Part III

Necessity of a protocol to enable teamwork between manned and unmanned ships

Meant is not said, said is not heard, heard is not understood, understood is not done.

– Marcus Rall [Rall and Dieckman (2005)]

It is not always possible to operate without communication, as is tried to accomplish in part II. This is often the case due to missing information, or a lack of understanding about the intentions of other vessels. With manned ships, verbal communication via VHF radio is used to acquire the missing information or discuss strategies with other ships. This communication is likely to be also necessary when unmanned vessel operate between manned vessel. This part discusses the development process of such a protocol. The relevance of such a protocol is proved using an experiment. This part will answer the following question:

> *Will a protocol based on existing maritime systems and communication protocols be sufficient to ensure safe navigation, while manned and unmanned vessels encounter each other?*

A protocol defines the format and the order of messages exchanged between two or more communicating entities, as well as the actions taken on the transmission or receipt of a message. Where it is mostly straightforward to use verbal communication, will it not be limited to this. As it might result in better situational awareness for both the manned and unmanned to use other means, such as visible signals or text messaging.

The communication which currently happens between vessels is most common when COLREGs do not result in clear strategies, or when intentions are not clear. Other communication which will not be within the scope of this research, is the communication with traffic controllers and how other vessels interpret conversations. Thus this research is a starting point to develop a full protocol needed for the acceptance of unmanned vessels.

This protocol is developed using an iterative process based on the situated Cognitive Engineering (sCE) method. Where reviews and prototype evaluations continuously refine a requirement baseline. How to apply sCE is described by Neerincx and Linderberg [Neerincx and Linderberg (2012)].

The first step is to create a foundation. The current situation of the problem is addressed. Thereby considering existing knowledge which might be relevant to solve the problem, which results in the envisioned technology. The next step is to define the system design specification. In which scenarios are described which show how the problem is solved. From this can be extracted what should be designed and why this is done. Using this a design is made which is being evaluated to make improvements in next iterative steps.

9 | Foundation

The foundation segment in the situated Cognitive Engineering (sCE) methodology describes the design rationale in terms of operational demands, relevant human factors knowledge, and envisioned technologies. Together, these three constituents describe the problem to be solved, the existing knowledge on ways to solve the problem, and the technology needed to implement that solution.

9.1 Operational demands

The operational demands describe the current practice as it is, i.e. without the envisioned technology. For the operational demands, the sCE method prescribes as main components the stakeholders with their characteristics and the problem description with an analysis thereof.

9.1.1 Problem scenario

Convention on the International Regulations for Preventing Collisions at Sea (COLREGs) have been developed long before bridge-to-bridge voice communication became available. They are supposed to be unambiguous. It is the responsibility of all bridge watchkeepers to know how to apply them instinctively, from observation by sight and radar. They work effectively when ships in an interaction obey them; they also specifically address circumstances where one ship does not.

However, as shown in the previous parts, are COLREGs not always sufficient to decide on the right strategy, for example, due to missing information. Problems with COLREGs happen already more often due to larger ships, more complex manoeuvers and more traffic. In those cases, operators can use the VHF radio for verbal bridge-to-bridge communication. Leading here are the Standard Maritime Communication Phrases (SMCP). The primary task of these phrases and surrounding protocol is to diminish misunderstanding in safety-related verbal communications. Beside this verbal communication, non-verbal communication might be used, such as light signals, sound signals and text messaging.

Three reasons for safety-related verbal communication are: As the operator you will deviate from the rules, you register deviant behaviour at other ships, or more information is needed to decide on the right strategy. Communication is often necessary due to the lack of visual information. For example, due to bad weather, obstacles like bridges and terminals, or the information received via AIS is not reliable. Besides the impact on the information you get by looking out of the window, is also the quality of the Electronic Chart Display Information System (ECDIS) and Automatic Radar Plotting Aid (ARPA) worse.

The different ways for communication are not developed to be used by unmanned vessels. On the other hand, it is not feasible to require all manned vessels to install new systems for communication, before introducing unmanned vessels. As this will require many more new regulations, development time and time to train seafarers.

Misunderstanding and problems with communication can be the result of changing the usage of systems over time, due to the evolving demands for operators. These changes can result in an information overload of the crew and communication channels, caused by the more frequent use of VHF. As it is a receiver or a transmitter, but can't be both at the same time, which means that in case two messages are sent at the same time, both senders will not receive the others message.

9.1.2 Problem analysis

To avoid misunderstanding which could result in hazardous situations. It is important that manned and unmanned vessels can communicate. A more extensive analysis is made to solve this problem. Describing the values of the different actors and discussing their related problems.

Primary actors

The focus of this research will be on bridge-to-bridge communication. For this communication are the most important actors for unmanned and manned vessels:

- Manned vessel
  - *Officer of watch*. He is the responsible person. He might work together with a helmsman and a lookout. He has to ensure proper functioning of all available systems. He does discuss with other crew members if there are any unusual activities. He is responsible for following a proper navigation plan while having his safe passage plan, to avoid collisions. He will use sight, Automatic Radar Plotting Aid (ARPA) and Electronic Chart Display Information System (ECDIS). Thereby is he aware of the ship's speed, turning circle and other handling characteristics to decide on the right strategy. He will monitor the VHF radio all the time while underway, to assist in emergencies if necessary, to hear Coast Guard alerts for

weather and hazards or restrictions to navigation, and to hear another vessel hailing you.

He wants to avoid information overload while being aware of the situation, which is only possible when he stays concentrated. This happens when the tasks are challenging, and he needs to have a form of autonomy [Porathe et al. (2014)].

  - *Helmsman and lookout*. Both monitor the situation and execute commands from the officer of watch. A risk for them is information overload or underload [Neerincx (2008)].

- Unmanned vessel
  - *Controller agent*. This agent is responsible for situational awareness. Thus getting safely from A to B. It will decide on the navigational strategy, it will do this based on the information acquired via all different means. Including newly developed communication protocols, computer vision and algorithms to transform sensor data into useful information. His duties are similar to the duties of the officer of watch as described for the manned vessel.
- Other vessels
  - *Crew and pilots on nearby vessels*. They might want to know the intentions of other vessels to base their strategy on, without receiving all discussions, as this might result in information overload.

Secondary actors

Beside the first group of actors, The new protocol could also influence others, besides the first group of actors mentioned above. Although they are not within the scope of this first design cycle, they should be considered to avoid problems such as information overload on current communication channels or confusion.

- Only recipients
  - Crew on vessels which are not travelling.
  - Shipowners of unmanned vessels, monitoring vessel from a remote location.
- Not within the scope of the research
  - Vessel traffic controllers
  - Crew which are in distress and require assistance

Goals

The main goal is to ensure reliable sharing of information, without the risk for information overload or misunderstanding, so that manned ships will trust unmanned ships to choose the right strategy, as manned ships can be informed, using natural language describing the reasoning of unmanned vessels. During communication should manned vessels only be updated when requested or in case of an unusual activity which could affect their strategy. Manned vessels should thereby be aware when unmanned vessels desire more information to decide on the right strategy. It might be possible to develop a protocol for communication with traffic controllers in later iterations, using the same philosophy. This communication becomes more critical when the development of a new system for traffic controllers also takes too much time to develop or implement.

Infeasible solutions

The easiest solution for unmanned ships would be to install a new system on every vessel. This is however not feasible as mentioned before. To implement this, it would mean that all ships which could encounter an unmanned ship will have to install this too. It might be possible to make it obligatory via regulations, which will cost much more time and money. Making the introduction of unmanned ships less likely. Time and money is also the reason to use a Non-visual User Interface (no-UI), as a GUI will require new screens or changes to the ECDIS which are only possible when regulations are changed.

9.2 Human factors

When designing technology, there are two driving questions that need to be well-thought out: (1) What tasks and/or values is the user trying to accomplish and how can the technology support the user in doing so?, and (2) How can the technology be designed such that the user can work with the technology?

The Human Factors segment of the sCE method describes the available relevant knowledge about, for instance, human cognition, performance, task support, learning, human-machine interaction, ergonomics, etc. Note that we emphasise that this knowledge should be relevant for the problem and its design solution: the knowledge described here should lead to a better understanding of either (1) or (2). The three elements relevant to the human factors analysis are the human factors knowledge, measures, and interaction design patterns.

