

Agent transparency and human performance in the context of autonomous collision avoidance

ICMASS – 2024

Koen van de Merwe, koen.van.de.merwe@dnv.com

30 October 2024



Decision support

“Fully” autonomous

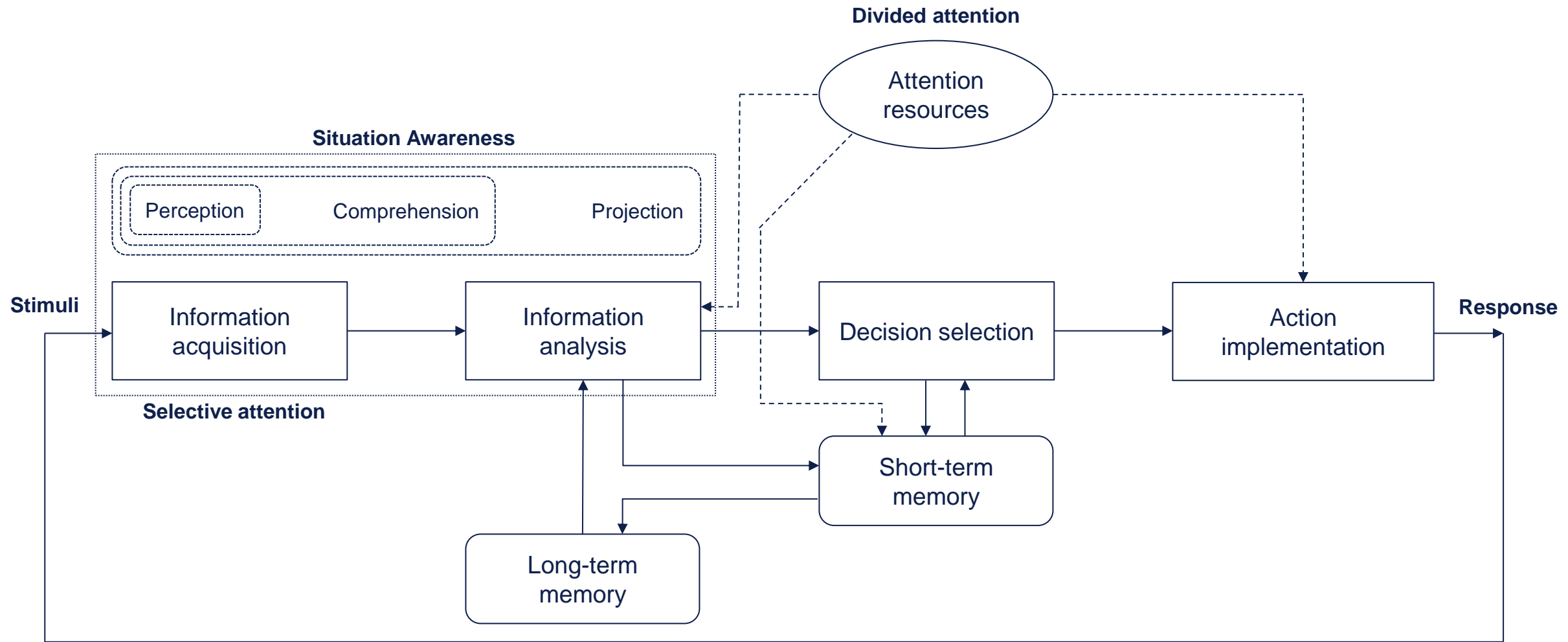
Human oversight



Out-of-the-loop

© Masterly

A model for human information processing



“Restoring the information loop”

- Transparency is about making the system understandable and predictable to its users
- Its goal is to enable operators to maintain proper Situation Awareness of the system in its environment without becoming overloaded

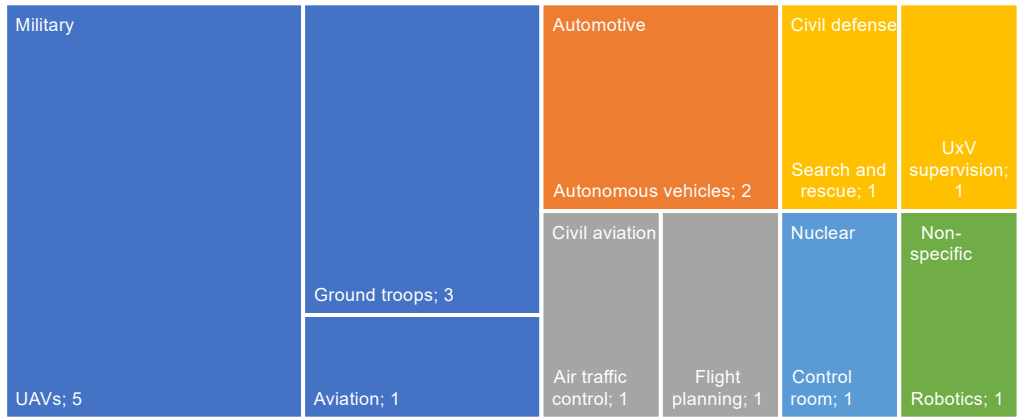
Endsley, Bolte, and Jones (2003)

