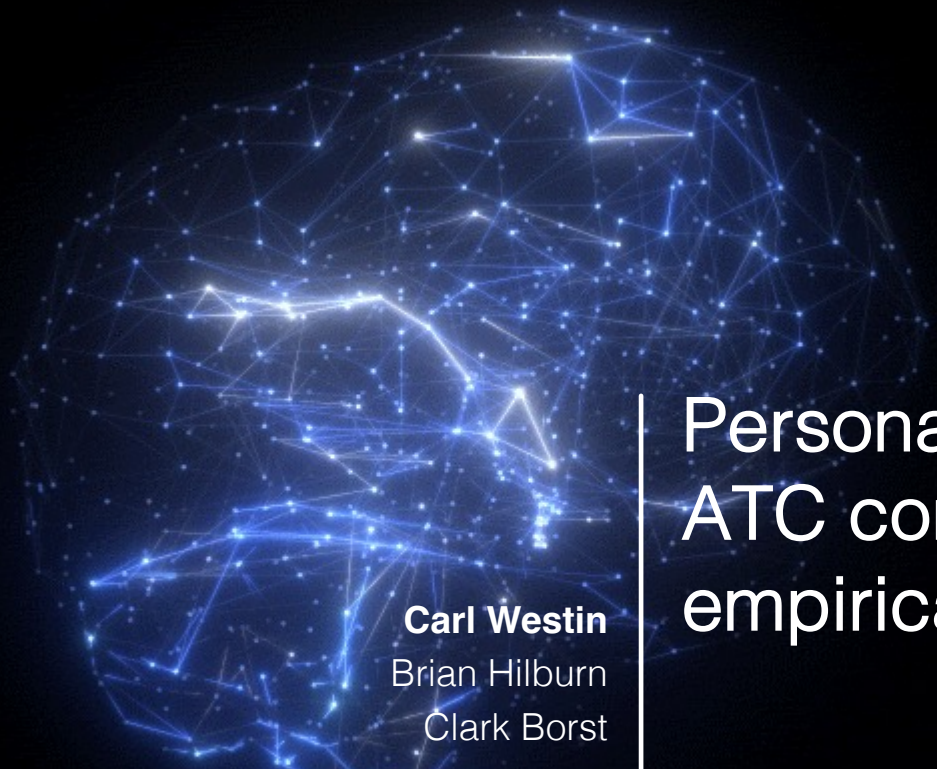


MAHALO
Modern ATM via Human/Automation Learning Optimisation

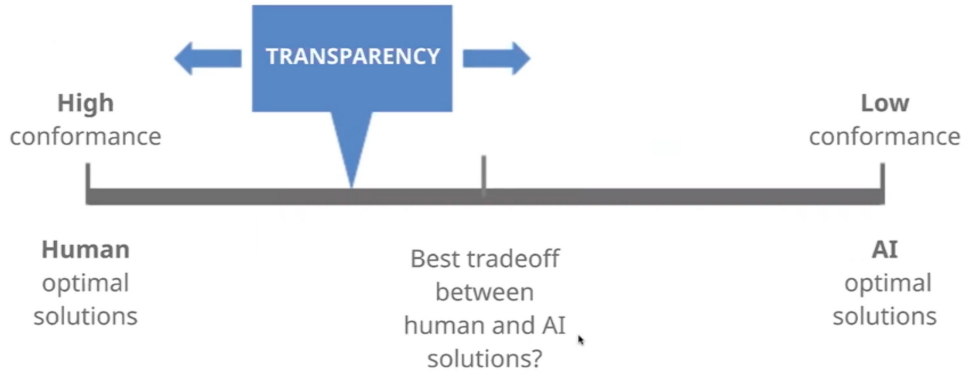
Personalized and transparent AI support for ATC conflict detection and resolution: an empirical study

Carl Westin
Brian Hilburn
Clark Borst
Erik-Jan van Kampen
Tiago Nunes
Supathida Boonsong
Matteo Cocchioni
Stefano Bonelli

sesar JOINT UNDERTAKING
12th SESAR Innovation Days
5-8 December 2022, Budapest
HungaroControl
sesar DIGITAL ACADEMY

MAHALO objectives

- Develop ML models for CD&R using:
  - **Supervised Learning** – to mimic ATCO solutions (*conformal*)
  - **Reinforcement Learning** – to generate ATCO independent optimized solutions (*nonconformal*)
- Empirically evaluate effects of **ML model conformance** and **advisory transparency** on ATCO advisory response
- Derive general design guidelines



Design challenge

Understanding automation

- Why does it propose that solution?

How should we build Machine Learning?

Transparency

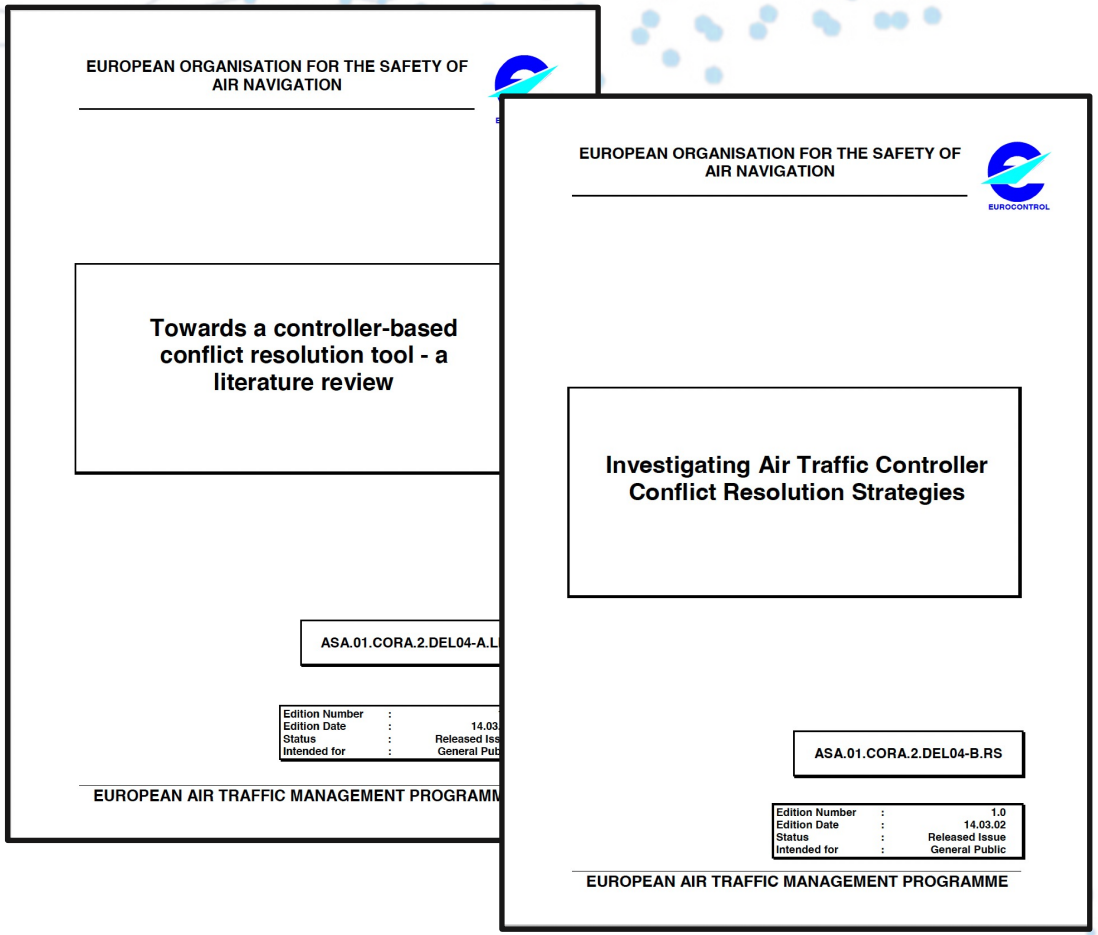
Is automation's inner process explainable to human?

