The effect of display type on the task of marine collision avoidance

Alexandros William Eftychiou

Project report submitted in part fulfilment of the requirements for the degree of Master of Science (Human-Computer Interaction with Ergonomics) in the Faculty of Life Sciences, University College London, 2012.

ACKNOWLEDGMENTS

I would like to express my sincere gratitude and thanks to Dr John Dowell who provided invaluable guidance, support throughout the project. Special thanks to Jenny Dragatsi who was always there with encouraging words and for her boundless patience. Without fellow members of UCLIC 2012 who were all great at providing a safety net of good advice and reassurance I feel the project would have been far more daunting. Finally a huge thanks to mum and dad for reasons that are too numerous to even begin listing.

ABSTRACT

This investigation is the first evaluation of the CDS direct perception display (CDS) designed by Liu and Pederesen (2004) to aid the navigators' of large vessels in the task of collision avoidance. A failing of this display is that it does not also represent the regulations (COLREGS) that govern ships at sea. Lee and Sanquist (2001) using evidence from the process control domain as outlined by Vincente and Rasmussen (1992) argue that the perceptual salience of the CDS makes it less likely that the information not displayed will be taken into consideration. Radar was used as a control as there is no evidence that radar produces this type of behavior. No significant difference in error rate whereby participants broke the COLREGS were found refuting Lee and Sanquist (2001). Participants' accuracy was recorded by inspecting trials where they abided by the COLREGS but chose either too soft or too hard a turn to avoid the collision. CDS was significantly more accurate though participants using the CDS and were more likely to consistently over-rotate the own ship icon to avoid a collision than when using the basic radar display. No significant effect of display type on task load was found though comparatively the CDS did significantly increase the task completion time.

CONTENTS

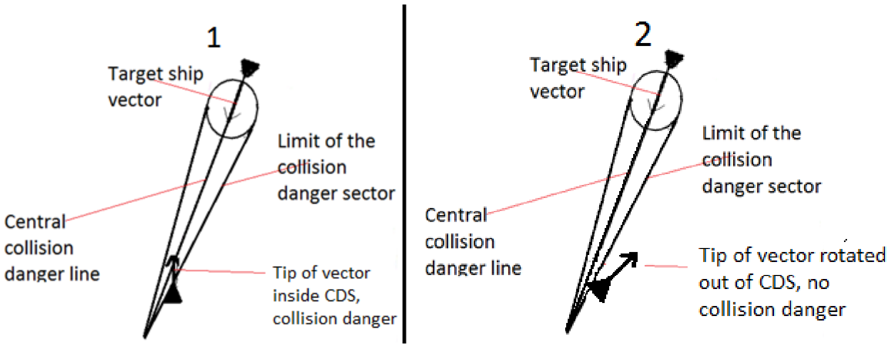
1 INTRODUCTION AND LITERATURE REVIEW

Collision avoidance is a central issue to maritime safety, collisions can result in the loss of human life and have grave economic and environmental consequences. As identified by Petersen and Nielsen (2001) the means available for maintaining the required degree of safety is limited which makes the issue of supporting navigators in avoiding collisions central to maritime safety. What makes this even more pressing is the fact that the seas are becoming ever more densely populated with giant vessels.

The tool known as the Advanced Radar Plotting Aid (ARPA) is designed to aid the navigator in his decisions and reduce the danger of collision. It displays the surrounding marine traffic flow within a specified range, augmenting the icons corresponding to other ships otherwise known as target ships with alphanumeric data when those targets are selected by the operator. The data gives the operator an understanding of the target ships position, bearing, speed, Time to Closest Point of Approach (TCPA) with own ship and Distance at Closest point of Approach (DCPA) with own ship. Target ships are also presented with vectors which are arrows protruding from the bow of the target ship and own ship. The vectors are a graphical representation of the heading of the target ship while the tip of the vector communicates where the target ship or own ship will be in a pre-set period of time. A safety envelope can also be determined by setting a minimum safety DCPA which if violated sounds an aural alarm to alert the navigator. The graphical ARPA display is therefore able to augment the human perceptual system making it very useful in understanding the surrounding marine traffic.