Mercado et al. (2016)

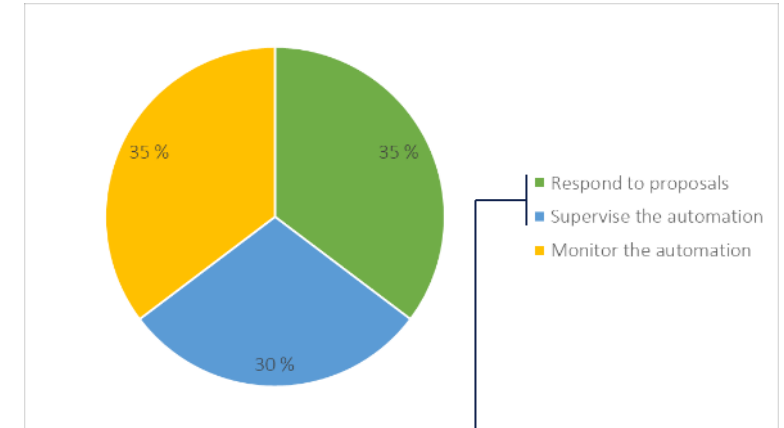
How does agent transparency support human performance in supervisory control?



Exploring transparency in the literature



Lack of maritime domain

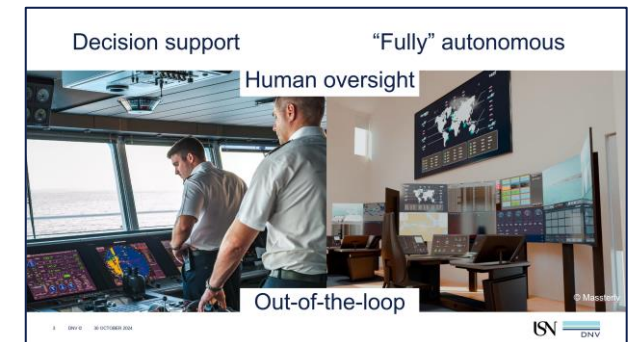


Apparent relation between task type and HF variables

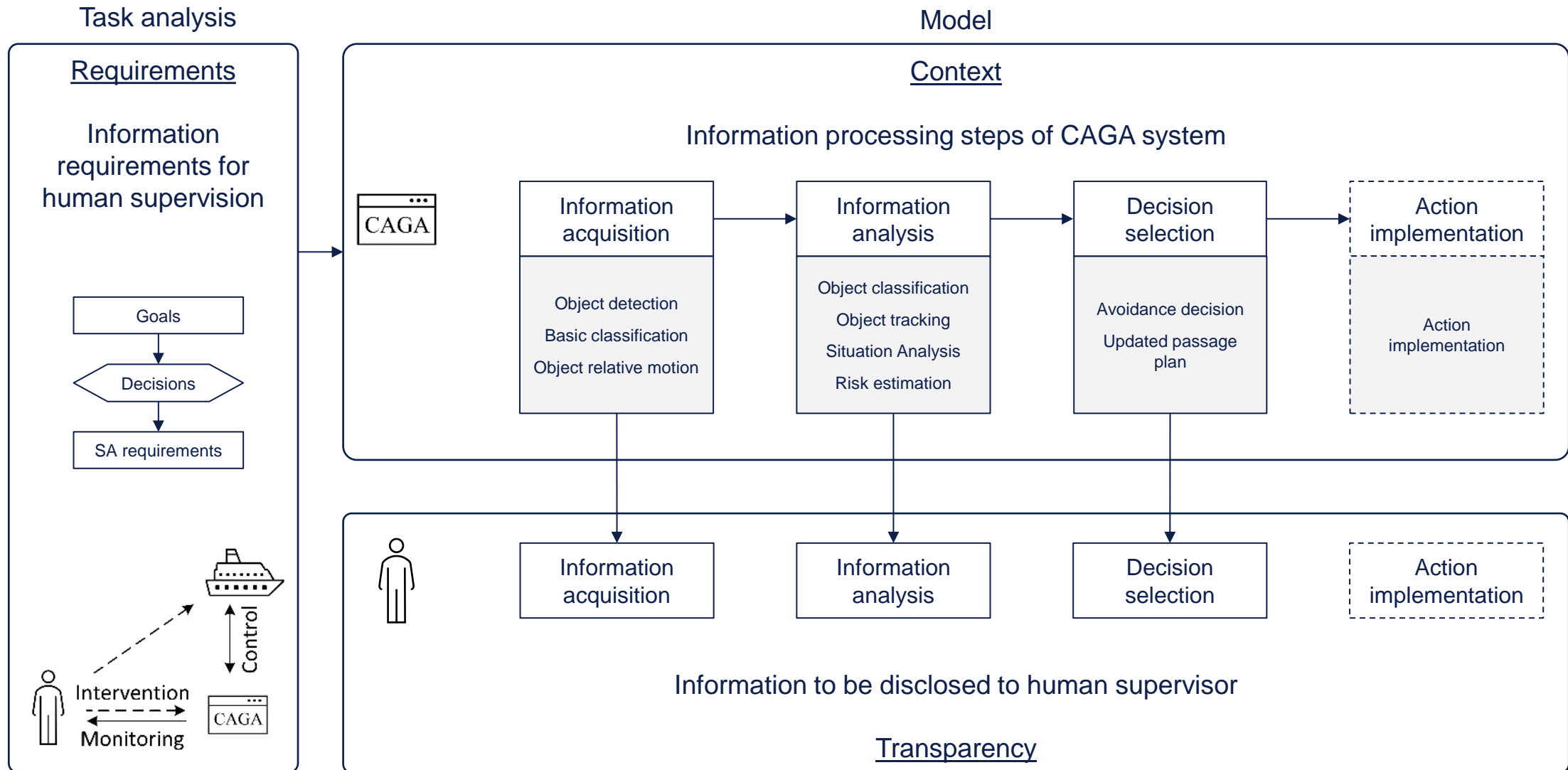
"...a promising effect of automation transparency on Situation Awareness and operator performance, without the cost of added mental workload, [...] where humans respond to proposals and where humans have a supervisory role".