Strategic conformance

Does automation seem to match human strategies?

CONFORMANCE	TRANSPARENCY: *Low*	TRANSPARENCY: *High*
Low	**Stupid automation** *"It's doing a strange thing, and I don't understand why…"*	**Peculiar automation** *"It's doing a strange thing, but I understand why…"*
High	**Confusing automation** *"It's doing the right thing, but I don't understand why…"*	**Perfect automation** *"It's doing the right thing, and I understand why…"*

Personalized automation support



EUROPEAN ORGANISATION FOR THE SAFETY OF AIR NAVIGATION

Towards a controller-based conflict resolution tool - a literature review

ASA.01.CORA.2.DEL04-A.L

Edition Number :
Edition Date : 14.03
Status : Released Iss
Intended for : General Pub

EUROPEAN AIR TRAFFIC MANAGEMENT PROGRAMM

EUROPEAN ORGANISATION FOR THE SAFETY OF AIR NAVIGATION

EUROCONTROL

Investigating Air Traffic Controller Conflict Resolution Strategies

ASA.01.CORA.2.DEL04-B.RS

Edition Number : 1.0
Edition Date : 14.03.02
Status : Released Issue
Intended for : General Public

EUROPEAN AIR TRAFFIC MANAGEMENT PROGRAMME

Will Controllers Accept a Machine That Thinks Like They Think? The Role of Strategic Conformance in Decision Aiding Automatio

Brian Hilburn, Carl Westin, and Clark Borst

In a series of real time trials, we simulated sophisticated air traffic m agement conflict resolution automation using unrecognizable replay controllers' own performance. Using a fairly novel experimental de and a prototype air traffic control interface, we explored with operati controllers the interactive effects of traffic complexity, level of automa and "strategic conformance" (defined as *the match between human machine solution strategy*) on a number of dependent measures. Confo advisories (exact replays of a given controller's previous solution) accepted more often, rated higher, and responded to faster than were conformal advisories (replays of a colleague's different solution). In end, one result stood out in particular: roughly 24% of conformal advis were rejected by controllers. How could it be that controllers, in ef disagreed with their very own solutions roughly one quarter of the t The project is currently exploring this and other related issues thro extended human-in-the-loop simulations.

Brian Hilburn and Carl Westin are with the Center for Human Perform Research, Voorburg, the Netherlands; Clark Borst is with the Delft Universi Technology, Delft, the Netherlands.

Received May 2014; accepted August 29, 2014.

Air Traffic Control Quarterly, Vol. 22(2) pp. 115–136 (2014)
© 2014 Air Traffic Control Association Institute, Inc. CCC 1064-3818/95/0301

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

IEEE TRANSACTIONS ON HUMAN-MACHINE SYSTEMS 1

Strategic Conformance: Overcoming Acceptance Issues of Decision Aiding Automation?

Carl Westin, Clark Borst, and Brian Hilburn

Abstract—Cognitive engineering researchers have long studied the complexity and reliability of human–automation interaction. Historically, though, the area of human–automation decision-making compatibility has received less attention. Paradoxically, this could in the future become one of the most critical issues of all, as mismatches between human and automation problem-solving styles could threaten the adoption of automation. This paper presents the concept of *strategic conformance* as a potential key factor influencing initial acceptance of automation, specifically decision aiding systems capable of guiding decision and action. Here, strategic conformance represents the match in problem-solving style between decision aiding automation and the individual operator. The theoretical foundation builds on the compatibility construct found in technology acceptance theories such as the innovation diffusion and technology acceptance models. The paper concludes with a critical discussion on the limitations and drawbacks of strategic conformance. It is proposed that the construct would be most applicable at the introductory phase of new decision aiding automation, in helping to foster operators' initial acceptance of such automation.

Index Terms—Acceptance, automation, compatibility, decision aid, decision-making, strategic conformance.

I. Introduction

SINCE the advent of the microprocessor nearly 50 years ago, numerous work environments have come to increasingly rely on some form of computer automation. Although we have come to accept automation taking over routine and low-level tasks, there remains some resistance to automation of safety-critical functions, especially in work domains that mandate automation use and rely on well-educated, well-trained, and highly skilled professionals [1]–[4].

Cognitive engineering (CE) researchers have studied automation use in relation to such underlying factors as situational awareness, trust, workload, risk, reliability, and level of automation [5]–[9]. Findings suggest that: 1) trust in automation develops over time as a result of prolonged experience [8]; 2) acceptance and operator performance decrease when the authority and autonomy of automation increase [10]–[12]; and 3) acceptance and operator performance benefit from automation actively involving the operator in the control and decision-making loops [13].

CE researchers have, however, historically paid less attention to factors affecting the *initial* acceptance of new technology, thus factors possibly preceding trust, reliability, and others. Notice that the rejection of new technology can begin at first exposure, perhaps even before an operator has actually used that technology [14]. Notice in this a potential paradox: An operator might only develop trust after using a system, but might also be unwilling to trust a system he/she has not used. For this reason, initial acceptance of advanced decision-making automation can play a critical role in its successful deployment.

Sociology, psychology, and information systems communities, on the other hand, have studied factors underlying initial acceptance. Here, the compatibility between human and technology is considered a key construct for overcoming the hurdle toward initial acceptance and technology adoption. "Compatibility" in this case refers to the perceived fit of a technology within the context in which it is used, driven by the user's values, experiences, and needs [14]. In general, the more compatible a technology is, the more likely it is to be accepted.

Presumably, compatibility can serve to mitigate initial acceptance issues of automated decision aids. Previous research has underlined preliminary benefits of matching automation's problem-solving strategies with the human, for example, by modeling human decision-making heuristics [15]–[18] tuned to a group of people. Would there perhaps be a greater benefit in terms of acceptance if automation's problem-solving style were matched to that of the individual? To our knowledge, no theoretical or empirical work has specifically focused on differences in decision aid problem-solving styles and its effect on individual operator acceptance.

In this paper, we introduce the concept of *strategic conformance* as a potential key factor (and subcomponent of compatibility) influencing the initial acceptance of decision aiding automation. We define strategic conformance as the degree to which automation's problem-solving style matches that of the individual human. A person's problem-solving style is made up of both the product (solution) and its associated process (underlying strategies). The latter is only apparent since the process cannot be determined by knowing the product, only inferred from observable behavior or output. We hypothesize that "strategic conformal" automation can, first and foremost, promote initial acceptance of new technology, but also improve

Manuscript received October 21, 2014; revised April 11, 2015 and August 28, 2015; accepted August 29, 2015. This work was supported by EUROCONTROL acting on behalf of the SESAR Joint Undertaking (the SJU) and the EUROPEAN UNION as part of Work Package E (project MUFASA: Multidimensional Framework for Advanced SESAR Automation) in the SESAR Program. Opinions expressed in this work reflect the authors views only and EUROCONTROL and/or the SJU shall not be considered liable for them or for any use that may be made of the information contained herein. This paper was recommended by Associate Editor J. D. Lee.