There are aspects of the system identified and outlined by Liu and Pedersen (2004) that could be improved. According to the authors the numerical display of anti-collision information is not very compatible with the capabilities of the human perceptual system, thereby hindering the process of information extraction. Secondly, the navigator has to judge the potential collision dangers from the analysis of large amount of numerical information of target ships, and then make decisions on how to act and take necessary evasive actions. The processing of large amount of data increases their mental workload and is time-consuming when the navigator is confronted with many targets simultaneously. The authors reference the findings of Cockroft (1984) that when the workload is high and beyond a navigator's processing abilities, it will result in poor decision-making and errors. Thirdly, current ARPA display does not provide sufficient information in the interface to support navigators' problem-solving, or say, knowledge-based behavior (Rasmussen 1986). A standard procedure for collision risk assessment with utilization of ARPA is firstly to observe true vector display to get an idea of the true traffic flow, then switch display to relative vector where the vector assumes that the target ship will hold true to the last data collected about it and finally turn on Trial manoeuvre mode and trial out the effect of a proposed course and/or speed change. The navigator then decides an evasive manoeuvre, switches off Trial Manoeuvre mode and once again observes the true vectors of targets before finally taking a manoeuvring action. Having to switch between the true and trial displays is not optimal but rather disruptive, wasting precious time and separation between vessels as the navigator decides the best course of action. Further it indicates the limitations of the representation provided as the navigator lacks enough information to build an updated mental model that could be used to reason about the situation and possible solutions.

Liu and Pedersen (2004) propose a direct perception ARPA display that would address the aforementioned issues. The authors argue that to improve navigators' collision avoidance assessment activities, the collision danger areas should be visualized in the display. This visualisation is based on the proposal made by Pedersen et al (2003,2004) in which collision danger areas are displayed as cone-shaped regions (Collision Danger Sector, hereinafter CDS) to acquired targets. With this type of display the risk of a collision can simply be assessed by viewing the tip of the own ships velocity vector in relation to the CDS. If the tip is within the CDS then there is a risk of a collision and any manoeuvre that takes the own ships velocity vector out of the CDS can be described as a collision avoidance manoeuvre. The is graphically demonstrated below.



Unlike the ARPA where navigators have to perform effortful and error prone analysis through the trial and error manoeuvres the CDS display is able to aid navigators in having a good projection of the future traffic states by allowing them to view the CDS and where their vessels vector is in relation to it. To date no controlled studies have been conducted to verify the claims made by the above authors, this investigation is therefore an important first in the evaluation of the CDS.

The importance of certain factors such as rules of the road in a maritime environment is not to be neglected, on many occasions ships will not be able to communicate with each other for various reasons and thus will not be able to know what the intentions of the other vessel are. The only way of being able to infer the actions of another vessel is by assuming that it will adhere to the maritime rules of the road (COLREGS) as set out by the International Maritime Association (IMO). If a navigator acts in such a way expecting that the target ship will also conform to the regulations and the target ship does not by breaking the rules, precious time and separation between the two vessels can be lost which consequently increases the chances of a collision. It is therefore crucial that systems such as the ones above do not undermine the likelihood of navigators taking COLREGS into consideration as they constitute primary factors in the navigators' task.

This does happen to be one of the main criticisms levelled against the direct perception CDS display by Liu and Petersen (2004). Tam, Bucknell and Greig (2009) stress that the above interface makes no attempt to incorporate the rules of the road (COLREGS) as specified by the IMO and might encourage the navigator to make incorrect decisions. Lee and Sanquist (2000) also level this criticism but against a different type of interface that allows direct perception of a collision situation. This interface shows safety zones of both the own vessel and the target vessel, the navigator then has to ensure that these safety zones do not overlap. The authors break down the task of considering a course change due to a collision danger being identified into its constituent cognitive components. These components are input select which is based on selective attention that can be affected by the salience of the information source. Inexperienced mariners can be drawn to the salience and seemingly reliable information provided by a display. While this is highly possible the literature referenced below by Vincente and Rasmussen (1992) contradicts this, they state that an increase in experience makes operators more able to complete the task based perceptual processing thereby making them more likely to be drawn to the perceptual salience of the display. Inexperienced operators would not be able to carry out the task through perceptual processing alone and would have to also rely on analytical problem solving where conceptual information not displayed would have to be incorporated to complete the task. The next component addressed by Lee and Sanquist (2000) is compute which is the process of calculating alternate courses, modern ARPA's help this process by allowing trials of possible manoeuvres on the system before the actual manoeuvres are carried out. Interpret is an important aspect as it involves the consideration of potential course changes in the context of the rules of the road which will constrain course selection by limiting the viable options and estimating likely behaviour of the target vessel. Some ARPA designs may undermine this process by encouraging overly simplistic heuristics, such as "manoeuvre to keep safety zones separated." or in the case of the simplified heuristic proposed by Liu and Pedersen (2004) place the vector outside the collision danger zone. Both these heuristics do not consider the rules of the road in interpreting the merit of a potential course. The final cognitive task "Decide/Select" identifies the most promising course change. Multiple sources therefore need to be considered in order to carry out the decision of consider course change many of which are missing in the ARPA display. Therefore any decision made purely on the information presented by the ARPA is not taking into consideration other important factors.