Van de Merwe, Mallam, & Nazir (2024)

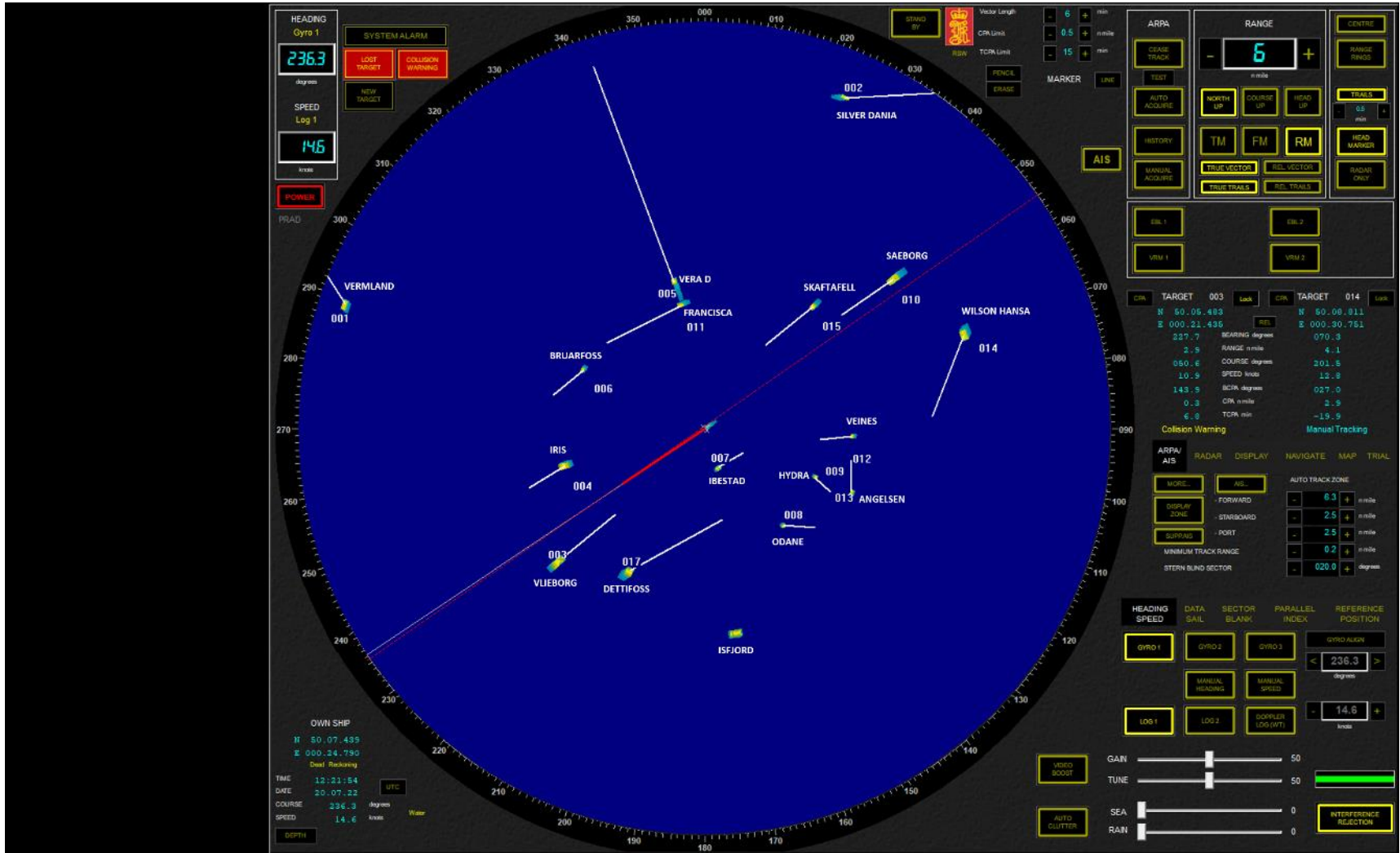
Collision avoidance systems?