C. Westin and C. Borst are with the Control and Simulation Section, Faculty of Aerospace Engineering, Delft University of Technology, 2629 HS Delft, The Netherlands (e-mail: C.A.L.Westin@tudelft.nl; c.borst@tudelft.nl).

B. Hilburn is with the Center for Human Performance Research, 2273 XC Voorburg, The Netherlands (e-mail: brian@chpr.nl).

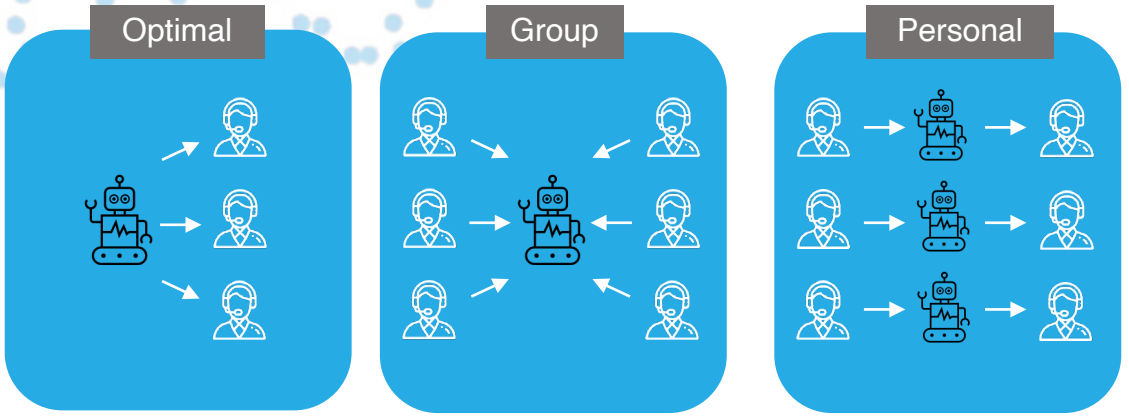
Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/THMS.2015.2482480