Lee and Sanquist (2000) make comparisons with process control domains as the task of ship navigation can be described as a process. Further navigation similarly to process control enables automation which allows a small number operating personnel to operate the process from a central control room. When making their comparisons to the process control domain Lee and Sanquist (2000) refer to Vincente and Rasmussen (1992) who discuss with reference to the findings of Klein(1989) and also of Kirlik (1989) that people naturally adopt strategies that allow them to reason about the state of a process using perceptual processing as opposed to analytical reasoning. They discuss three levels of cognitive control, Skill based behaviour (SBB), Rule based behaviour (RBB) and Knowledge based behaviour (KBB) and the way in which information is interpreted defines which level of cognitive control will be activated. KBB involves analytical problem solving while RBB and SBB are concerned with perception and action. In order for the SBB and RBB levels of cognitive control to be activated the operator must be attuned to the perceptual properties of the environment. Without a particular level of experience the operator would not be familiar enough with the perceptual aspects of the process in order to be able to complete the task at the SBB and RBB level. In novel situations the operator would rely on analytical reasoning or KBB in order attend to the task. Vincente and Rasmussen (1992) discuss that in fact perceptual processing is able to out-perform analytical reasoning in terms of both speed and decrease in errors while being the natural level of choice for experienced operators. The issue with this level of cognitive control similarly to controlling the process of ship navigation is that in some instances in process control where displays failed to describe the system completely the operators act as if the process was physically structured as it was shown on the display. These instances were found to induce poor decision making because the operators failed to consider the properties of the process not presented in the display. This poor decision making results in under-specification and errors (Vincente and Rasmussen 1992). Lee and Sanquist (2000) note that this also seems to occur with sophisticated ARPA's where perceptually salient features of the display capture the mariners attention while their conceptual knowledge of rules of the road goes unused because the display does not contain this information. They also state that based on interviews with mariners, using a standard radar is more likely to ensure that navigators follow the rules of the road as the perceptually salient aids provided by advanced ARPA's are not present and thus do not encourage rule breaking behaviour.

A typical example of this pattern of behaviour is reported by Vicente and Rasmussen with reference to a study by Hollnagel (1981). Hollagnel (1981) found evidence that participants relied on the concrete perceptual characteristics of the process of the display while demonstrating a tendency to disregard the abstract functional properties of the process being controlled. Effectively participants were treating the process as if the display was a complete veridical representation of the process, the issue is that this is not the case. There was no direct consistent relationship between the perceptual characteristics of the display and the constraints that described the process behaviour. Therefore it is difficult for the operator to consistently be able to control the system by only considering the surface features presented in the display.

An interesting study referenced by Vincente and Rasmussen (1992) related to this phenomenon was conducted by Smith (1989). His study compared the performance of two groups of subjects, one that was given a deliberately incomplete problem representation in the form of a decision tree diagram, and another that was not given any representation aid. The findings revealed that the incomplete representation actually impaired performance because subjects tended to rely on it as a comprehensive and veridical representation of the problem, thereby participants failed to take into account important factors that had been purposefully left out of the representation. Thus, the author concluded that being provided with an incomplete problem representation can actually lead to worse performance than having no representation at all. Vincente and Rasmussen propose that this can be attributed to the "out of sight, out of mind" phenomenon identified by Fischoff et al (1978). An interesting parallel could be drawn with regards to the use of radar and CDS- ARPA which provides a representation aid in the form of CDS and vectors extending out of the target and own ship, but these aids are only an incomplete representation of the situation by not including the