From information requirements to transparency



Transparency level: None

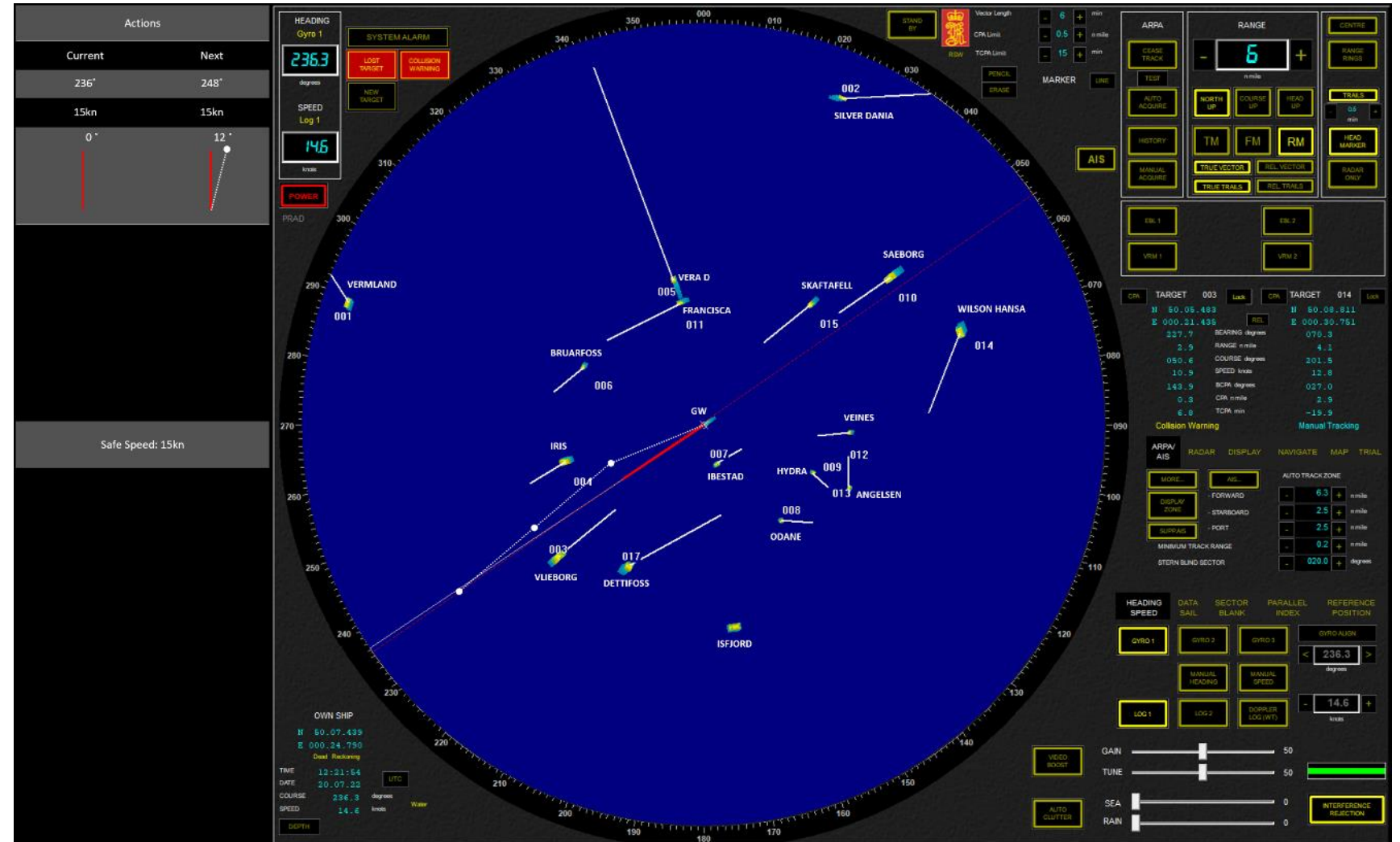


Information processing steps

Level of transparency	Information acquisition	Information analysis	Decision selection
Low			X
Medium (A)		X	X
Medium (B)	X		X
High	X	X	X