2168-2291 © 2015 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

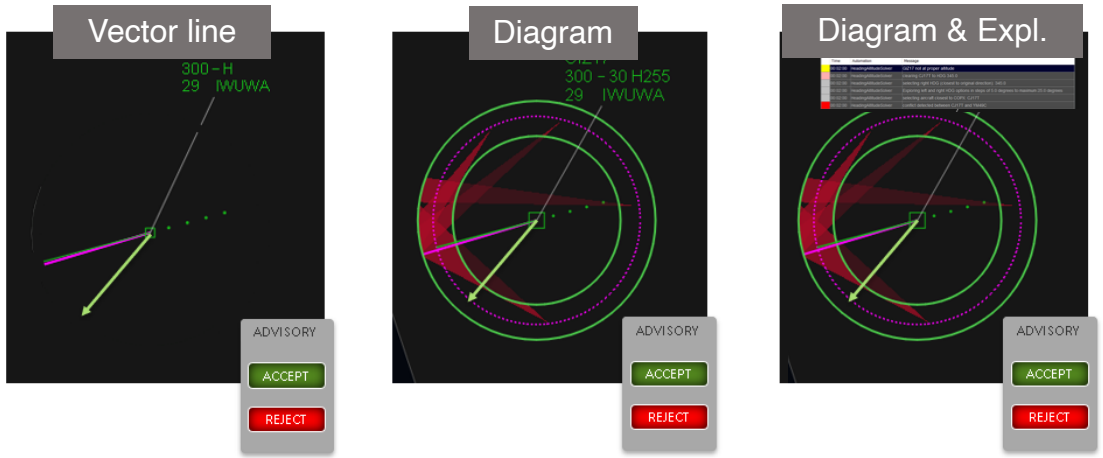
Conformance & Transparency variables

ML model conformance



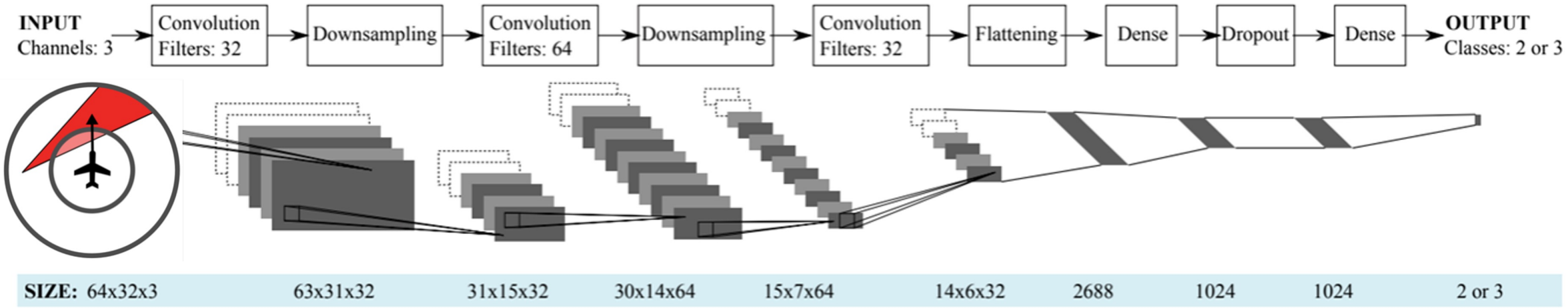
LOW → HIGH

Advisory transparency



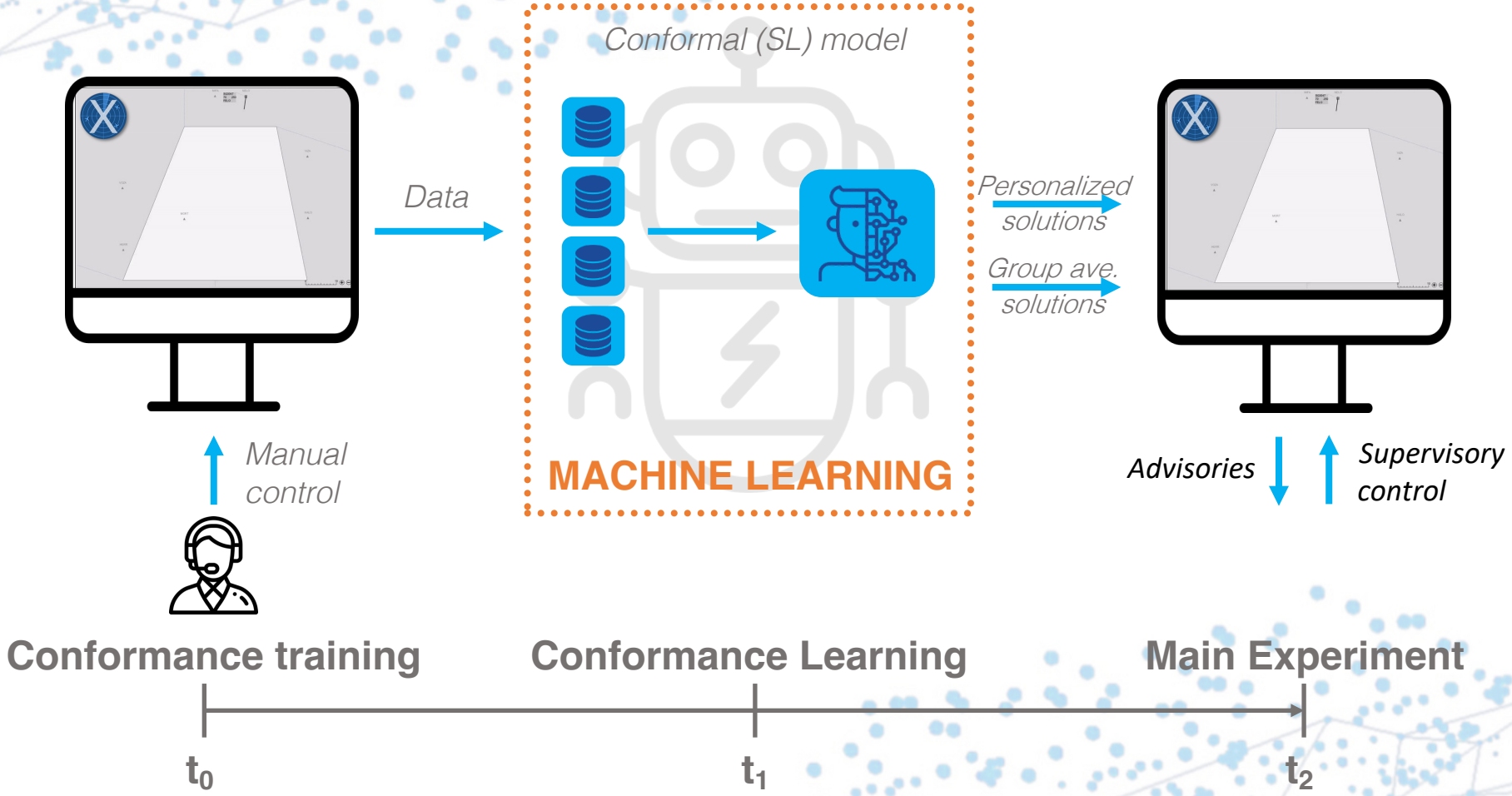
Supervised Learning (conformal advisories)

Convolutional Neural Networks (CNN) – good for processing image data

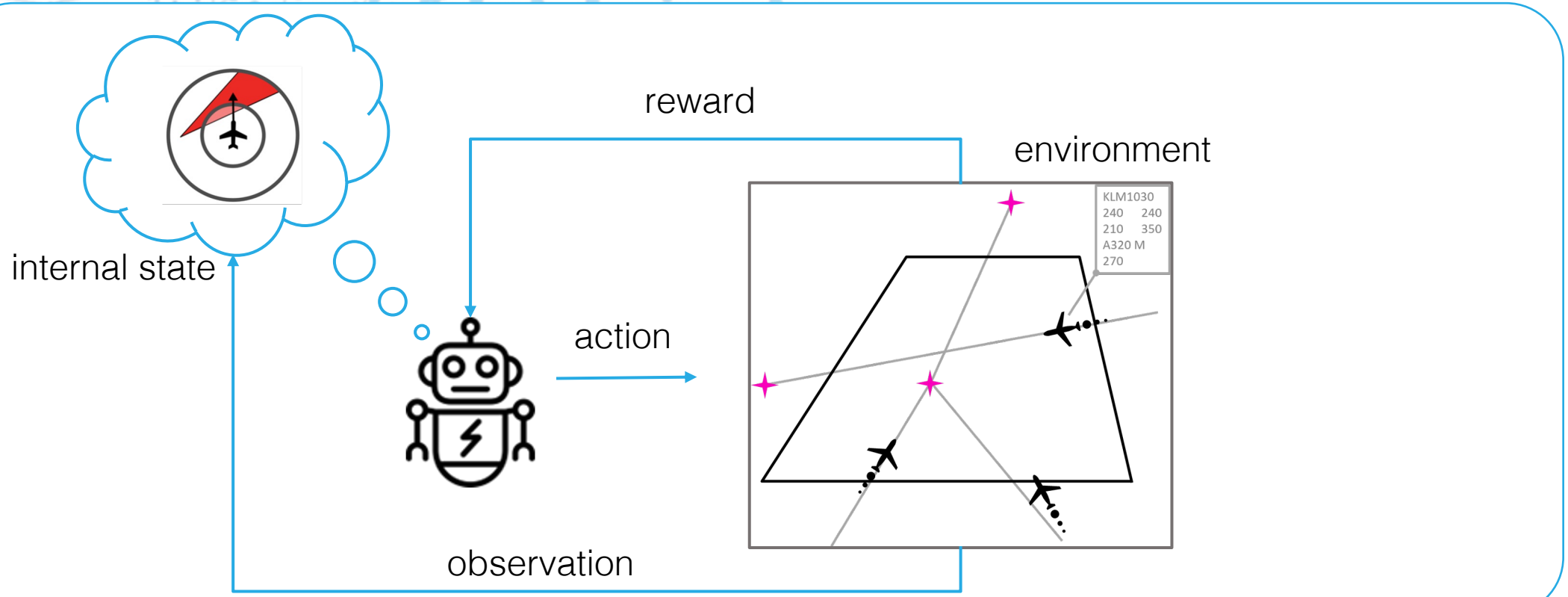


Goal: Build ~~personal~~ and group prediction model for conflict resolutions

Experimental setup



Reinforcement Learning (optimized advisories)