Human factor knowledge

Human factors knowledge describes available knowledge coming from previous research about how to solve the problems that have been identified in the problem analysis. The key problems relevant for human factors are information overload, situational awareness, autonomy, and learning a new protocol. Thus the following questions should be answered:

- When does information overload occur?
  - In case of divided attention, there is a high risk for information overload and distraction by low priority messages. Therefore the developed system should be context-aware so it can limit this risk by adapting the message to the situation [Arimura et al. (2001)].
  - Overload might appear due to a competition for the operator's attention that is going on between different information items. If automated systems handle many tasks, the operator can deal with high workload circumstances but will suffer from severe underload during quiet periods, probably losing his or her situational awareness [Neerincx (2008)].
  - The information acquired at one particular moment does not necessarily serve for high-level situation awareness, for the user needs to recall the previous related information to understand the situation thoroughly. But constantly providing information might not be the solution because there will be a huge risk for information overload. Admittedly it is plausible to deliver needed information for the future task by task detection. The user might still fail to keep pace to the rapidly changing system and fulfil multi-threaded tasks[Porathe et al. (2014)].
- Which information do operators need for situational awareness?
  - Understanding the current picture is not enough for full situational awareness. Expert decision makers must be able to project their understanding into the future. This projection enables experts to make the decision which results in the best options in the future. Projection requires to have good mental models of the dynamic relationships between the relevant parts of the environment over time. Experts focus a lot on creating their futures via present decisions. In turn, experts do form these decisions out of their comprehension of the likely interactions of all the elements they deem both relevant and important [Gregory and Shanahan (2010)].
  - Situational awareness can be enhanced by feedback, perceived information from the environment, information from other agents, as well as remote sensors. [Carver and Turoff (2007)]

- How is information perceived when acquiring it passively or actively?
  - Attention profoundly modulates the activity of sensory systems, and this can take place at many levels of processing. Imaging studies, in particular, have revealed the greater activation of auditory areas and areas outside of sensory processing areas when attending to a stimulus [Palmer et al. (2007)].
  - Good teamwork involves anticipating the needs of teammates, and that means pushing information before operators request it. Therefore, if things are going well, there should be little need for pulling information. In this study task, participants were instructed to push information to others, and over time master the specific timing of information sharing to the intended recipient. Findings indicate that pushing information was positively associated with team situation awareness and team performance, and human-autonomy teams had lower levels of both pushing and pulling information than all-human teams [Demir et al. (2017)].
- What is needed for successful teamwork between human and a computer?
  - People need to understand what is happening and why when a teammate tends to respond in a certain way. They need to be able to control the actions of an agent even when it does not always wait for the human's input before it makes a move, and they need to be able to reliably predict what will happen, even though the agent may alter its responses over time [Bradshaw et al. (2003)].
  - Effective team communication, a fundamental part of team coordination, is crucial for both effective team situation awareness and team performance [Demir et al. (2017)].
- Do people trust automated systems?
  - When using automation, the role of the human changes from operator to supervisor. For effective operation, the human must appropriately calibrate trust in the automated system. Improper trust leads to misuse and disuse of the system. [Walliser (2011)].

Human factor measures

Measures describe how to operationalise the quality of the intended behaviour or performance, i.e. how well is a user working with the design able to reach his/her objectives and what is the quality of the collaboration between the human worker and the technology?

- Is the system used correctly?
- Will the protocol solve the problem of missing information?
- Does the protocol act as expected?
- What is the impact on attitude towards unmanned ships?

Interaction design patterns

Interaction Design Patterns (IDPs) focus on the Human-Computer Interaction (HCI), such as usable interface design and control options. IDPs offer generic solutions to recurring HCI design problems that have been proven to be effective. Relevant IDPs are given in table 9.1. Keywords are often seen as the new buttons to interact. For the new protocol are message markers the keywords. These make it easier to train the conversational agent, and more clear for operators on manned vessels what the options are. Whereas the conversational skills of the agent are the core of general communication, the usage of other methods of communication will improve redundancy and effectiveness. A multimodal conversational agent also includes visible signals such as masthead light signals, flags and AIS messages. These will enable operators to see immediately if the vessel is unmanned. The last two methods are for extreme situations: audible and distress signals. The agent for the unmanned vessel should understand what these mean and how to use them before manned vessels trust them.

Radio communication	Usage of message markers and conversational agent
Visible signals	Mast head signals, flags and AIS information
Audible signals	Horn and speakers
Distress, urgency and safety signals	Flares and smoke

Table 9.1: Interaction design patterns

9.3 Envisioned technology

The envisioned technology describes the available options of using existing technology and the need to develop novel technology to come to a system solution. The sCE method asks to specify what devices (hardware) and software the designers could use in the system design. In addition, for each type of technology, an argument should be provided as to why this technology might be of use and what the possible downsides might be of that particular type of technology.

The envisioned technology will use only existing systems to develop a no-UI. Different systems which are currently used, are described in appendix A. Below different systems and protocols are mentioned which can be used in the new protocol. Using these already existing systems will shorten the development, learning and implementation time. Table 9.2 gives the used systems, equipment and protocols.

Using already existing protocols makes it easier to learn, such as Standard Maritime Communication Phrases (SMCP) and COLREGs. These systems make it also recognisable, which means that

Radio communication	Conversational agent Negotiating agent Usage of message markers Availability on VHF Natural language variations on SMCP NATO phonetic alphabet and numbing Addressed AIS message to exchange information or interrogate
Visible signals	Light signals Mast head signals Flags Heading, position and movements
Audible signals	Horn Speakers

Table 9.2: Envisioned technology

users will understand the benefits quicker. Show that it is useful and easy to use, as this is key to the acceptance of technology [Davis (1989)].

The type and amount of information presented to users must be tailored to the unique situation in which users use the information. Prior research on trust in automation found that providing human operators with information related to the reliability of an automated tool promoted more optimal reliance strategies on the tool. Further, information related to the limitations of an automated tool aids in trust recovery following errors of the automation. This added information appears to be useful in deciphering the boundary conditions under which the tools are more or less capable. Thus, providing human operators with information related to the performance of an automated tool appears to be beneficial [Lyons and Havig (2014)]. Therefore it is beneficial for the cooperation between manned and unmanned vessels to show if it is an unmanned vessel, which will first be done using visible signals, and also at the start of radio communication. Telling the user that you are unmanned or automated also happens in industry projects, such as Google Duplex [Nieva (2018)].

In the next chapter the system design specification will be presented for the envisioned technology. Most important will be the definition of the usage of radio communication and how this should be supported by other forms of communication, such as visible and audible signals.

10 | System design specification

The system design specification describes the solution to the problem in the form of a system design that makes use of the identified relevant human factors knowledge, and the envisioned technology. The system specification consists of design scenarios, use cases, requirements, claims, and ontology. This chapter will answer the question: How a protocol should look like to ensure safe navigation when manned and unmanned vessels meet.

10.1 Design scenarios

The sCE method prescribes the specification of design scenarios. Design scenarios are short stories that provide a clear description of how the user will work with the technology thereby enjoying the solution offered to one of the problem scenarios. Together, the problem and design scenarios provide a contextualised view on:

1. The problem the design aims to solve.
2. The people that are currently affected by this problem.
3. The way in which the current system design aims to solve this problem.
4. How people will use the system.

Manned ships can understand intentions for unmanned vessels, resulting in good situational awareness. In cases they desire more information, they can acquire this by using existing systems. Without the risk for information overload. Thereby are the additions to existing protocols for those systems easy to understand, as they use the same philosophy as current protocols.

Trust in autonomous ships is formed, as the information is reliable, the interaction is similar to the interaction with other manned ships, which results in the acceptance of unmanned ships on the general waterways. Where the risk for collisions and perceived risk reduces or at least does not increase.

Information extraction from problem scenario

The previous chapter describes the problem which has to be solved with the envisioned technology. The following issues have to be tackled to solve the problem:

1. Different actors are afraid of information overload.
2. Officer of watch is afraid to lose situational awareness.

3. Officer of watch is afraid to lose autonomy.
4. Current systems are not designed to be used by unmanned vessels.
5. Manned ships want to ask for support or information.
6. Unmanned ships want to ask for support or information.

Envisioned effect of system implementation

How the problems as mentioned above are tackled is discussed below. This shows what the result is after implementing the envisioned technology:

1. The system will send only on demand or when it has tried any other solution, as this will reduce the probability for information overload. As a threshold to check if the system is successful in solving the problem of information overload, is the criteria the current amount of communication.
2. The protocols which are currently used by the officer of watch are the same. The purpose is to make it more easy to get information, which means that the situational awareness will minimally be affected by the system on board of the manned vessel. Also here the current level of situational awareness can be used as a threshold.
3. By introducing an negotiating agent, decisions can be made similarly to current ships. Using the decision tree, the agent will have a favourite strategy, but it might be possible to use others if this is better for the encountered manned ship. This choice will ensure that the officer of watch at the manned ship, still has a feeling of autonomy, similar to that of the current situation.
4. Although the current systems are not designed to be used by unmanned vessels. There are significant developments on conversational agents in the last few years. Voice communication is becoming easier to develop, especially when considering that a lot of VHF conversations are recorded. These make it more easy to train the conversational agent. The importance of transparent communication must thereby be acknowledged. Thus also being open about being a robot.
5. Using addressed AIS messages and a conversational agent at unmanned vessels. Manned ships will be able to ask for support or information at all time. It is even likely they will receive the information faster compared to manned-manned ship communication.
6. The conversational agent is most relevant for unmanned ships when they want to ask for support or information, as operators do not use addressed messages often at the

moment. They will use the SMCP in a similar way to how they use it right now. So not much will change compared to the current situation for manned ships.

10.2 Functional requirements and claims

The functional requirements and claims, describe specific functionalities the technology should provide to its users, this is followed by the system's objectives, and the hypotheses to be tested during system evaluations. All functional requirements are annotated with their underlying objectives (called claims).

This explicit linking of requirements to claims enables designers to formulate hypotheses that need to be tested in system evaluations to justify the adoption of the functional requirement. If the claim cannot be proven to be valid through system evaluations, the designers need to refine their system design, for instance, by trying to improve the functionality, replacing the functionality with a different one, or dropping the functionality and the claim altogether (i.e. by deciding that the objective is not reachable at this point). Either way, there is no use of including functionality that does not achieve its underlying claims. User stories are used to do this, these are usually in a form: "As an $\langle actor \rangle$, I want to $\langle what? \rangle$, so that $\langle why? \rangle$". Followed by acceptance criteria, to determine when this part is correctly implemented. The primary actors who will be taken into account in this first iteration are Officer of Watch (OoW), unmanned vessels, and nearby vessels.