COLREGS. This could cause preferential attendance to the salient aid on the display at the expense of considering the COLREGS. This would be similar to the out of sight out of mind phenomenon identified by Fischoff et al. The Radar does not provide any representation aids such as a CDS or vectors, instead the signal for a collision danger with a target vessel will be that the target vessel is on a constant bearing with own vessel. To ascertain whether own ship and target ship are on a constant bearing forces the user to continuously track the progress of the other vessel so as to build a mental representation of the situation. The investigation will therefore use radar as a comparison to test the hypothesis that the users will be drawn to the perceptually salient depiction of the CDS display. If radar causes less rule breaking behaviour this could provide evidence for the importance of making certain that a perceptually salient display such as the CDS takes into consideration all the factors including the COLREGS. The findings could support the argument that the benefits of playing on the human operators strength of perceptual processing by presenting the information in perceptually salient manner might be negated by its increase in adverse effects on users performance through a lack of ability to take important information that is not presented on the display into consideration.

With Vincente and Rasmussen (1992) reference to SBB, RBB and KBB it was deemed important that since the current experiment was not able to recruit specialist navigators as participants but rather laymen’s from the psychology subject pool. That the paradigm be simple enough in order to help increase the likelihood that participants would become skilled and experienced enough to rely on lower levels of cognitive control and thus be more reliant on the perceptual saliency of the displays as opposed to having to fall back on their knowledge base in order to solve the task. Ease of task was also important so that participants’ lack of experience and skill did not have adverse effects in terms of translating into errors due to lack of understanding rather than being drawn to the perceptual saliency of the display.

A delay system will also be introduced with the objective of simulating conditions that could affect the navigator's decision making process. Lee and Sanquist (1996) had previously mentioned that certain situations increased the likelihood of rule breaking behaviour. The particular situation given by Lee and Sanquist (2000) was that the CDS system indicated that a left turn would increase the distance of own ship from the land as well as keeping own ship's CDS separated from the target ship's CDS. The problem was that a left turn would break the "rules of the road" which has been found to increase the likelihood of a collision as the target ship is not anticipating such a turn. The result of this could be that the target ship also turns left to continue on its trajectory of passing in front of own ship as the target ship has the right of way and should pass in front. This results in a near miss situation as target ship has now turned onto a collision path with own ship expecting own ship to correct in order to stay in lines with the rules of the road, which the navigator of own ship has no intention of doing. This situation results in the distance between the two vessels closing until one vessel decides to make a correction resulting in a near miss. In the case of this investigation the delay system was designed to create similar scenarios that could increase the likelihood of exhibiting rule breaking behaviour. Encouraging distance from land was replaced in this investigation with encouraging an on time arrival at the ship's destination. Participants were instructed that a Soft turn response would incur a delay of an hour while hard turn incurred a delay of two hours. It was of interest to also investigate whether participants would be more likely to treat the display as an absolute when there was an increased delay attached to their response particularly in the CDS-X factor. The CDS-X where the tip of own ships vector is protruding out of the target ships CDS simulated the aforementioned CDS display by not taking the COLREGS into consideration thus informing the navigator that a violation of the rules of the road would be acceptable in order to keep the target and own ship separated while keeping delay to a minimum. The difference with the CDS displays in this experiment is that they use the simplified heuristic of keep tip of vector out of CDS to avoid collision. In the aforementioned CDS display both target and own ship CDS are shown in order to allow the user to track and keep them separate. In the case of the CDS-X display in the current experiment the vector crosses through and is protruding out of the target ship's CDS, but the same level of simplified heuristic is being communicated. The CDS-X display is communicating to the navigator that the tip of the vector is out of the CDS and so it is safe to increase speed and cross targets path (Against COLREGS). While the Lee and Sanquist (2000) CDS display communicates that a left turn (Against COLREGS) would keep the CDS of both ships separate while increasing distance from land.