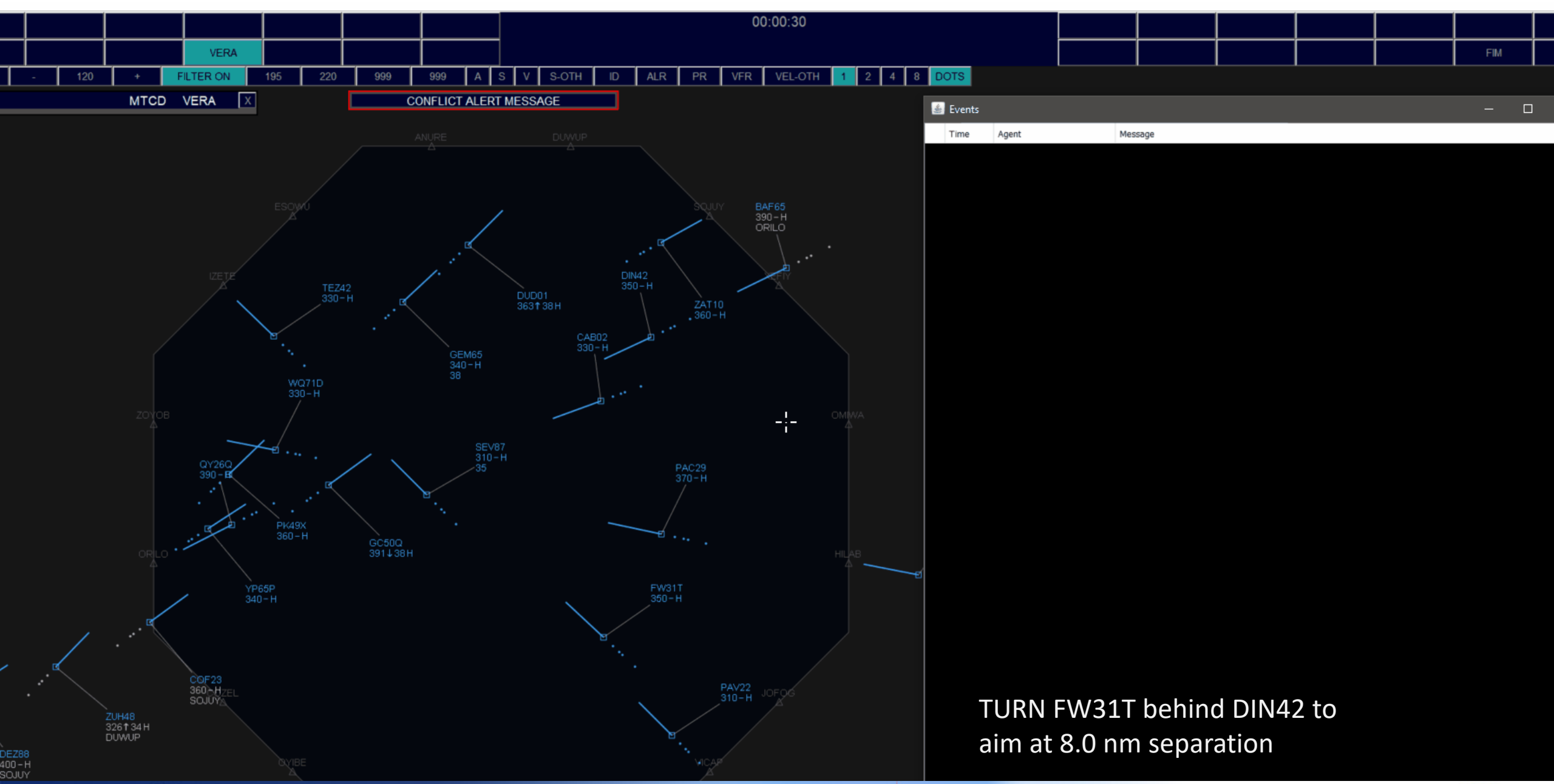
Goal: Build optimized prediction model for conflict resolutions

Experiment

Participant task

- Supervise automation controlling all traffic
- inspect/accept/rate resolution proposals issued by automation.

	SIM1	SIM2
Participants	19 (-1)	15 (-1)
Data (*conflicts solved*)	342 (*18 per participant*)	270 (*18 per participant*)



Dependent measures

The system solved the conflict the same way I would have.

1	2	3	4	5	6
Disagree highly					Agree highly

I can understand why the system suggested that solution.

1	2	3	4	5	6
Disagree highly					Agree highly

AFTER EACH SOLUTION

- Acceptance response
- Agreement rating
- Advisory conformance rating
- Advisory understanding rating
- Response time
- Delta closest point of approach (CPA) distance
- Workload rating

Results

Conformance effects

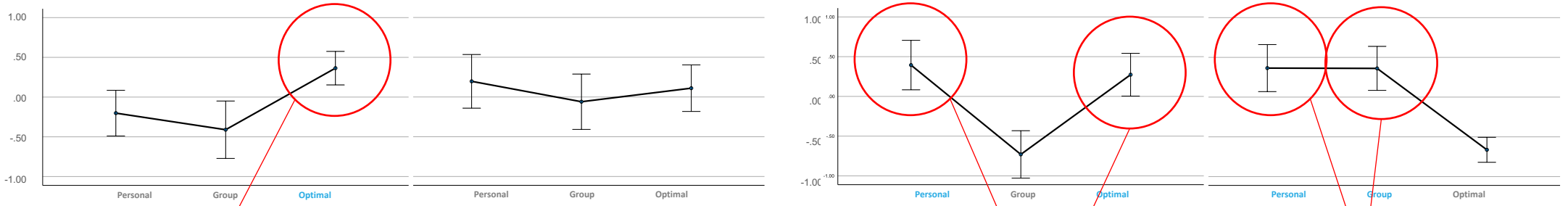
Measure	SIM1 Scen. A	SIM1 Scen. B	SIM2 Scen. A	SIM2 Scen. B
Agreement ratings	👍		👍	👍
Workload ratings		👍		
Delta CPA distances (nm)	👍		👍	👍
Response time (s)			👍	

Conformance and transparency effects

Acceptance response	👍		👍	👍
Advisory conformance rating	👍		👍	👍
Advisory understanding rating			👍	👍

SIM1, Scen. A | *SIM1, Scen. B* | *SIM2, Scen. A* | *SIM2, Scen. B*

Agreement ratings (z-scored)



Acceptance response

Optimal preferred | -- | Personal preferred | Personal and group preferred

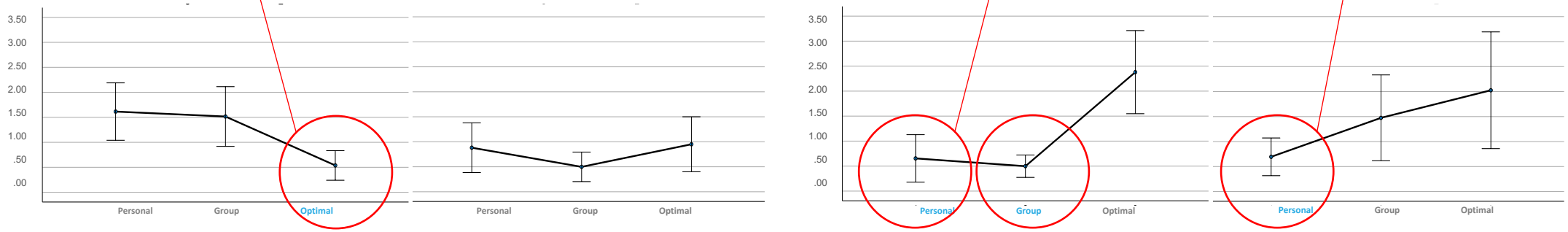
Conformance rating (*The system solved the conflict the same way I would have*)

Optimal preferred | -- | Personal and optimal preferred | Personal and group preferred

Understanding rating (*I understand why the system suggested that solution*)