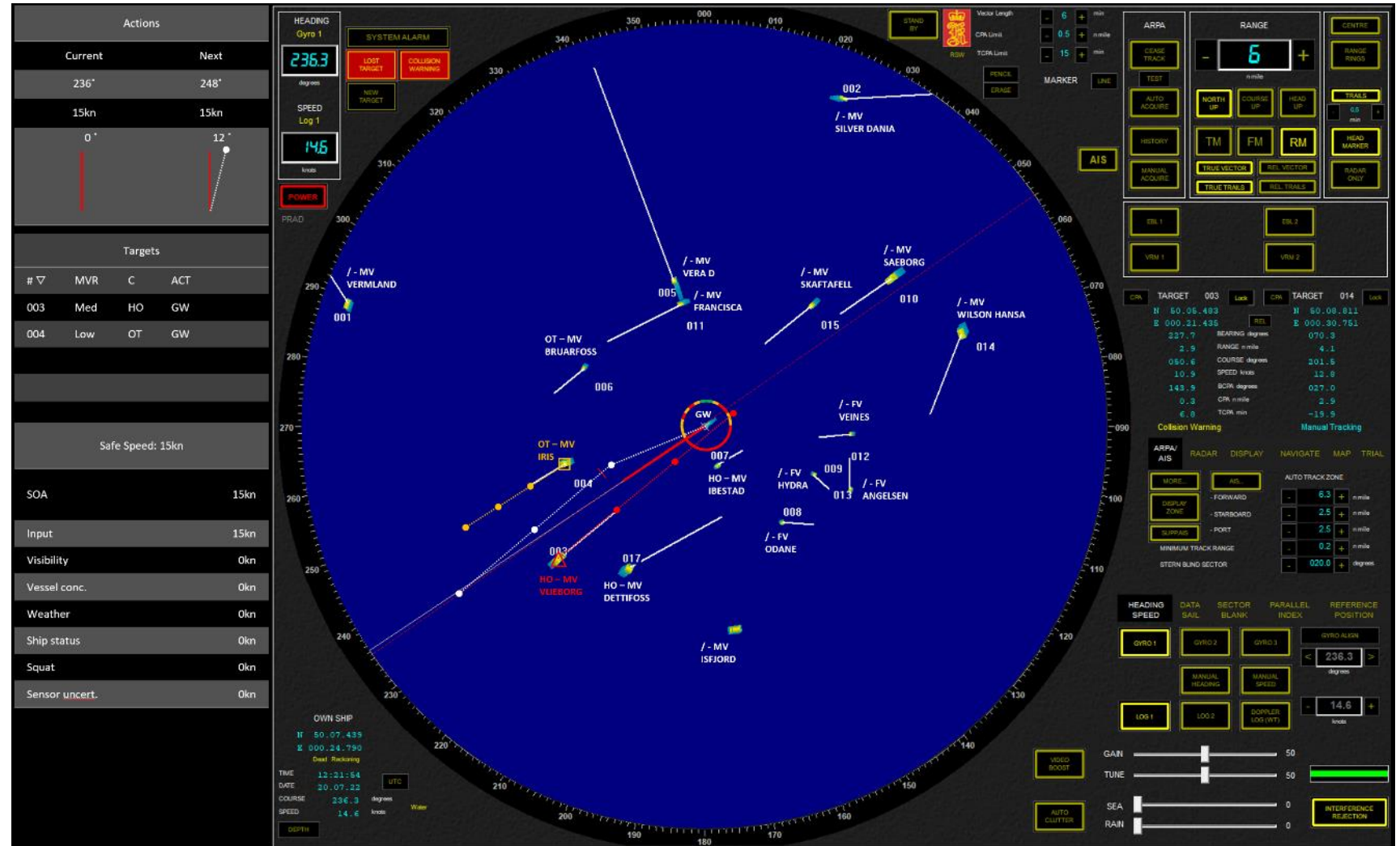
Transparency level: Low



Information processing steps

Level of transparency	Information acquisition	Information analysis	Decision selection
Low			X
Medium (A)		X	X
Medium (B)	X		X
High	X	X	X

Transparency level: Medium (A)



Information processing steps

Level of transparency	Information acquisition	Information analysis	Decision selection
Low			X
Medium (A)		X	X
Medium (B)	X		X
High	X	X	X