User stories

- As an officer of watch, I want to know if there are unmanned ships in the area, so that I know what to expect from the communication. AIS shows if the ship is unmanned, and what its status is.
- As an officer of watch, I want to validate if the information received via AIS is correct, so that I can base my decision on accurate information. When an officer of watch asks unmanned ship via VHF, the answer should be reliable and based on the live information.
- As an officer of watch, I want to make my intentions clear towards all other ships, so that they can anticipate this. The agent should incorporate the shared intentions into the decision-making process.
- As an officer of watch, I want to be able to make small mistakes when following a protocol, so that I can still act fast when I do not know the exact SMCP sentence. The

unmanned vessel should understand natural variations in a message, compared to the SMCP sentence.

- As an officer of watch, I want to use existing protocols, so that the extra effort to communicate with unmanned vessels will be kept to a minimum. Current seafarers should be able to understand what they should be doing without an explanation on the protocol.
- As an unmanned vessel, I want to initiate communication, so that I can exchange information or ask questions to another specific ship. The unmanned ship needs situational awareness based on a digital model of the reality to know which vessel sails where and what the interactions could be to make the right decision.
- As an unmanned vessel, I want to validate if acquired information is accurate, so that I can base my decision on accurate information. Check via communication if the digital representation is accurate.
- As an unmanned vessel, I want to be able to check if they understand my intentions when other ships do not act as expected, so that I know if I should change my strategy. The communication should be incorporated in the decision-making process and unexpected actions by others ships should be registered.
- As a nearby vessel, I want to receive only information which is relevant to me, so that the risk for information overload is limited. By switching VHF channels for full conversations, this is similar to the current way of working.

10.3 Use case

Scenarios are used to create more specific descriptions of step-by-step interactions between the technology and its users (i.e. use cases). Use cases include actors, to specify which stakeholders/agents are interacting with each other in a given action sequence. Use cases make the design scenario more concrete by describing exactly how technology makes sure that the problem is solved. Use cases are informed by human factors theories as described in the previous chapter.

The purpose of the use case, as described in this section are: To give insight in all interactions during a common critical situation at sea. Tags are used to relate them to the situations and scenarios as described in chapter 3. By making it very specific, better insight is acquired in factors which should be taken into account when defining functional requirements.

Autonomous fast crew supplier crossing shipping lane in front of cargo ship

Tags: Crossing, Move away from other, Evasive manoeuvrer now, Crossing distance, CPA, Intention, Messaging

A 26-meter autonomous fast crew supplier (FCS2610) is heading towards a wind farm at the north sea with a speed of 22 knots. To get there, she has to cross a busy traffic lane. In which she will pass a 150-meter container ship (Reefer), sailing at 14 knots. The FCS2610 has noticed the Reefer late and has to make an evasive manoeuvrer, to pass in front of the Reefer with a passing distance of 900 meters or 0.5 Nautical miles, which is just accepted according to the safety domains [Szlapczynski and Szlapczynska (2017b)], using criteria from chapter 5. Communication is necessary to ensure the Reefer understands the intentions of the FCS2610. This will take place in the following manner:

- The AIS, masthead and flags are showing the vessel is sailing autonomously, which means there is no crew, but the autonomous systems listen to the VHF.
- A conversation is started by the FCS2610, calling the station on board of the Reefer and updating status in AIS to communicate intention.

    Reefer, C-6-Z-G-7
    Reefer, C-6-Z-G-7
    This is unmanned FCS2610, 2-F-F-P-4
    Unmanned FCS2610, 2-F-F-P-4
    Switch to VHF channel seven-two
    over.

- The FCS2610 waits for a response from

    Unmanned FCS2610, 2-F-F-P-4
    This is Reefer
    Agree VHF channel seven-two
    over.

- At VHF channel 72, FCS2610 communicates her intentions.

    Reefer, C-6-Z-G-7
    This is unmanned FCS2610, 2-F-F-P-4
    Intention. I intend to pass in front with a distance of 0.5 Nautical mile.
    over.

- At VHF channel 72, Reefer confirms intention.

Unmanned FCS2610, 2-F-F-P-4
This is Reefer, C-6-Z-G-7
Intention received. You intend to pass in front. Distance is 0.5 Nautical mile.
over.

- Close communication and pass in front.

Reefer, C-6-Z-G-7
This is unmanned FCS2610, 2-F-F-P-4
Nothing more. Have a good watch.
Over.

Unmanned FCS2610
This is Reefer
Thank you.
Over and out.

- Update AIS status of FCS2610 to show it has no questions, and is listening.

10.4 Specification of terms used in protocol

Lastly, the sCE method prescribes the construction of an ontology, i.e. a vocabulary describing a common language to be used throughout the system specification to avoid miscommunication, misunderstanding, and inconsistencies. Furthermore, the ontology can serve as the basis for the technology's data structure. By specifying important concepts in the ontology and also choosing to use only one word instead of various ambiguous synonyms, communication becomes clearer, and misunderstandings can be reduced to a minimum. The terms specified in the ontology are consistently used throughout the entire project. For this project, they are categorised in status, messages and situations.

10.4.1 Status

The system will know which functions and protocols it should execute, by defining different states for the system. The list below describes the different states:

Listening Listening to the radio without taking action.

Waiting Waiting for a response by other ship.

Negotiating Deciding on the right strategy by discussing this with other ship(s).

Messaging Sending a message. While sending it is not possible to receive a message.

Updating Adjusting the information stored within the system, which will consecutively be sent to others ships via AIS.

Unavailable There is a problem with the system, which makes it unable to communicate.

These states will also be communicated via AIS, to ensure transparency between different agents and avoid confusion.

10.4.2 Types of messages

Both in the messaging and negotiation states are messages send. These messages form the conversation. The agent will send different messages, during the phases of the conversation. The types of messages are described below:

Call Start of conversation, in which a ship only requests contact with another ship.

Acknowledge Accept the invitation for conversation.

Message Starts with "marker word" to clarify communication purpose, followed by the actual message and ended by a request for confirmation. The SMCP use the following marker words: *advice*, *information*, *warning*, *intention*, *question*, *instruction* and *request*.

Response Response to the previous message in the conversation.

Close End conversation with a greeting.

The Standard Maritime Communication Phrases (SMCP) will be used, as this is a known protocol for seafarers. This protocol has its ontology. This is described by IMO in the regulations [International Maritime Organization (2000)]. A summary of the SMCP can be found in appendix A.1.4.

10.4.3 Speech acts

As can be seen in the use case in section 10.3, does a conversation contain several steps, which relate to the different message types. Where every message within the conversation ends with *over*, and the other vessel should answer with a response.

The speech act theory of Austin [Austin (1975)] is used to validate that each message and conversation are useful. As communication should be kept to a minimum, which means that every message sent, should have a locutionary, illocutionary and perlocutionary speech act.

Where the locutionary part is the sound, the illocutionary part is the intrinsic message, and the perlocutionary speech act aims to trigger an action. Thus in case of the locutionary act, *What are your intentions?*. The illocutionary message is: *I want you to tell me that your intentions are*, while the perlocutionary aim is that the other vessel will say what its intentions are. Below the different **perlocutionary**, illocutionary, and (locutionary) acts in a conversation are shown. It should be considered that the MESSAGE itself also has these various acts.

Other ship pays attention I want other ship to listen to me.
(${<}\, name\ \&\ call\ sign\ other\ vessel \,{>}$, ${<}\, name\ \&\ call\ sign\ other\ vessel \,{>}$. This is ${<}\, name\ \&\ call\ sign\ own\ vessel \,{>}$)

Other ship switches to right VHF-channel I want to communicate via specific VHF-channel.
(Switch to VHF channel ${<}\, channel \,{>}$)

Other ship takes action or shares information I want other ship to understand the message.
(${<}\, marker\ word \,{>}$, *MESSAGE*)

Other ship closes message too I want other ship to know, the conversation has ended.
(over and out)

To make it more clear what the illocutionary act is, does the SMCP define seven marker words, these introduce the content and purpose of the communication. The marker word is placed in a message after calling the other vessel and introducing yourself but before the real message. Examples of messages with different marker words are shown below:

- *Advice*. Stand by on channel 6 - 8.
- *Information*. The fairway entrance is: position: bearing 1-3-7 degrees true from North Point Lighthouse, distance: 2 decimal 3 miles.
- *Warning*. Buoy number: one - five unlit.
- *Intention*. I intend to reduce speed, new speed: eight knots.
- *Question*. What are your intentions?
- *Instruction*. You must alter course to starboard.
- *Request*. Immediate tug assistance.

11 | Design evaluation

The last part of the sCE method is the design evaluation. The design evaluation aims to test and validate the system's design, such that the current design can be improved in incremental development cycles.

The evaluation method will be an experiment, where participants have to decide on the actions in a simulation environment, together with a questionnaire related to the experiment and communication in general. The participants are experienced seafarers. The interviews aim to answer the following question:

> *Will the described protocol ensure safe navigation and more situational awareness, when manned and unmanned vessels encounter each other?*

Different measures are used for validation and verification. Key variables are performance, trust, situation awareness and satisfaction. This chapter describes these in more detail. This chapter aims to test if the earlier described protocol, will indeed result in more situational awareness. The situational awareness should on its turn result in safer navigation.

The performance variable measures if the protocol does not influence the decision-making negatively, thus do the participants follow COLREGs and use SMCP correctly. This is validated by looking at the CPA and questioning the participants about their reasoning. The main question is: Does the protocol influence the performance of seafarers?

The trust variable is about the confidence of seafarers in the system. The protocol will only work effectively when seafarers trust the protocol. The aim of this first iteration is to find out what worries the participants, and if they want to cooperate. Based on this can be answered: Are seafarers confident that the protocol will act as they expect?

The third variable is situation awareness, as this should not be influenced negatively by the protocol. This means that seafarers predict future states correctly and are aware of everything happening around them. This is rated by the participants themselves using an observer rating system, and by questioning them on their awareness of key characteristics of other vessels (e.g relative speed, colour, course changes). This answer the question: Has the protocol a negative impact on the situation awareness?