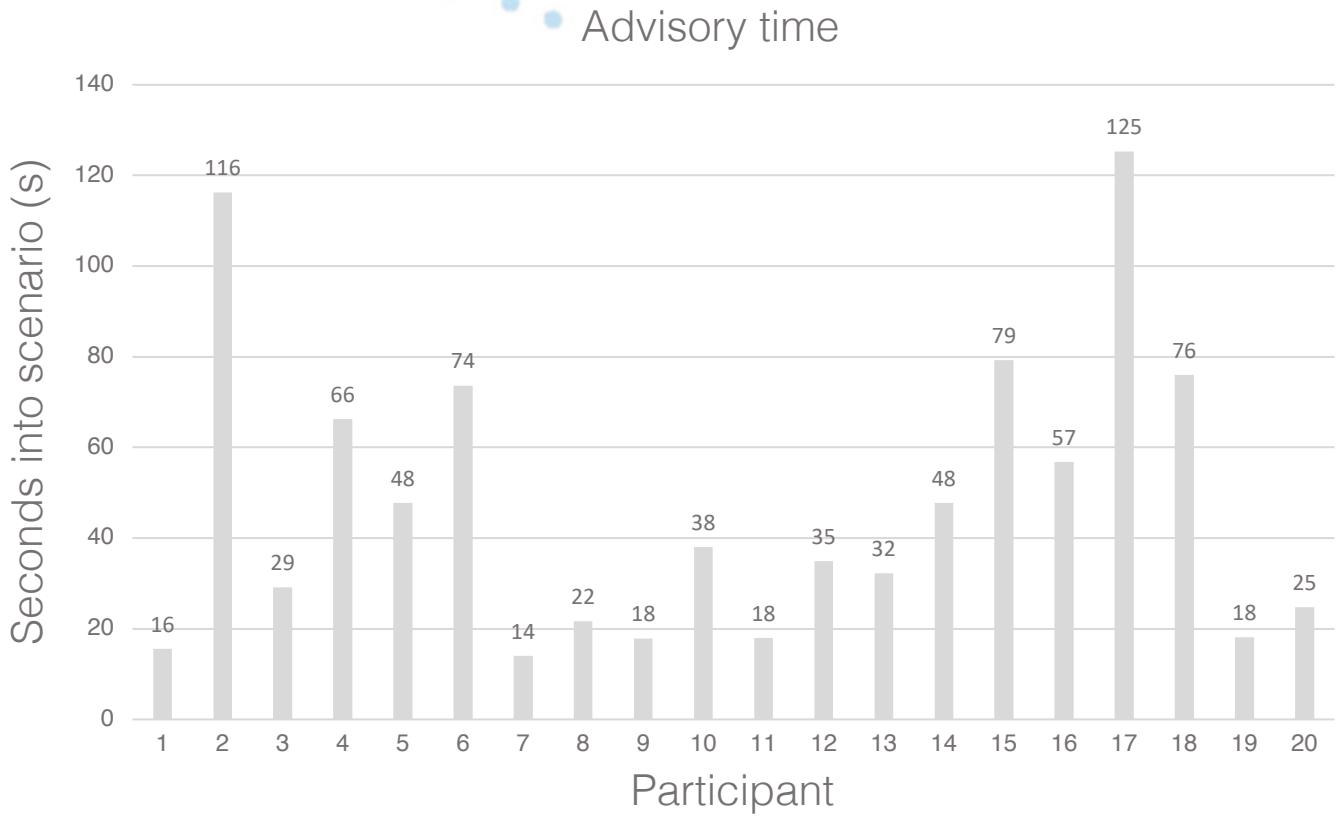
-- | -- | Group less understandable | Optimal less understandable

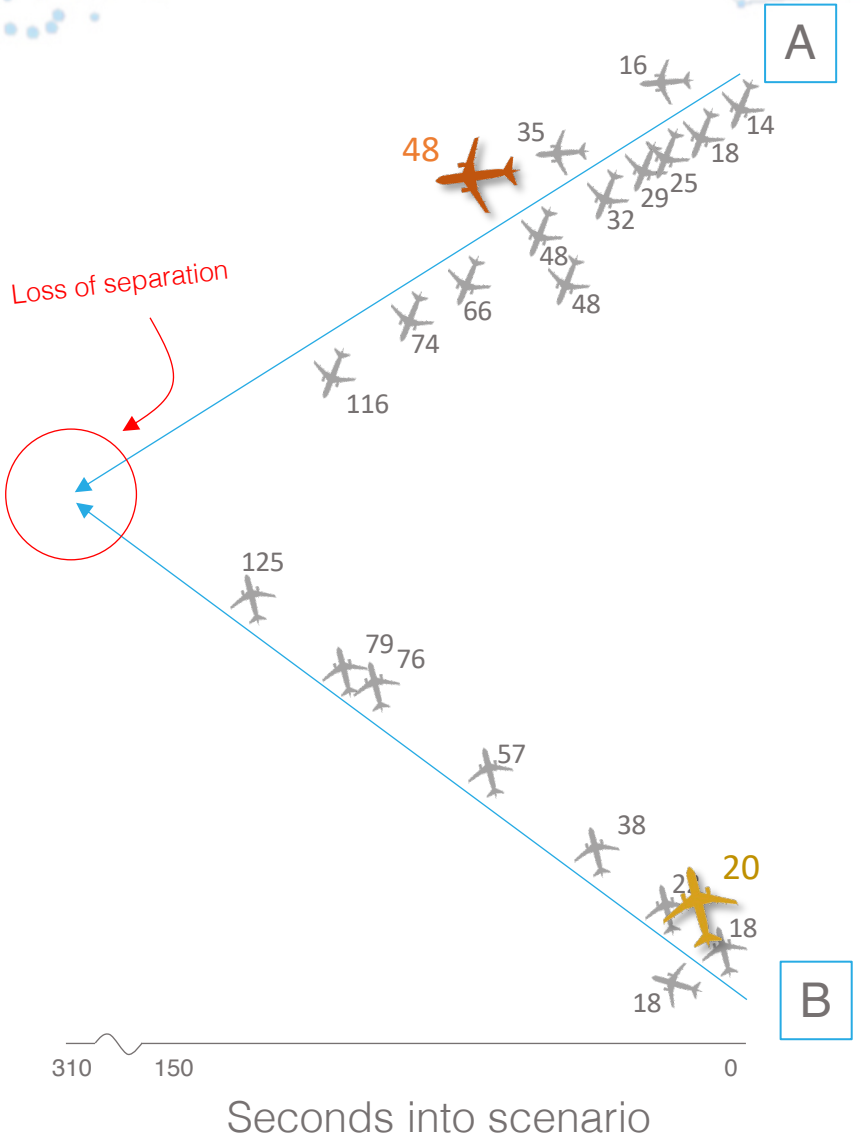
Delta CPA distances (nm)