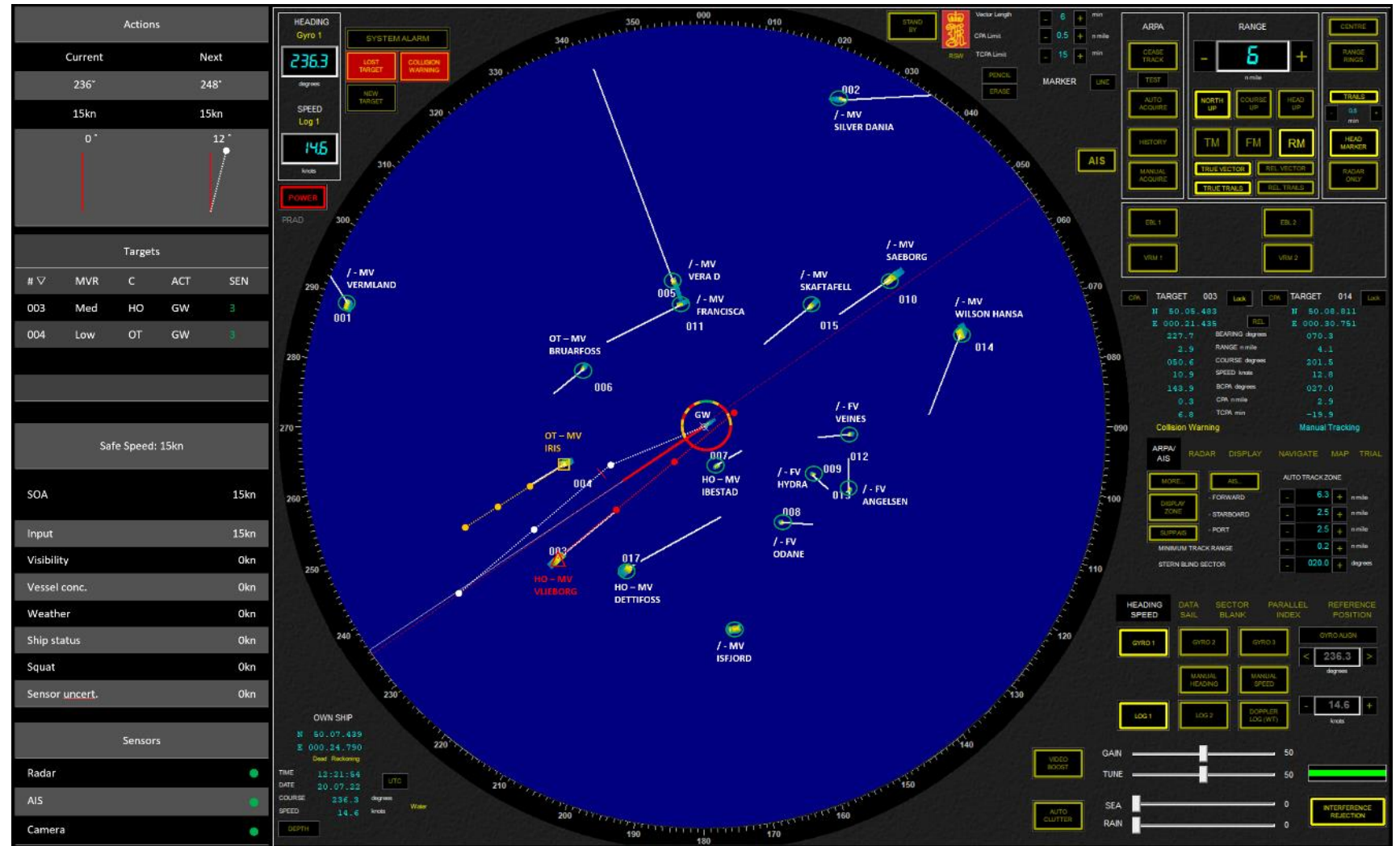
Transparency level: Medium (B)



Information processing steps

Level of transparency	Information acquisition	Information analysis	Decision selection
Low			X
Medium (A)		X	X
Medium (B)	X		X
High	X	X	X

Transparency level: High



Information processing steps

Level of transparency	Information acquisition	Information analysis	Decision selection
Low			X
Medium (A)		X	X
Medium (B)	X		X
High	X	X	X

Controlled experiment

- 34 (licensed) navigators
- Dependent variables
 - Transparency (4)
 - Complexity (2)
 - Within-subjects design
 - Randomised stimuli
- Dependent variables
 - Situation Awareness (SAGAT)
 - Task performance (Time)
 - Mental workload (NASA-TLX)
 - Preference (Ranking)
- Post-experiment interviews