This investigation will attempt to gather empirical evidence in the form of error rates to test the claims made by Vincente and Rasmussen (1992), Lee and Sanquist (2000) that displays that do fail to describe a system completely induce poor decision making expressed in this case as an increase in error. Error in this investigation will be discussed in the context of crossing errors. Crossing errors are described as when participants choose a manoeuvre that breaks the COLREGS. The COLREGS in the current experiment state that you are not allowed to cross targets path. Therefore when participants choose a manoeuvre that causes own ship to cross targets path this will be logged as a crossing error. The investigation will look particularly at the direct perception interface (Liu and Pedersen 2004) as it falls into the description given by Lee and Sanquist (2000) of an interface that does not present all the properties of the process. This is true of the direct perception interface which is incomplete as it does not represent traffic protocols alongside the CDS as identified by Tam, Bucknell and

Greig (2009)The direct perception interface will be compared with a standard radar interface in various collision avoidance scenarios. The target ship will approach at own ship at trajectories of 10, 20, 30, 40, 50, 60, 70 and finally 80 degrees from the vertical from both the left and right hand side. The first four angles fall into the response category of soft turn which cuts off at forty five degrees. The second set of angles fall into the response category hard turn.

The direct perception interface will be broken into a further two factors where tip of own ships vector is inside the CDS and where tip of vector is crossing through and protruding out from the top of the target vessels CDS, this factor will be labelled CDS-X. As previously mentioned with the CDS display, if the tip of the own ship vector is outside the CDS then there is enough separation for the two vessels to safely pass each other. In the CDS-X factor the tip of the vector has crossed completely through the target vessels CDS and is protruding out of the other side, the system is communicating that at this point in time staying on the current course and crossing targets path is a safe option. The issue is that this goes against the COLREGS created for this particular experiment. Therefore the trials in the CDS-X factor are not taking the rules of the road into consideration and instead are communicating that crossing targets path is an appropriate option. This is therefore a simulation of a display that falls into Lee and Sanquist (2000) description of a display that does not present all the properties of the process. The CDS factor where the tip of the vector is inside the CDS is not communicating that crossing targets path is the appropriate response but the simple heuristic of rotate to keep vector out of CDS is still communicated. This heuristic does not take the COLREGS into consideration therefore CDS also falls into the description of a display that does not present all the properties of the process. The difference between CDS and CDS-X factor is that CDS-X is saliently communicating a rule breaking behaviour is appropriate while the CDS communicates that the user has to rotate to get the tip out of the CDS, but the choice of what degree of rotation and in which direction is left up to the user. The importance of providing empirical evidence for this is that this investigation can support the argument for the development of CDS displays that represent context in the form of the rules of the road into consideration and thus increase navigators decision making performance.

The Liu and Pedersen (2004) also argue that the appropriateness of a manoeuvre can be analysed based on whether it takes the tip of the vector outside the CDS. They state that this makes the system transparent and decreases workload compared to more basic displays that do not produce CDS. In order to assess these claims participants will be requested to complete a NASA-TLX in order to measure the load of carrying out the task on a radar and CDS display. Reaction times would also be noted for each trial.

The reaction time data will also be used to investigate whether there is an effect of degree of mental rotation of the own ship icon on the task completion time. Shepard and Metzler (1971) found a linear increase in participants' reaction times that was a function of the degree of mental rotation. Of particular interest is to examine whether the trials where a rotation of 30-60 degrees which is nearer the discrimination threshold of 45 degrees will take significantly longer than trials at the extremes. Trials at the extremes are described as trials where either a 10-20 or 70-80 degree turn is appropriate. These trials are further away from the discrimination threshold of 45 degrees which is the cut-off point between selecting a hard or a soft turn. An interesting finding would be if rather than participants exhibiting the linear increase as a function of their mental rotation that instead they demonstrate an increase in reaction time closer to the cut-off point. This could intuitively be possible as closer to the cut-off point the harder it is to recognise whether the appropriate turn falls into the hard or the soft turn category. Further away from the cut-off point it may become easier for participants to recognise that the appropriate response is a hard or a soft turn. Such a finding would demonstrate that recognition overcomes the need for participants to mentally rotate the image in order to gauge whether their response falls into the soft or hard turn category.