The last variable is satisfaction, seafarers should like to use the protocol, as this is necessary to ensure that seafarers indeed use the protocol. Thus answering the question: Do seafarers like to use the designed protocol?

Two situations will be simulated to get relevant feedback and answer the above mentioned questions. The situations are based on the accident reports as described in appendix C, everyday situations around the port of Rotterdam and cases used in literature. The situations are simulated and visualised using the tool as described in appendix B. This visualisation will enable the experts to gain situational awareness and give useful feedback on the protocol. The protocol itself is mostly knowledge-based and not automated during the evaluation. Thus the interviewer has to know the Standard Maritime Communication Phrases (SMCP) relevant to the experiment, and usage of systems like Automatic Identification System (AIS) and Automatic Radar Plotting Aid (ARPA).

11.1 Evaluation method

The evaluation method is a so-called Wizard of Oz evaluation. This technique enables unimplemented technology to be evaluated, by using a human to simulate the response of an automated system. As the technology itself has not yet been implemented. The "wizard" simulates the system responses in real-time. Using seafarers and the Wizard of Oz method, an expert evaluation can be acquired on the proposed protocol without implementing it.

11.1.1 Experiment set-up

A participant is needed as Officer of Watch (OoW), and tools to execute the experiment. During the experiment, different variables will be tested. The description of these is part of the experiment design and is described in this section.

Participants

The 16 participants in this experiment are classified, based on their experience and expectations from autonomous shipping. All participants are Dutch, but their experience differs in certification, types of vessel they have operated, and their years of experience at the bridge.

Officer in charge This category has the lowest ranked officers at the bridge. In this case, they were trained to be officers of watch but only had limited sailing experience. Therefore they are not yet allowed to be chief mate or master. They are currently studying at STC-Rotterdam. The advantage of this group is that they studied SMCP and COLREGs and worked as an officer of watch, within the last year.

Chief mate is head of the deck department of a merchant ship. He is responsible for the deck crew and cargo. He reports to the master or captain. The officer directs the helmsman to carry out a course or speed change.

Master is the highest ranked officer. He is ultimately responsible for the safety and security of the ship. The master ensures that the ship complies with company policies, local regulations and international laws. The captain is ultimately responsible for aspects of the operation, including the safe navigation of the ship.

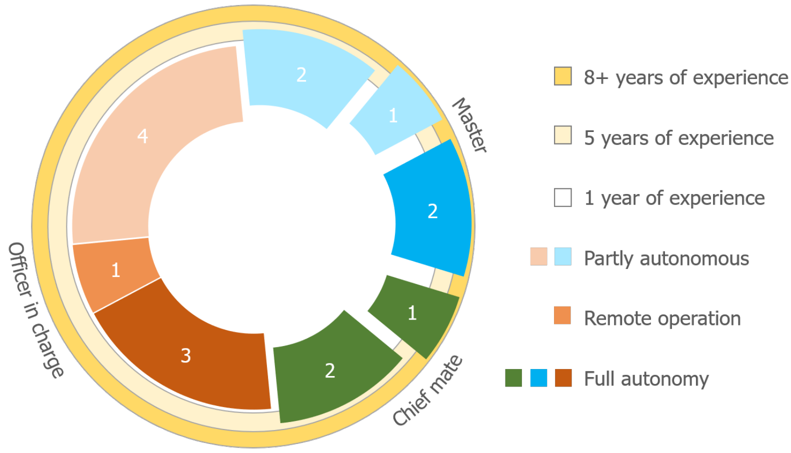


Figure 11.1: Classification of participants based on experience

The classification of participants is shown in figure 11.1. The different colours indicate the participants highest rank according to their certification. The maximum radius corresponds to their years of experience, while the shading the autonomy participants think ships will eventually have. The lightest colour is for participants who expect vessels to be partly autonomous, which means that they expect there will always be crew on board. Remote operation means that they expect ships to sail without a crew, but that every ship will be monitored 24/7 from a remote location, where the operator can intervene at any time. The darkest colours show the participants who think ships will eventually be able to operate fully autonomous, even in the busiest regions, such as the Dover and Malacca Strait.

The number of operators and their role per ship-type is shown in figure 11.2, some participants sailed on different ship types. Tugs and fast crew suppliers are under small vessels. Complex workboats are crane barges and dredgers. Coasters and general cargo vessels are most common. Ro-Ro and Ferries are more complex than the coasters and often also a bit longer. Cruise ships bring much more responsibility due to the number of passengers, resulting in larger safety domains. Tankers are the least manoeuvrable vessels operated by one of the participants.

Tools

Beside the participants, tools are needed to do the experiment. The tools needed are:

- Screen to show the simulation environment

- Questionnaire to be used before, during and after the experiment
- Room without distractions to do the experiment
- Possibility to store and later process actions during the experiment

The simulation environment is discussed in more detail in appendix B. Figure 11.3 shows the simulation environment. The environment has the map on the left, the side-bar or arrow-keys can be used to control the selected vessel. The status bar at the bottom of the screen gives information on possible actions, errors and status of simulation.

Experiment design

The experiment will consist of an interview and two assignment. In only one of the two assignments, the participant is allowed to use the protocol. This set-up means that the experiment is a within-subject design. The most significant benefits of this type of experimental design are that it does not require a large pool of participants. A within-subject design can also help to reduce errors associated with individual differences. A major drawback which should be taken into account during the experiment is that the result of the first assignment may influence the result of the second assignment. A problem which is known as the carryover effect. This effect is mitigated by counter-balancing the participants. Thus some are allowed to communicate using the protocol during the first assignment, others during the second.

Thereby are variables measured on different levels. The same set of questions is used for this both times. The questionnaire uses both questions with a linear scale and open questions. The linear scale questions are on an ordinal level, it is not possible to do calculations, but making a histogram of the results will show the tendency of the participants. The level of measurement for the open questions is at a nominal level. Conclusions are drawn based on the grouping of questions and answers per subject.

11.1.2 Experiment task

The participant will act as an Officer of Watch (OoW) in two situations. The duties of the OoW are to keep watch and navigate the vessel. He is the representative of the ship's master while keeping watch on the bridge, and has the total responsibility for safe navigation. This responsibility means that he has to follow a proper navigation plan to avoid any collision according to COLREGs. He is thereby aware of ship's speed, turning circles, and ship handling characteristics. He also communicates with other vessels when that is necessary.

More specific will he sail a vessel in different scenario's. To answer the research question for this chapter, will there be two cases. One in which the participant can communicate using the

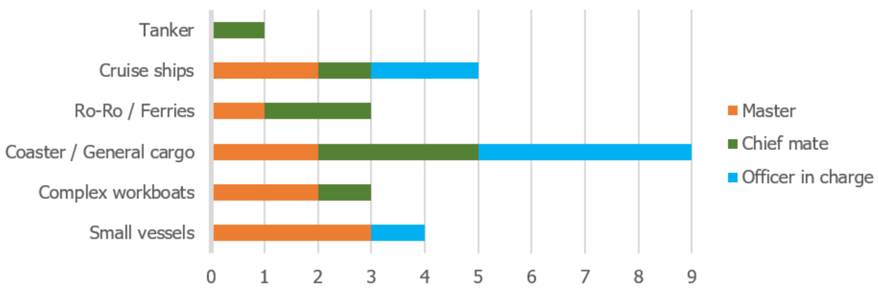


Figure 11.2: Ship-types operated by participants

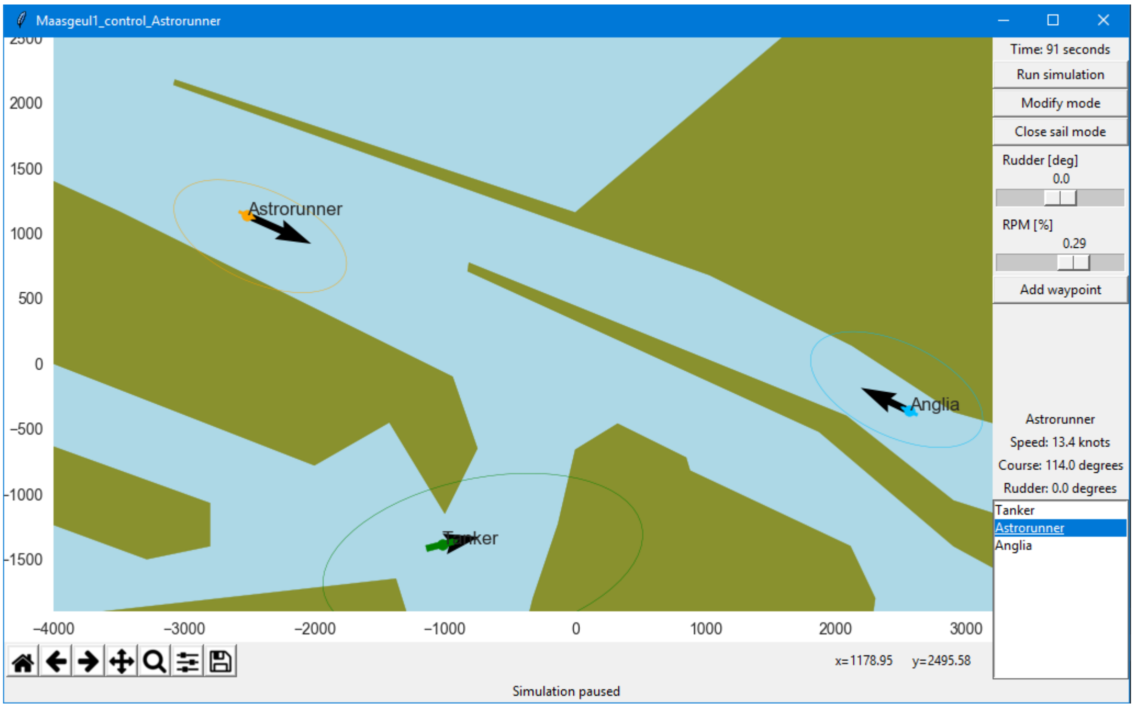


Figure 11.3: Simulation environment

designed protocol, in the other scenario he is not able to communicate. During each scenario will he direct a vessel on a 2D-map, as shown in the simulation environment. Within the experiment, tests will be done for different situations and scenarios. Each participant is thus able to communicate in either situation 1 or situation 2. The resulting strategies are listed below.