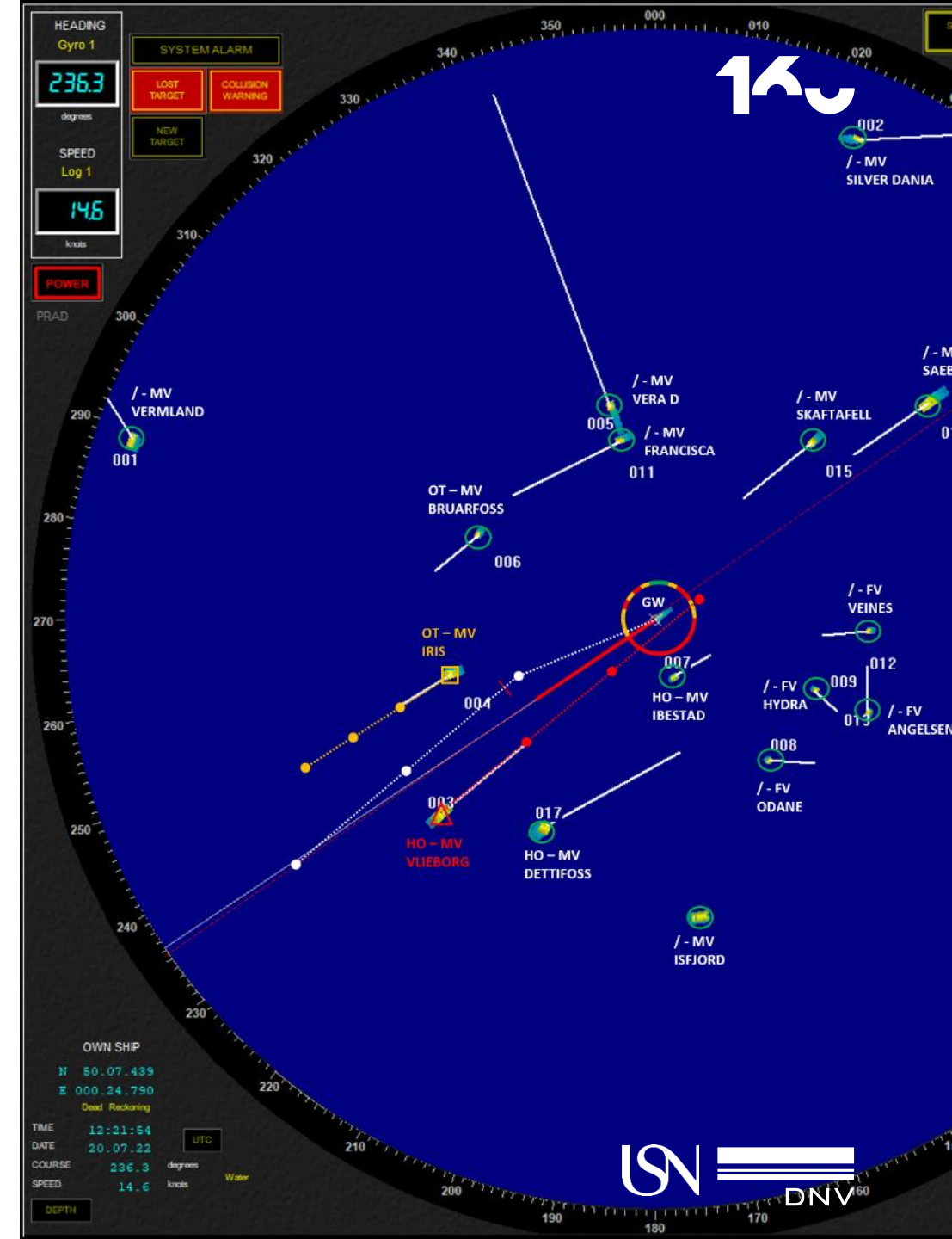
Experimental results

- Improved Situation Awareness with increased transparency
- Additional time needed to perceive and comprehend the traffic situations
- No effect on mental workload observed
- Participants clearly preferred the medium (A) and high transparency HMIs*, i.e. those that present the system's risk analysis
- Thematic analysis on interview data showed:
 - Insight into the CAGA system's real-time understanding of the situation
 - Insight into how the CAGA system understands the risk picture

Van de Merwe et al (2024b)

Van de Merwe et al (2024c)