A related area that will also be investigated is whether the number of errors increase depending on proximity to the cut-off point. This is again intuitively possible as the participants' ability to distinguish between a hard or soft turn response should decrease as proximity to the cut-off point increases. Finally the number of total errors will be compared between the two displays to investigate any differences between participants ability to make the appropriate response. Specifically the error rates for all angles of approach in each display will be compared to attempt to gain a higher resolution understanding of how each display at each angle of approach affects participants' responses.

Hypotheses:

1. Trials in the perceptually salient CDS-X factor where the display actually communicating that no response is necessary will produce more crossing errors than both the CDS and Radar factors. An interaction with degree of turn is also expected where harder turns in the CDS-X are expected to produce more crossing errors than normal turns.
2. In line with the Liu and Pedersen (2006) findings the CDS and CDS-X displays will score lower on the NASA-TLX task load index than the radar display. The reasons for this being the display of both vector and CDS which the authors propose increase transparency and decrease work load

3. Reaction times are hypothesised to be lower for CDS displays as they play on the strengths of rapid perceptual processing as described by Vincente and Rasmussen (1992).
4. Time taken to taken to mentally rotate will be in line in line with Shepard and Metzler (1971)

2 METHOD

2.1 Participants:

Twenty four participants were recruited using the UCL Psychology subject pool. Participant ages ranged between 20 and 57 years with a mean age of 34. Participants had no prior experience using the Radar and collision danger sector (CDS) displays.

2.2 Design:

A 5x5x5 within subjects design was utilised, the three factors were Radar, CDS and CDS-X display. The CDS displays communicate their information in the same way with the only difference being that in the CDS-X display the arrow crosses through the CDS as opposed to the other CDS display where the arrow is inside the CDS. Each factor was comprised of five levels, a hard left and hard right turn and a soft left, soft right turn and no turn speed up. The no turn speed up trials were used as filler and not included in the analysis as any response that participants made would not break the COLREGS. Order of display presented was counter balanced so as to negate any order effects that might arise.

2.3 Measures:

The main dependent variable being measured was error and particularly crossing errors. A crossing error is described as an error when participants break the rules of the game by crossing in front of the target vessel's path. It was of interest to investigate whether any of the three factors mentioned above would cause a significant difference in the number of crossing errors that participants were making. A Crossing error was recorded in the CDS-X factor when participants trying to decrease their delay chose the no turn speed up response as the display was communicating that this was the appropriate response. Crossing errors in the CDS factor were recorded when participants tried to decrease their delay by choosing a response that rotated the tip of own ships vector out of the CDS by crossing the target ships path. Along with crossing errors all other errors were also measured such as responding with a hard turn instead of a soft and vice versa as well as rotating in the incorrect direction. These errors will be used to investigate whether type of display has an effect on participants' performance in terms of the accuracy of their response.

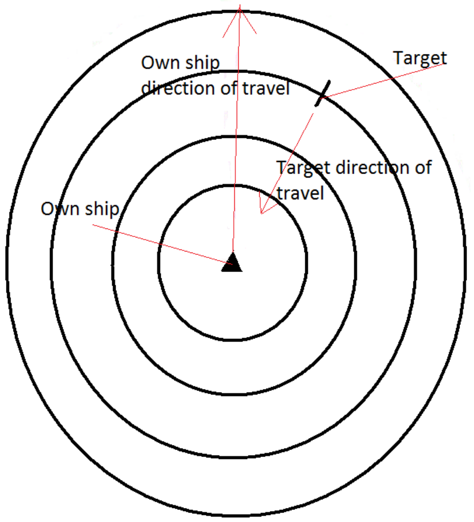
Other dependent measures taken were participants' reaction times (RT), task load for each factor which was measured using the NASA-TLX questionnaire and participants' subjective responses during a retrospective think-aloud.

2.4 Apparatus and Materials:

The display images for all trials were created in paint.NET. To accurately measure the angle at which the target vessel would be approaching, a 180 protractor was used. The angles at which target would approach, ranged from 10-80 degrees from the vertical, from both the left and the right hand side. The angle of approach range of 10-40 degrees is classed as the soft turn range while the 50-80 degree range is classed as the hard turn range. The experiment was run on a Dell desktop in the UCL psychology department and Power point was used to present the images to participants during the experiment. Participants' response times were measured using a stop watch. The images presented to participants in both the Radar and CDS displays are purposefully of own ship and target ship in close proximity to each other as it is at high proximity that the importance of rule following behaviour increases in its ability to decrease the likelihood of a collision.