1. Crossing situation at North-Sea
   (A) Follow COLREGs strictly
   (B) Cross in front
   (C) Cross at the back
2. Entering Maasgeul from Maasvlakte
   (A) Cross in front
   (B) Cross at the back
   (C) Pass without crossing

Crossing situation at North-Sea

The first situation for the experiment is a crossing situation based on the accident between MV ARTADI and MV ST-GERMAIN (appendix C.2). Where both ships followed COLREGs, but due to a lack of communication and wrong presumptions on the intentions, did the accident occur.

The traffic in this simulation consists of three ships: a 250-meter tanker (GULF VALOUR), a 140-meter cargo vessel (ASTRORUNNER) and a 400-meter container vessel (EMMA MAERSK). Figure 11.4 shows the situation. The relevant information for these ships is given in table 11.1.

An example is given below for the communication in the situation when operating the ASTRORUNNER while crossing the GULF VALOUR at the North-Sea:

- Call

  Gulf valour, Gulf valour. This is Astrorunner. Over.

- Acknowledge

  Astrorunner. This is Gulf Valour. Over.

- Message

  Gulf Valour. This is Astrorunner.
  Question. Is your intention to alter course to Starboard?. Over.

- Response

  Astrorunner. This is Gulf Valour
  Question received. We will alter course to give way and pass one nautical mile astern. Over.

- Close communication.

  Gulf Valour. This is Astrorunner.
  Understood. Have a good watch. Over.

  Astrorunner. This is Gulf Valour. Over and out.

	GULF VALOUR	ASTRORUNNER	EMMA MAERSK	
Length	249.0	141.6	397.7	m
Width	48.0	20.6	56.4	m
Draft	13.2	6.5	12.6	m
Deadweight	114900	9543	156907	ton
Type	Oil tanker	General cargo vessel	Container vessel	
Position	[-1400, -1400]	[3250, 0]	[400, 2400]	m
Speed	16.0	15.2	12.0	knots
Course	45	278	225	degrees
Previous port	Singapore	Zeebrugge	Rotterdam	
Next port	Rotterdam	Dover	Hongkong	
Direction	North-east	West	South-west	

Table 11.1: Relevant information for crossing situation at North-Sea

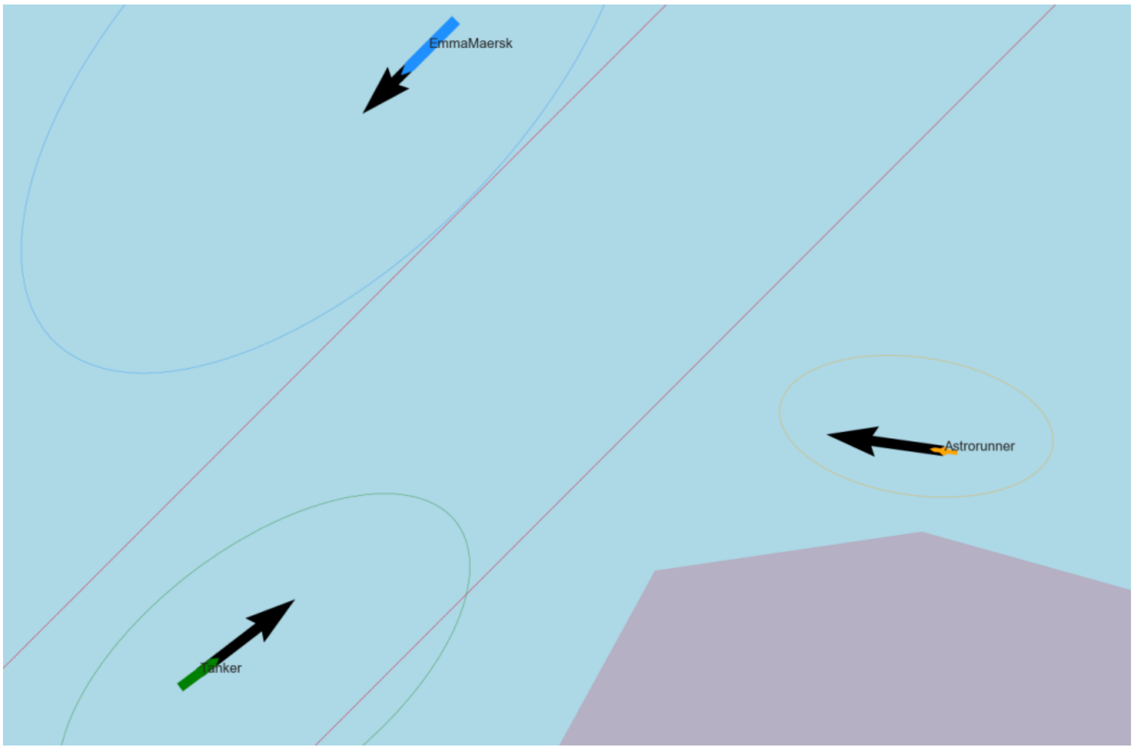


Figure 11.4: Situation sketch for crossing situation North-Sea

Entering Maasgeul from Maasvlakte

The second situation is a common situation at the port of Rotterdam. The situation is based on the description by Pilots from 'Nederlands Loodswezen'. The big challenge here is that ships are accelerating and decelerating. Therefore do traffic controllers notify ships about others intentions. But in the case as presented does this not always happen, or too late.

The traffic in this simulation consists of three ships: a 250-meter tanker (GULF VALOUR), a 140-meter cargo vessel (ASTRORUNNER) and a 140-meter Ro-Ro vessel (ANGLIA SEAWAYS). Figure 11.5 shows the situation. The relevant information for these ships is given in table 11.2.

The following conversation is likely for the situation where the GULF VALOUR is leaving the port of Rotterdam and has to cross or pass the ASTRORUNNER:

- Call

  Astrorunner, Astrorunner. This is Gulf Valour. Over.

- Acknowledge

  Gulf Valour. This is Astrorunner. Over.

- Message

  Astrorunner. This is Gulf Valour.
  Question. What is your port of destination? Over.

- Response

  Gulf Valour. This is Astrorunner.
  Question received. My port of destination is the Vulcaanhaven. Over.

- Message

  Astrorunner. This is Gulf Valour.
  Instruction. You are the stand-on vessel and should keep course and speed.
  Over.

- Response

  Gulf Valour. This is Astrorunner.
  Instruction received. We will stand on. Over.

- Close communication.

  Astrorunner. This is Gulf Valour.
  Nothing more. Have a good watch. Over.

  Gulf Valour. This is Astrorunner. Over and out.

	GULF VALOUR	ASTRORUNNER	ANGLIA SEAWAYS	
Length	249.0	141.6	142.4	m
Width	48.0	20.6	23.0	m
Draft	13.2	6.5	5.0	m
Deadweight	114900	9543	4650	ton
Type	Oil tanker	General cargo vessel	Container vessel	
Position	[-1372, -1377]	[-3090, 1395]	[3000, -550]	m
Speed	7.8	13.4	10.3	knots
Course	98	114	291	degrees
Origin	Princess Arianehaven	North-Sea	Vulcaanhaven	
Destination	North-Sea	Beneluxhaven	North-Sea	
Direction	Leaving	Entering	Leaving	