*Human Machine Interface



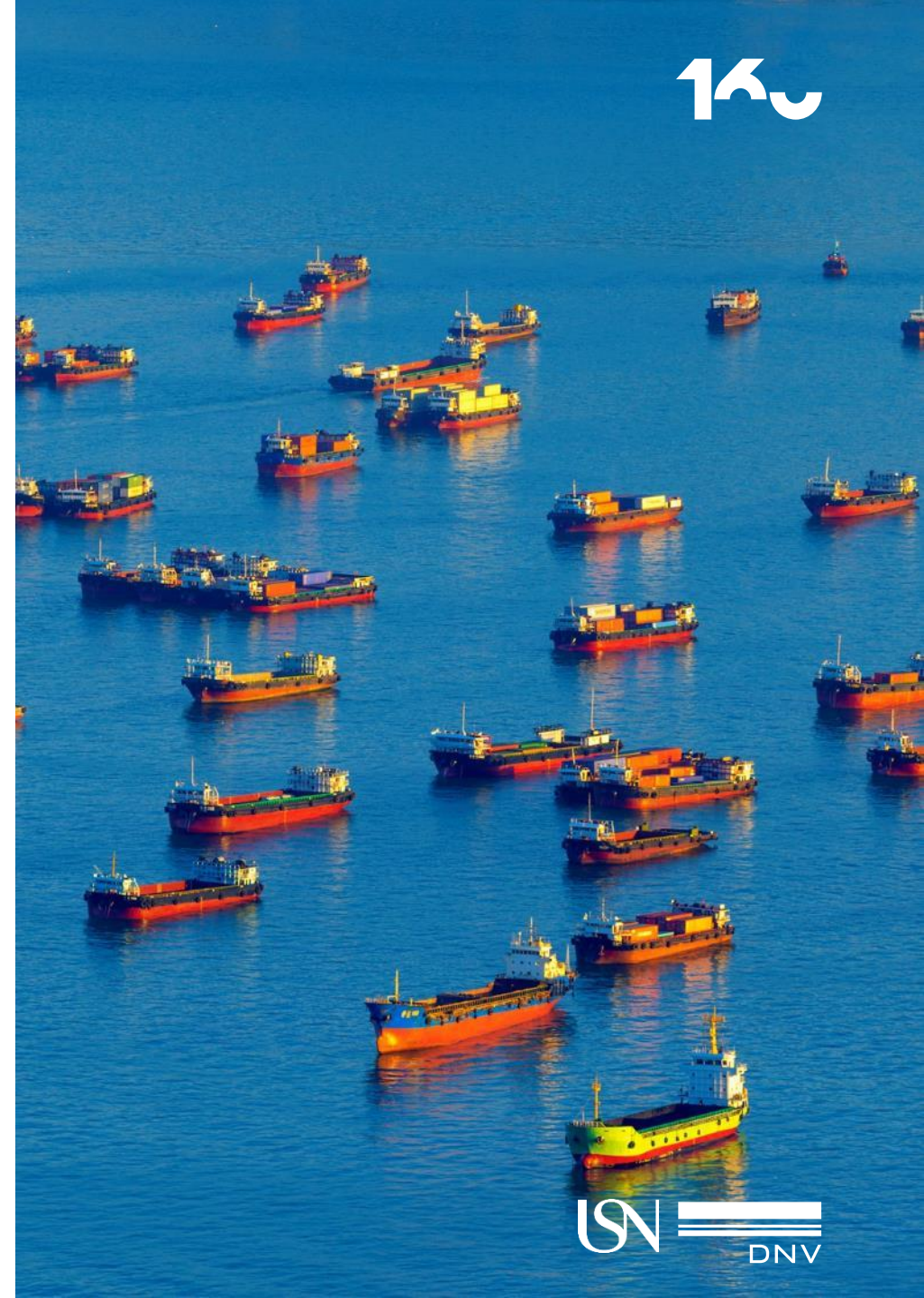
Experimental results

- Mixed methods approach illustrates the role of transparency to address out-of-the-loop performance problem:
 - Experimental results showed which LoT had effect
 - Medium (A) and High
 - Interview results showed why
 - Understand the ship's understanding of the situation
- Transparency and safety critical tasks
 - Ensuring safety is main responsibility of navigator
 - System's analysis = most critical information



Further work

- What are the limits of transparency?
 - Time-limited operations
 - Dynamic transparency
 - Integration of transparency information vs noise
 - Inclusion of grounding avoidance information
 - How to represent uncertainty?
- What is the role of transparency in human-AI teaming?
 - Improved task engagement with transparency?



Designing for Human-Autonomy systems



Automate only if necessary	Use automation for routine tasks, not high-level cognitive functions	Provide support for Situation Awareness, not decisions	Keep operator in control and in the loop	Avoid too many automation modes
Make modes and system states salient	Enforce automation consistency	Avoid advanced queuing of tasks	Avoid use of information cuing	Use methods of decision support to create symbiosis
Provide automation transparency	Ensure logical consistency across features and modes	Minimize logic branches	Map system functions to the user's goals and mental models	Minimize task complexity
Support comprehension of information (L2 SA)	Provide assistance for projections (L3 SA)	Use information filtering carefully	Support assessment of system confidence level	Support assessment of system reliability

Endsley, 2017

Endsley, Bolte, and Jones, 2003



Wrapping up



“This dissertation advocates the relevance of affording human operators with insight into the reasoning of autonomous systems and established transparency as an important prerequisite on the path towards safe and effective human-supervisory control” Van de Merwe (2024)



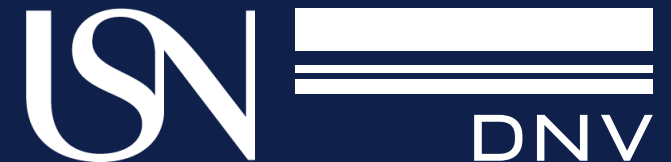


WHEN TRUST MATTERS

Thank you!

koen.van.de.merwe@dnv.com

www.dnv.com



References



- van de Merwe, K., Mallam, S., & Nazir, S. (2024). Agent Transparency, Situation Awareness, Mental Workload, and Operator Performance: A Systematic Literature Review. *Human Factors*, 66(1), 180–208. <https://doi.org/10.1177/00187208221077804>
- van de Merwe, K., Mallam, S., Nazir, S., & Engelhardtson, Ø. (2024a). Supporting human supervision in autonomous collision avoidance through agent transparency. *Safety Science*, 169, 13. <https://doi.org/10.1016/j.ssci.2023.106329>
- van de Merwe, K., Mallam, S., Engelhardtson, Ø., & Nazir, S. (2023). Towards an approach to define transparency requirements for maritime collision avoidance. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 67(1), 483–488. <https://doi.org/10.1177/21695067231192862>
- van de Merwe, K., Mallam, S., Engelhardtson, Ø., & Nazir, S. (2023). Operationalising Automation Transparency for Maritime Collision Avoidance. *TransNav, International Journal on Marine Navigation and Safety of Sea Transportation*, 17(2). <https://doi.org/10.12716/1001.17.02.09>
- van de Merwe, K., Mallam, S., Nazir, S., & Engelhardtson, Ø. (2024b). The Influence of Agent Transparency and Complexity on Situation Awareness, Mental Workload, and Task Performance. *Journal of Cognitive Engineering and Decision Making*, 18(2), 156–184. <https://doi.org/10.1177/15553434241240553>
- van de Merwe, K., Mallam, S., Nazir, S., & Engelhardtson, Ø. (2024c). Navigators' Perspective on Information Requirements for Supervisory Control of Autonomous Ships. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 10711813241261686. <https://doi.org/10.1177/10711813241261686>