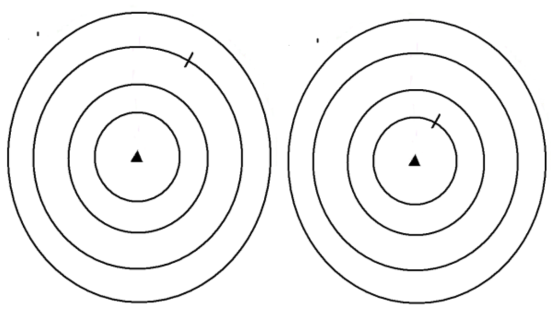
2.4.1 Radar trials:

The own ship icon was positioned in the centre of the screen. Using a 180 degrees protractor, the angles of approach for each trial ranging from 10-80 degrees on both the left and right hand side were marked.



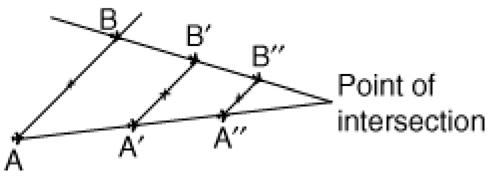
In order to simulate movement on the radar display, participants would be shown two consecutive images with target at differing distances both lying on the same approach angle from the own ship icon. The images were arranged so that when presented to participants sequentially they would simulate movement and thus participants would perceive that target is approaching on a constant bearing.

1 2



Participants would be first presented with the image on the left and then with the image on the right. This will simulate target moving towards the own ship icon on a constant bearing.

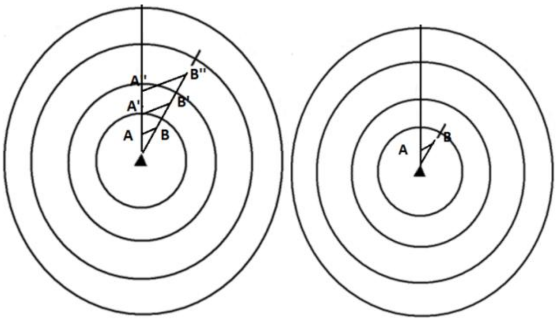
2.4.2 Constant Bearing



AB, A′B′, A″B″ = Lines of constant bearing

In navigation, a line from a fixed or a moving point to a moving object or a fixed point that retains a constant angular value with respect to a reference line is called a line of constant bearing

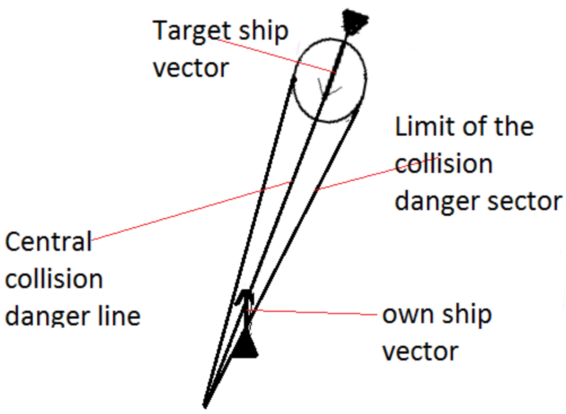
The lines of constant bearing have been super-imposed on the below images to help make things clearer.



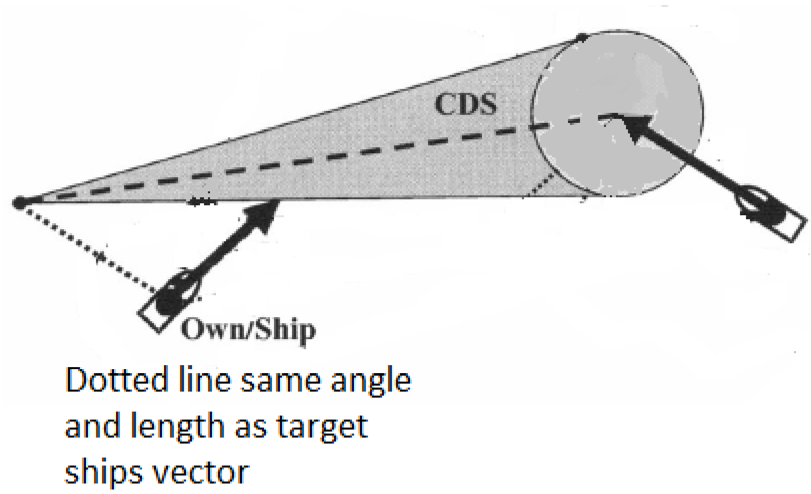
The example above shows target approaching at a constant bearing on a trajectory of 30 degrees from the vertical, therefore a soft turn response would be appropriate in order to avoid a collision situation.