Table 11.2: Relevant information entering Maasgeul

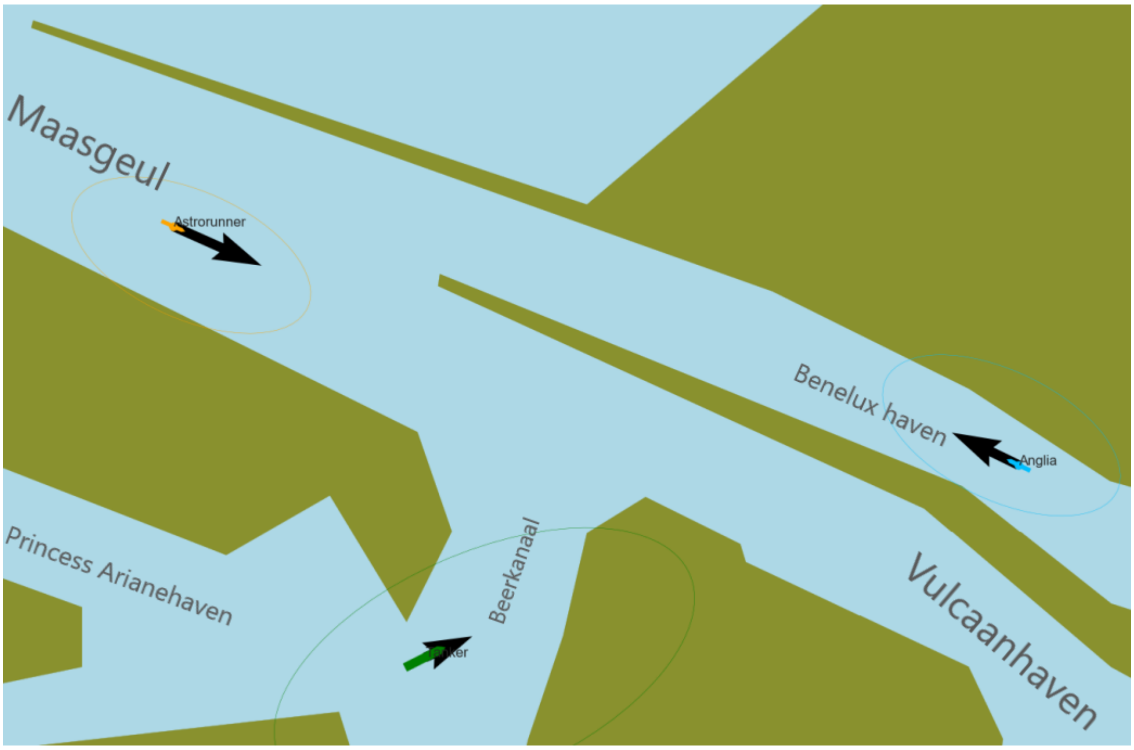


Figure 11.5: Situation sketch port of Rotterdam

11.1.3 Dependent variables

During the experiment, different variables will be evaluated. These are based on the human factor measures, as described in section 9.2. To answer research questions during the experiment, will a combination of both quantitative and qualitative measurements take place. Thus combining numerical values with non-numerical arguments. The measurements are done during the experiment via an interview and observations. Using the variables can be concluded if the system acted as expected and will result in safe navigation when using existing protocols. In the next section will these variables be linked to questions asked during the experiment, using the symbols as shown behind the variable name. The dependent variables within the experiment are:

Performance ♣Evaluation if the participant operated safely. Thus did he safely navigate the vessel by making the right decisions, which means that the participant followed the applicable COLREGs and showed good seamanship. The resulting closest point of approach (CPA) is a good measure for this. Also, the reasoning and situation recognition which results in choosing the right strategy is a measure. Questions which will be answered are:

1. Does the participant follow COLREGs?
2. Does the participant stay well clear of other ships?
3. Does the participant communicate using SMCP?

Trust ♢The system does not act only out of self-interest but to acquire a pareto-optimal solution. The participants must have a feeling of confidence that unmanned vessels operate as they expect. Therefore they must be confident that the system works as it is supposed to do. In later stages, this will also be supported by evaluations of reputable institutions. [Ozawa and Sripad (2013)]. For now, the survey is leading here, where questions will be asked on the participant's trust in autonomous systems, their trust in SMCP and how they thought the communication went during the experiment. Questions which will be answered are:

1. Is the participant confident that the system works?
2. What worries a participant when using the system?
3. How well does a participant trust unmanned vessels compared to manned vessels?

Situation awareness ♠The perception of environmental elements and events with respect to time or space, the comprehension of their meaning, and the projection of their status

[Naderpour et al. (2016)]. The situational awareness is measured using an observer rating system, showing if participants have noticed changes in course, the colour of different vessels and relative speed. Questions which will be answered are:

1. Does the participant predict future states correctly?
2. Is the participant aware of other vessels (e.g. speed)?
3. Has the participant free cognitive capacity (e.g. colour)?

Satisfaction ♡The participant should like to use the protocol. This is measured by questioning them on the effectiveness of SMCP, observe their usage of the protocol during the experiment and their reaction to vessels using the protocol. Questions which will be answered are:

1. Does the participant enjoy using SMCP?
2. Does the protocol change the behaviour of participants?

During the experiment these questions are answered, some are asked directly to the participants. Other answers are based on the behaviour of the participants. The answer to all of these question is on a nominal level, as these are all "Yes" or "No" questions. However, the results of the questionnaire are on an ordinal level, as these answers explain the "why". For situational awareness and performance are also higher level measurements used. Such as the metrics CPA and runtime, which are on a ratio level. Estimating the speed of vessels is on an ordinal scale measured. Statistics for those metrics will however not give more insights in the dependent variables, as there are too many factors influencing these results, such as the chosen strategy and accepted risk.

11.1.4 Experiment procedure

To execute the experiment. Several steps are taken together with the Officer of Watch (OoW):

1. Explain how the OoW can take actions, such as steering, change speed, set way-points or engage in communication.
2. Ask general questions on attitude and basic information.
3. Explain the situation to OoW in a similar way, to a usual watch hand-over. Only discussing relevant issues for navigational duties.r
4. Start simulation.

5. Depending on the simulation, let autonomous ship take actions or wait for the OoW to engage in communication.
6. End simulation.
7. Question OoW why he made decisions.
8. Evaluate the simulation.
9. Repeat step 3-8 for more situations.
10. Question OoW about advantages and challenges of the protocol.

The answers to the questions are collected using a Google form. This form includes the same questions as described below. These questions are yes/no, a yes to no scale of four steps or answers can be selected in a list. Often followed by an open question to explain the answers in more detail. All questions aim to gather information for one of the dependent variables, get to know the participant or gather expert opinions and feedback on the designed protocol. Using symbols behind each question shows what this aim is.

♣ Performance	♠ Situation awareness
♢ Trust	♡ Satisfaction
⊙ Get to know the participant	⋆ Protocol

Explanation and basic information (1, 2)

The participant is not explicitly informed about the exact purpose of the research. The participant will, however, get a short introduction on how to use the simulation environment. This introduction includes the commands it can give as an officer of watch. The environment is easy to use, as action are similar to the ones currently happening at the bridge. Thereby is some information about the participant acquired, followed by some general questions relevant to the protocol, such as there knowledge of Standard Maritime Communication Phrases (SMCP).

- Which certificates do you have? ⊙
- What is your experience as captain and mate? ⊙
- Which type of vessels did you operate? ⊙
- What do you expect from the developments towards autonomous shipping? ⊙⋆
- What do you see as the biggest challenge for introducing autonomous and unmanned vessels? ⊙⋆
- Do you expect to trust autonomous ships more? ⊙♢

- Do you want to know if a ship is unmanned? ♢
- Are ship's horns still used to communicate intended manoeuvers? ⋆
- What are the most important forms of communication? ⋆
- Do you still use SMCP consciously? ⊙
- How does a standard conversation look like, according to the SMCP? ⊙⋆
- Is the protocol around SMCP easy to use? ⋆

Situations and scenarios (3, 4, 5, 6)

The next steps are repeated several times for the different situations. The communication in that situations should be about the intentions of the other vessel. When using the designed protocol, the conversation should be similar to the examples given in section 11.1.2. It depends on the experiment order, in which situation it is possible to communicate:

1. Crossing situation at North-Sea
2. Leaving port of Rotterdam via Maasgeul

During these simulation are all actions logged together with the CPA. Thereby are notes made on the way decision are made, these are verified with the participants when filling in the questionnaire in the next steps. This includes feedback on the risk taken.

Relevant questions for situation (7, 8, 9)

Different questions will be asked, to gain insight into the quality of the experiment and the effectiveness of the protocol. Thereby is a link made to the decision process as discussed in section 2:

- What type of situation is this? ♠
- Which criteria are relevant? ♣
- Which strategy did you choose? ♣
- Which actions were taken? ♣
- What was the speed of different vessels? ♠
- How often did the ships change their course? ♠
- Which colour did every ship have? ♠
- Did other ships behave as expected? ♠♣
- Were you in control over the situation? ♣♢♡
- Did you miss any information to come up with the right strategy? ⋆♡

- Was it necessary to communicate? ♠◊
- If there was communication, was this as you expected? ♡
- Would you act differently, if you knew there was a human officer of watch? ◊♡

General questions on protocol (10)

After running the different situations, an interview is held. This part of the experiment is intended to answer the following questions from the participant perspective and explain the purpose of this research:

- Is the protocol around SMCP easy to learn? ⋆
- Is the protocol around SMCP a complete protocol? ⋆
- Do you have any other comments on SMCP? ⋆

11.2 Evaluation results

The evaluation results describe the outcomes of the test. Because of the iterative and rapid research cycles, the evaluation does not necessarily include all requirements/claims/use cases available in the system specification. Often the evaluation investigates a subset of the system specification. Therefore, it is often useful to also specify what claims were tested, with the use of what evaluation method, and what artefact was used during the evaluation (i.e. which requirements, technology, and interaction design patterns were included in the artefact).

The results of the experiment are evaluated in a systematic way for the 16 respondents. First, a summary is made of the reactions to the Google form. This summary is made using both a statistical analysis on the ordinal level and classification of answers on the nominal level. Dependent variables discussed are discussed, using this summary. Six of the participants were only able to communicate in the crossing situation at the North Sea. Ten participants had the possibility of communication only in the second situation when leaving the port of Rotterdam via the Maasgeul.

11.2.1 Observations during simulations

Situational awareness is tested by questioning participants about the relative start speed, colour and course changes. The participants focused on possible future risks. Speed estimations did they often base on normal behaviour for different ship types, instead of taking the speed vector into account.

The first case is the crossing situation at the North-Sea. The situation was implemented in such a way, that the container ship was the slowest, and tanker the fastest. As can be seen in figure 11.6 did the participants think the opposite, which can be explained by the usual speed of different vessels. A large container ship (Emma Maersk) usually goes the fastest, the smaller general cargo vessel a bit slower (Astrorunner), and the tanker (Gulf Valour) is often the slowest. Thus opposite to how the scenario was set-up. Between the two test-group, protocol and no-communication was there no difference.