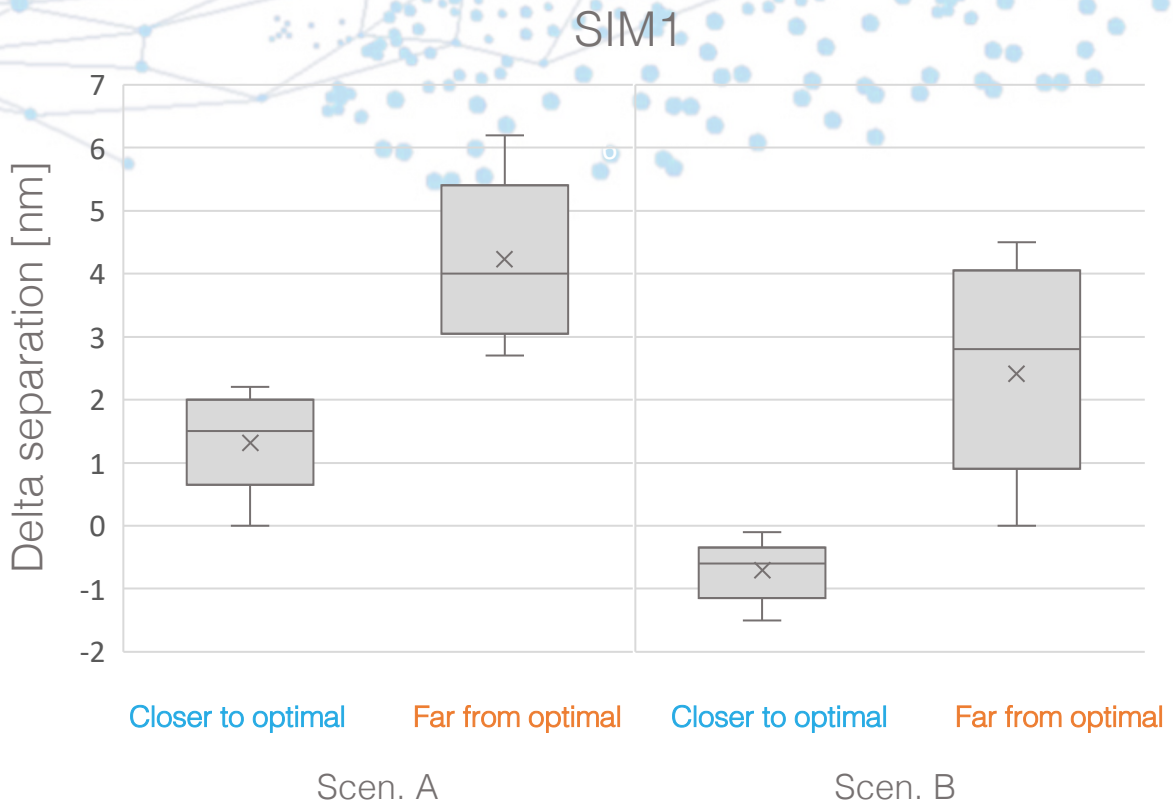
Advisory conformance

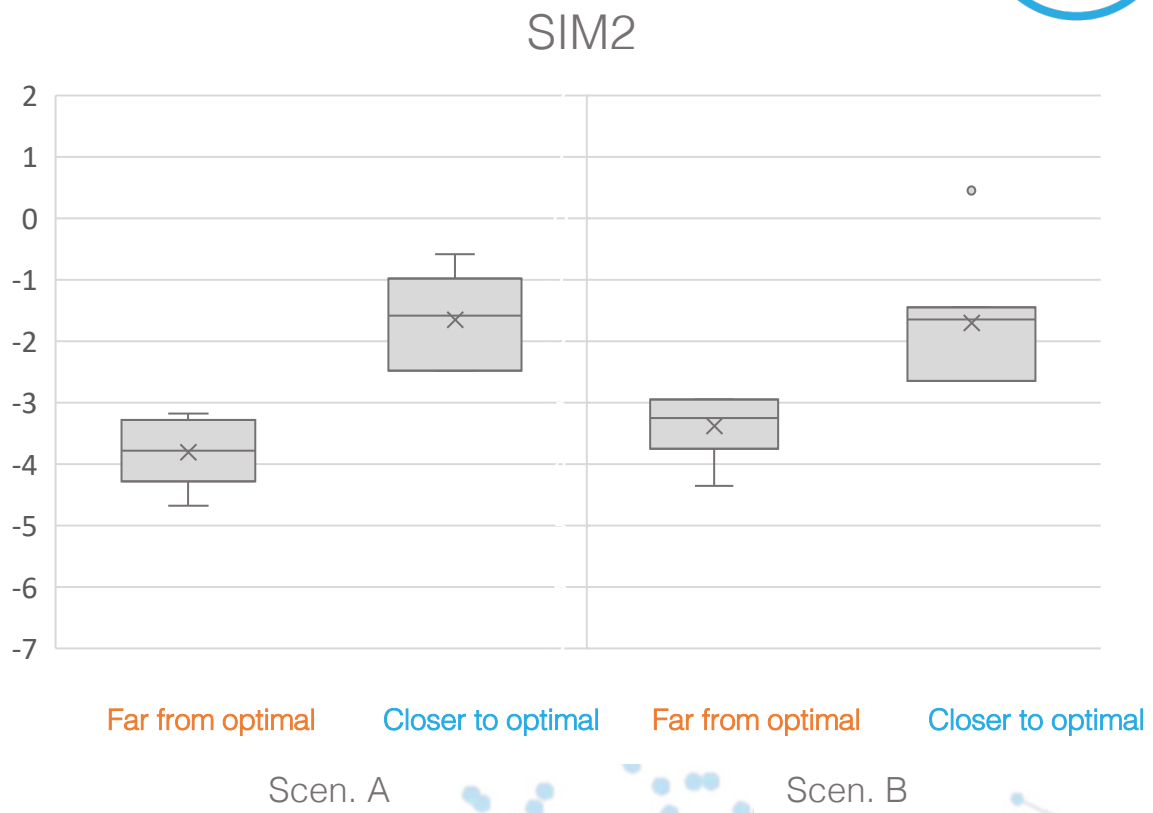
SIM1, Scen.A





Individual differences in CPA vs optimized advisory





Delta Separation: *Difference in nautical miles (nm) between personal model separation distance and optimal model separation distance.*

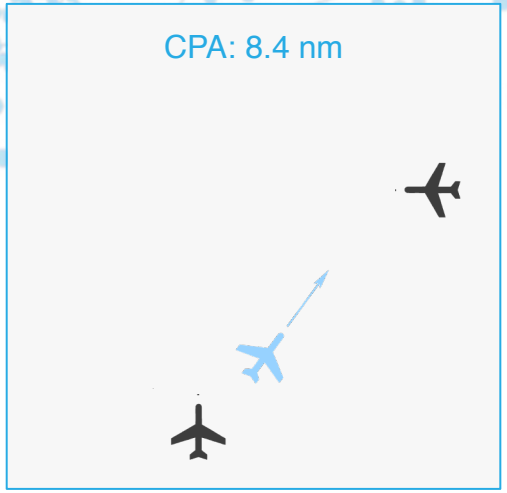
Results

Differences between ATCO groups depending on how close their separation distance preferences are relation to the target CPA in the optimal advisory

	SIM1		SIM2	
Measure	Scen. A	Scen. B	Scen. A	Scen. B
Agreement ratings	👍 T0	👍 T2		
Workload ratings			👍 T1	
Delta CPA distances (nm)		👍 T2		👍 T2
Response time (s)			👍 T0	
Acceptance response		👍 T2		
Advisory conformance rating		👍 T2		👍 T0
Advisory understanding rating	👍 T2			👍 T0