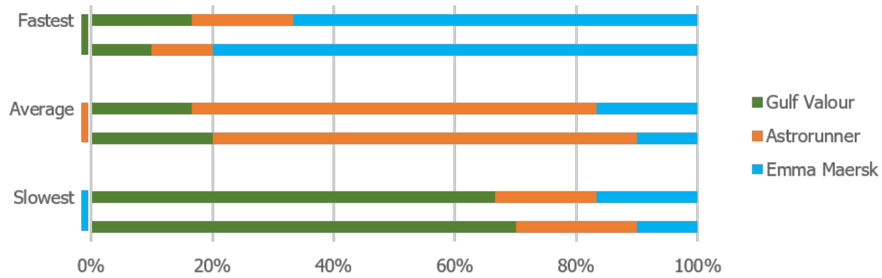


Figure 11.6: Estimation of relative start speed at North Sea (protocol vs no-communication)

Wrong estimations could also be the result of the simulation environment. But the estimates of the participants are opposite to the vertical-horizontal illusion when considering the speed vectors as measurement [Prinzmetal and Gettleman (1993)]. Thereby did participants estimate the speed correctly, when they were able to see the speed vectors and were questioned about the speed during the experiment, instead of making an estimation based on memory.

The relative start speed in the second situation when leaving the port of Rotterdam via the Maasgeul, had results as could be expected. Figure 11.7 shows this, where most participants did estimate the relative speed correctly. Also, for this case, was there no difference if a participant was able to use the protocol.

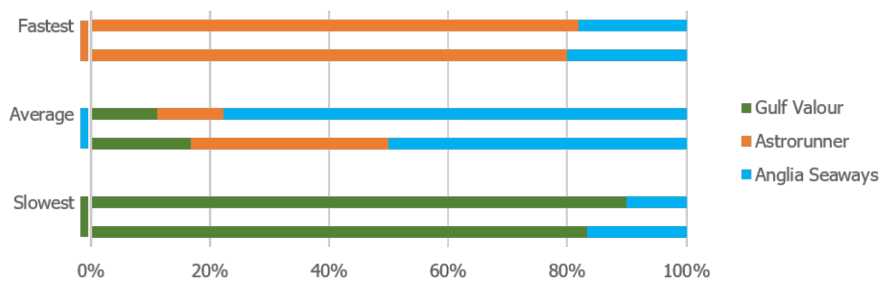


Figure 11.7: Estimation of relative start speed at Maasgeul (protocol vs no-communication)

The estimation for the number of course changes went a lot better for the first situation. The

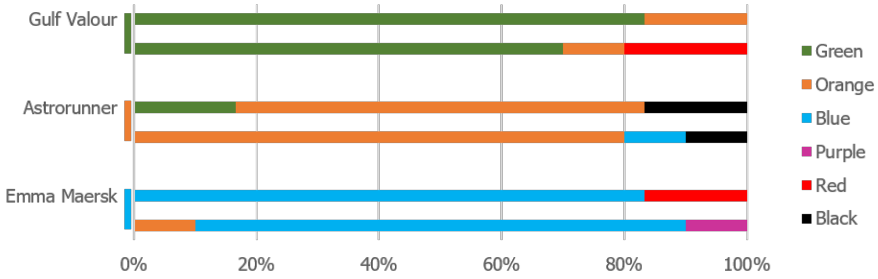
Emma Maersk did not alter course, which participants correctly observed, just as the single course change for the Astrorunner to ensure a perpendicular passing of the traffic separation scheme. The course changes of the Gulf Valour depended on the situation. If the participant controlled the Gulf Valour, they said that the Gulf Valour made two course-changes: First alter course to starboard, and second return to original course. While in the case where the participant operated from the Astrorunner, did they often not count the last alteration. Thus did the participants count only one course-change for the Gulf Valour. The attention of the participants can explain this observation. When the Gulf Valour alters course for the second time, the Astrorunner is close to the Emma Maersk, and the participant pays therefor attention to the Emma Maersk and not the Gulf Valour.

In the second situation are the course changes much more clear, as two vessels go straight (Astrorunner and Anglia Seaways), while the Gulf Valour make one clear turn. It depends on the exact actions the participant took, how many course changes the vessel made. But there were no wrong answers given. Being able to communicate, did affect the strategy and thereby actions, but not the ability to register the right amount of course changes.

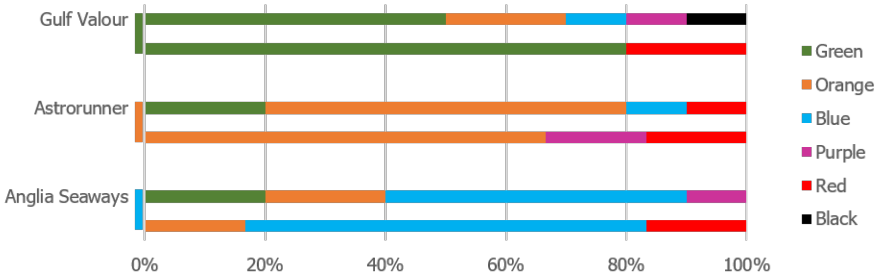
The colour is something which does not have any effect on the decision. It is, however, information which shows if the participants remembered irrelevant details about the simulation. These irrelevant details are a measure for the free cognitive capacity, as participants remember more of these when the situation is less complex, also known as inattentional blindness [Most et al. (2000)]. This difference can be seen between the situation when crossing the North Sea, and entering the Maasgeul. Figure 11.8a and figure 11.8b show that more participants have the colour correct in the crossing situation at the North Sea, which is also the less complex situation when it comes to making the right decision.

After each simulation did the participants also answer questions on actions of other vessels, communication and missing information. The results of this are shown in figure 11.9 and figure 11.10. It should be noted that the experiment was not fully counter-balanced and there was a limited number of participants, which means there is a high margin of error. Therefore are conclusions drawn only when there is a significant difference between the results for protocol vs no-communication.

Interesting is that 40% of the participants would have acted differently when no communication was possible. Their explanation for this is that they could not anticipate to other vessels, thus slowed down, hoping that they could deduce from the actions of the other ships what their plans were. If the protocol works well, this is not needed. The other participants said they followed the COLREGs, and expect unmanned vessels to do the same. This reasoning means that the participants would act the same when the other vessel was manned.



(a) Situation when crossing at the North Sea



(b) Situation when entering Maasgeul

Figure 11.8: Response to color of vessels for protocol vs no-communication

A conclusion which was expected is that the participants would miss information when they were not able to communicate using the protocol. In the second situation when entering the Maasgeul, this is the case. However, most participants did indicate they were also missing information on the CPA during the experiment as is usually shown on the ARPA. With the first situation at the North Sea did more than half of the participants think it was not necessary to communicate. Therefore did a similar amount of participants miss the same information in both cases (protocol vs no-communication).

Also did the type of ships participants have operated, influence their behaviour, as operators who used to sail on small vessels did often take more risk and expected higher accelerations and manoeuvrability. While the operators who used to operate large vessels, such as tankers and cruise ships, prefered to wait for other ships to act first if they were missing information.

In case the participants used the protocol, the communication was in both situations as expected for most of them. This shows that seafarers expect to be able to use natural language variation on SMCP.

The participants are questioned about their trust in autonomous systems, in all cases, they want to know if a ship is unmanned, so that they can anticipate to it. They gave as an example that it is likely that an autonomous vessel will follow COLREGs more strict, than ships operated by a Filipino crew. But in emergencies, such as failure in the engine room, do the participants trust autonomous ships less. Therefore there was not a definitive answer if autonomous vessels will be trusted more or less, compared to manned vessels.

11.2.2 Evaluation of the protocol

The questionnaire tries to validate design choices for the protocol. A design choice which is based on COLREGs, but unknown if it works in practice is the usage of the ship's horn to communicate the intended manoeuvrer. All participants said that ships only use the ship's horn in two situations. When a ship wants to get attention from another ship, in case other means of communication do not work. Or in case of fog, then is the fog horn sounded every few minutes to make other ships alert of presences [IMO (1972)].

The most important form of communication is the VHF radio, although new systems are used more often, such as AIS text messaging and INMARSAT-C. INMARSAT C provides two-way data and messaging communication services to and from virtually anywhere in the world. Which improves the ship-to-ship communication, the disadvantage is that surrounding vessels are not able to listen to the conversation and thus will not receive the shared information.

According to the participants are conversations via VHF based on SMCP, but do not follow the protocol strictly. The participants describe a few key characteristics:

- Start of every message has the purpose to get attention, by calling specific vessel: "($< name\ other\ vessel >$, $< name\ other\ vessel >$. This is $< name\ own\ vessel >$)"
- Common practice is to use the words as defined by SMCP (alter course, instead of change course). The sentences are often not used as described by SMCP.
- Marker words are not used.
- When responding, the previous message is repeated.
- Message is ended with over, conversation with over and out.

In general, they do state that the protocol is easy to use, easy to learn and complete. as figure 11.11 shows. The drawback is that the protocol covers too much. Therefore people can't always remember how they should follow the protocol and make variations to it.

Less relevant for SMCP, but more for the final design of the protocol is the knowledge of English. As the level of and pronunciation of English is not at the right level for all officers, certainly for example with Filipino, Indian or Pakistani crew. This problem means that voice recognition should be tested for a variety of accents. It is also essential to know the names of different systems (e.g. LORAN, DSC, etc.) and places (e.g. camping, seal beach, etc.), which vary per ship and area and cannot always be found on maps.