Results

Close to optimal



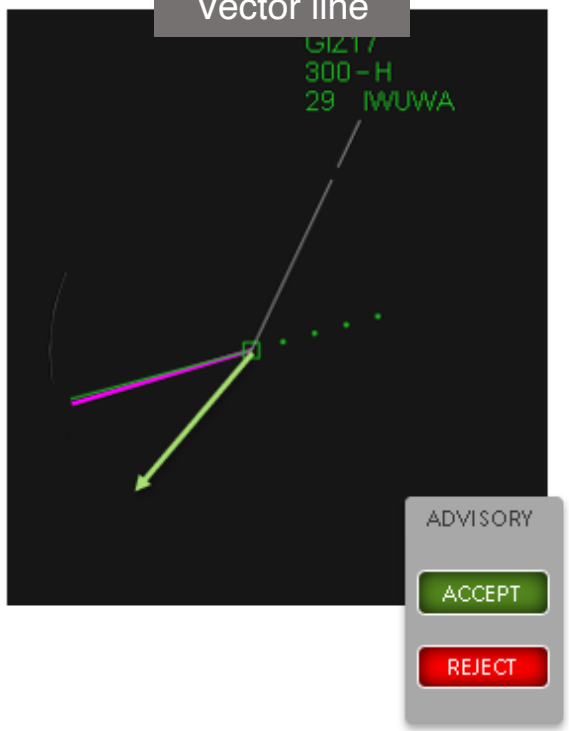
Far from optimal

Group with a preferred separation distance closer to optimal CPA:

- Accepted advisories with less interference
- Higher agreement ratings
- Higher conformance ratings
- Higher understanding ratings
- Smaller CPA distances
- Lower workload ratings
- Faster response time

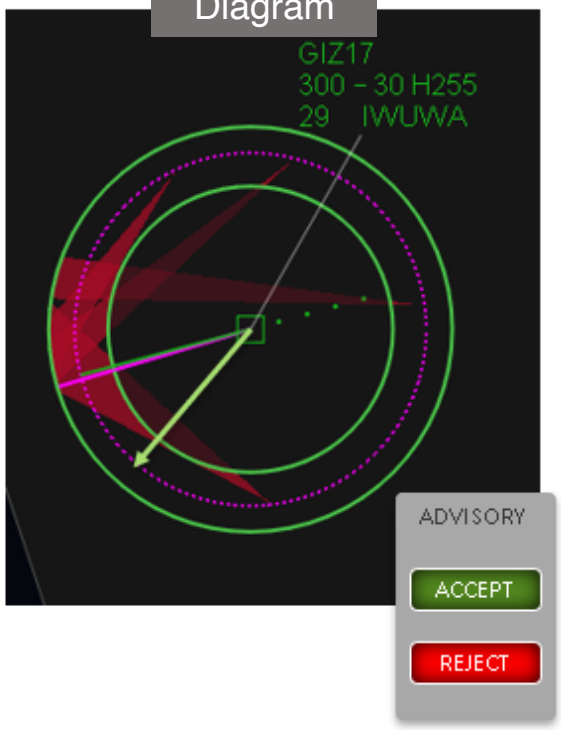
Transparency & separation distance

Vector line



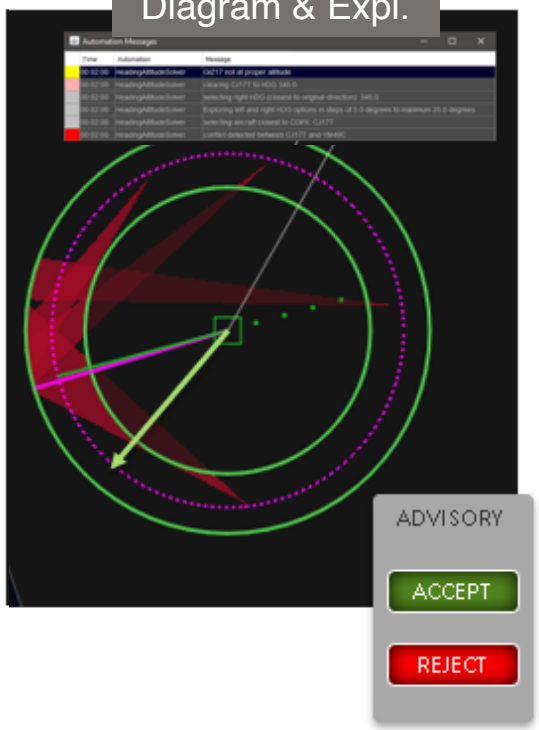
No information on separation distance

Diagram



Separation can be judged by comparing distance between vector with red triangle

Diagram & Expl.

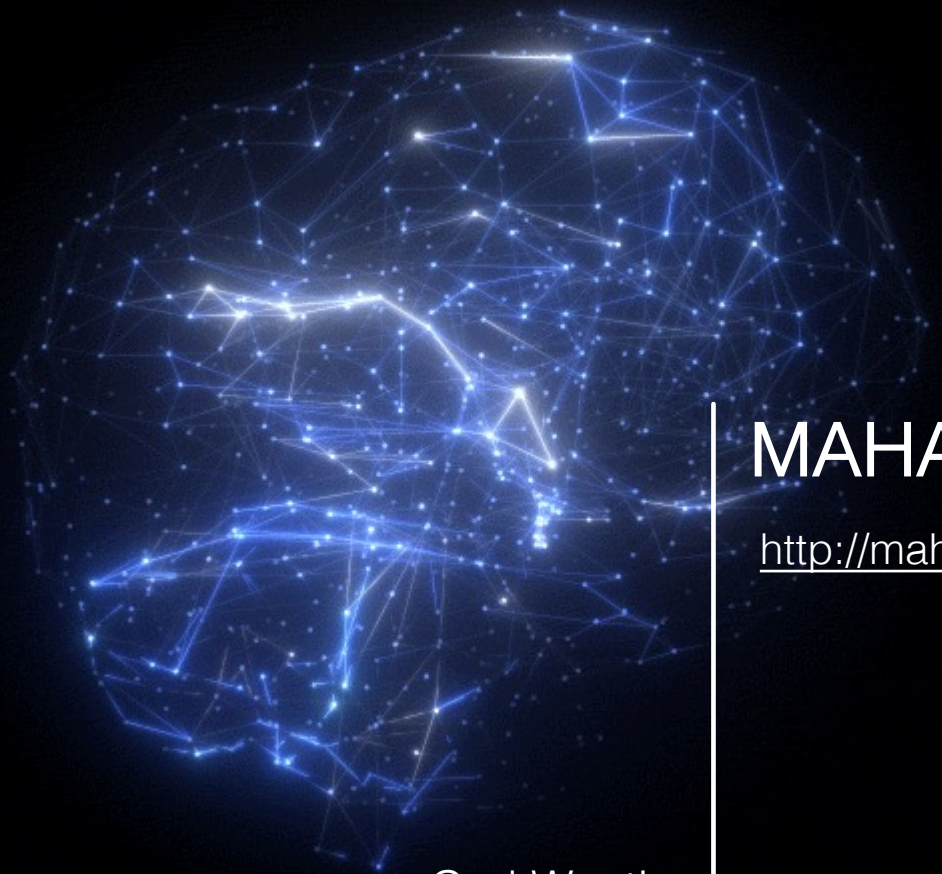


Target separation part of explanation

Guidelines for future AI systems in ATC

ML/AI Design	**Personalisation**	Transparency	HCI	General

- Future ATC systems should acknowledge individual differences.
- Future ATC system should explore personalisation mechanisms to benefit human-AI teaming.
- Future systems should consider individual preferences in problem solving only when appropriate.
- If the system goes against the individual's preferences, the system should be able to provide an explanation for why the system believes its solution to be better than the individual's.



MAHALO for listening

http://mahaloproject.eu

Carl Westin
Linköping University, Sweden,
carl.westin@liu.se