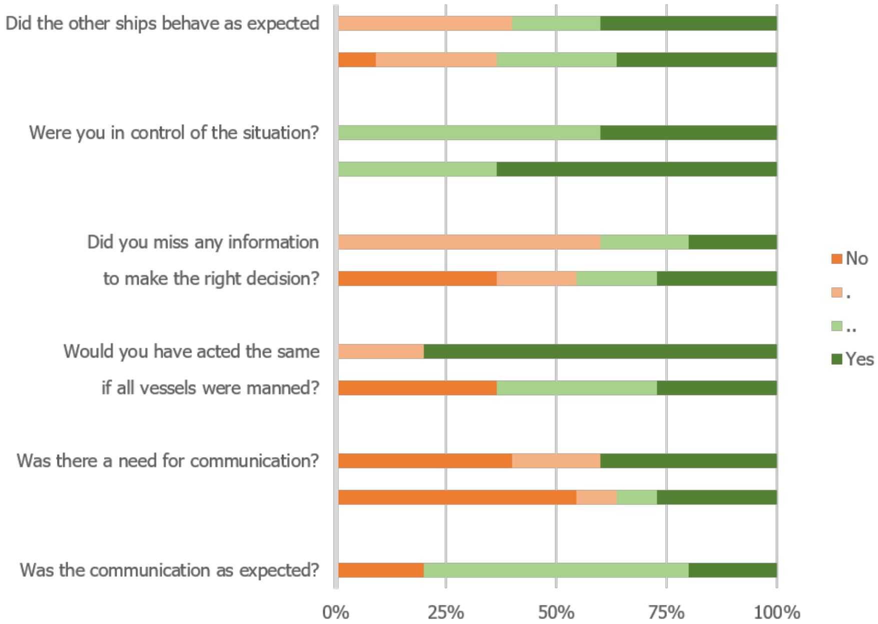


Figure 11.9: Protocol vs no-communication when crossing at the North Sea

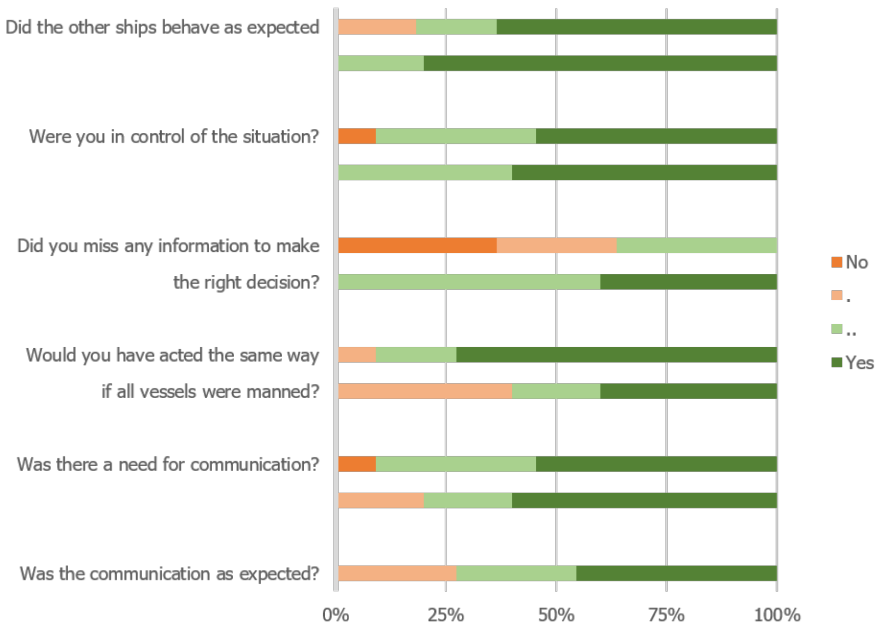


Figure 11.10: Protocol vs no-communication when entering Maasgeul

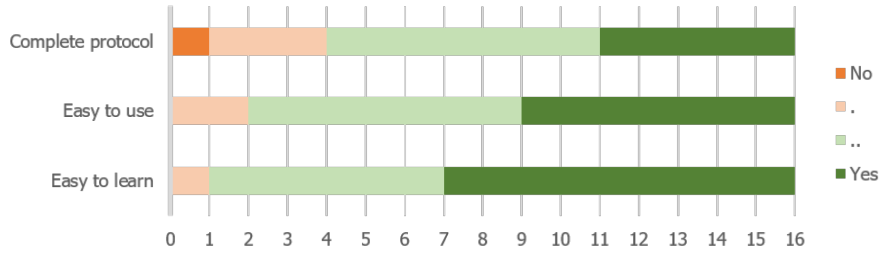


Figure 11.11: Opinion of participants on SMCP

Thereby was the identification of vessels hard before the introduction of AIS, which should be considered during the development of such a protocol. As it should also be possible to identify ships when the AIS fails, this means it should be possible to identify vessels without knowing the call-signs and name via AIS.

Due to the speed and starting point of the simulation, the operators did have to decide if they wanted to communicate quickly. Therefore was commented by some of them, that they would have liked to communicate earlier, than the moment the simulation started. So they would have had enough time in case the other vessel did not respond immediately, although the designed protocol and the current usage of VHF differ in the way how operators should use SMCP strictly. Does it take the same amount of steps, and a similar amount of words. Therefore it is not expected that the conversations will take longer if the voice recognition works correctly.

11.3 Lessons learned

The experiment aims to determine if the designed protocol will ensure safe navigation and more situational awareness when manned and unmanned vessels encounter each other. This can be answered by looking into different measures on how the participants and protocol performed during the experiment, which shows what should be taken into account in the next iterative steps while developing the protocol.

The *performance* is an evaluation if the participant followed regulations and made the right decisions. None of the cases did result in an accident, even though there were some close encounters. There was no clear correlation between the ability to communicate and the closest point of approach (CPA), as some participants took more risk when the intention of another vessel was known. The protocol did help to acquire the right information to chose a strategy earlier in the process. This time gave participants more control over the situation and ensured them that the other vessels would also follow COLREGs. Thereby was the communication as expected when using the protocol, after reminding participants how the communication should be according to SMCP. This corresponds with the opinion of participants that it is a complete protocol, which is easy to use and learn.

Trust is the second measure which is evaluated. This means that the participants are confident that the system works as expected. The participants were most worried about voice recognition, definitely for Filipino, Indian or Pakistani crew, as these have an accent which is already hard to understand for humans. The words and sentences within the protocol itself shouldn't be the problem, as it is optimised to be well understandable via radio. The recognition of speech acts could be difficult, but it has the advantage that the protocol uses explicit keywords. Thereby is every response started with a confirmation of the previous message, which helps to increase trust in the protocol and unmanned vessels in general.

The third measure is *situation awareness*. This is measured by checking if the participants were aware of relevant and irrelevant details such as estimating relative speed, number of course changes and the colour of a vessel. The relevant details went better in the complex situation which demanded high attention. Irrelevant details such as course changes of vessels which already passed, or the colour of a vessel were remembered less in these situations. These results were however not influenced by the protocol, which means that the protocol has a small effect on the free cognitive capacity.

The last measure is *satisfaction*. The participants like to use the protocol. Here is the voice recognition most important, if this will understand natural language variations to SMCP, are most participants very positive. As SMCP itself is an 'idiot proof' system. If it is possible to communicate this way with vessels, will strategies not differ from interactions with manned vessels.

Critically looking at the way the results are acquired shows the relevance of these results. The main strength of the experiment is that a participants are used with different backgrounds, as the participants are both experienced and inexperienced on small and large vessels. The group of 16 participants is sufficient to draw conclusions for this first iteration and shows that it is feasible to develop a protocol using SMCP and a conversational agent. When implementing the full protocol more evaluations are necessary. These evaluations should consider more different situations. These will not only say if there is an impact on the performance, trust, situation awareness and satisfaction of seafarers. But the results of these evaluations should also what this impact is. This means it is possible to mitigate a negative impact, or exploit the advantages.

12 | Conclusion

This part has shown that it is possible to solve the challenge of communication and ensure safety by enabling unmanned vessels to communicate with manned vessels. We solved this challenge by designing a protocol for communication, based on existing maritime systems and protocols, using the situated Cognitive Engineering (sCE) method. Where the first phase includes a description of operational demands, relevant human factors and envisioned technologies. The second phase consists of design-scenarios and use-cases, which result in a list of requirements. These requirements are used to define an ontology for the system and a description of the capabilities of the system. With experts is evaluated if this is indeed enough to ensure that unmanned vessels can operate safely between manned vessel. All this combined will answer the question:

> *Will a protocol based on existing maritime systems and communication protocols be sufficient to ensure safe navigation, while manned and unmanned vessels encounter each other?*

Using Standard Maritime Communication Phrases (SMCP) as the core for a new protocol, will ensure safe navigation. As the advantage of SMCP is that it is already optimized for radio communication. The clear speech acts will make it more easy to automate the communication. Which means than unmanned vessels are able to use the same protocol as is currently used by manned vessels. Thereby did experts confirm that it is an idiot proof system, which is easy to use and learn, certainly when the speech recognition tool will be able to handle natural language variations on SMCP. This will be more easy when officers use the protocol as it is designed. After recent accidents between manned vessels was already decided that pilots and masters should use English speech and SMCP in all cases, to avoid misunderstanding. Avoiding misunderstanding will result in better situational awareness, which results in safer navigation. Thus it is possible to ensure safe navigation when manned and unmanned vessels meet, by developing a protocol. The designed protocol uses existing maritime systems and communication protocols. For this design is the situated Cognitive Engineering (sCE) method used.

This research was a first iteration for developing a functional protocol for communication between manned and unmanned ships. It has showed that it is possible and what should be taken into account. The next step is to map different speech acts, to specific situations. One of the key challenges is thereby the recognition of situations. The mapping can eventually be used to define the full ontology which should be implemented in the conversational agent.