2.4.3 CDS Trials:

CDS trials were based on the direct perception interface for ship-ship collision avoidance created by Liu and Pedersen (2006) whereby the user can avoid a collision situation by ensuring that the tip of their vessels vector remains outside the CDS of the target vessel. Below is an example of a trial that participants would be presented with.



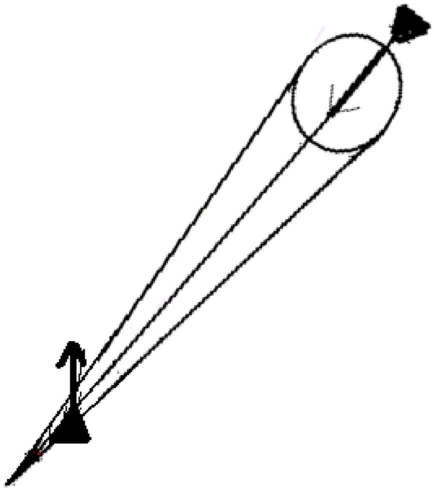
Creating an image of this display was possible by measuring the length and angle of approach of the target vessel’s vector. A line of same length and angle as the target vessel’s vector was then drawn coming out of the centre of the own ship. The line is graphically represented below in the form of the dotted line.



From the tip of this line other lines are drawn which end on each side of the safety envelope circle that surrounds the vector of the target vessel; this creates a conical shape. A final line is drawn linking the tip of the first line drawn with the tip of the target vessels vector. The conical shape that has been created represents the CDS which participants need to keep the tip of their own ship’s vector out of in order to pass a safe distance from the target vessel. If the tip of the own ship’s vector is lying on the central line then the two vessels are on a direct collision course.

The first CDS of the two CDS graphical representations shows target approaching at an angle of 30 degrees from the vertical. The tip of own ship vector is inside the CDS therefore a soft turn response from participants would be appropriate to move the tip of the own ship’s vector out of the CDS.

2.4.4 CDS-X Trials:



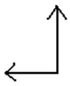
The CDS-X is the same display as the CDS but these images present participants with a scenario where the tip of the vector is protruding through the collision danger sector.

2.4.5 Delay System:

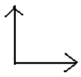
A soft turn left or right would incur a delay of one hour while a hard left or right turn would incur a delay of two hours. The no turn speed up option would decrease delay by 1 hour. Where participants erroneously under-rotated, resulting in crossing target's path or not turning hard enough to avoid a collision, an extra 4 hours of delay would be added to the existing delay that their option incurred. Therefore participants were encouraged to abide by the rules and prevent collisions. In order to prevent participants from being risk averse and always over-rotating, an extra penalty delay was introduced of an additional two hours of delay for every over-rotation a participant made. Participants were informed that the best score would be the one that incurred the lowest level of penalty delays and that there was a 10 pound prize for the best performance.

Participants were provided with a graphical representation of the angle of rotation that a soft turn response would rotate the own ship icon and also the angle of rotation that a hard turn response would rotate the own ship icon. The delay incurred by each response was also written alongside each representation, these images were placed next to the screen for quick reference if participants needed to do so.

Turn Left (45 degree turn from vertical) incurs a delay of 1 hour 

Turn hard left (90 degree turn from vertical) incurs a delay of 2 hours 

Turn Right (45 degree turn from vertical) incurs a delay of 1 hour 

Turn hard right (90 degree turn from vertical) incurs a delay of 2 hours 

Speed up (keep going straight) decreases delay by one hour 

2.4.6 Other apparatus:

In order to measure the task load of the two displays, NASA-TLX task load evaluation forms were provided to participants. Participants were provided with tally cards to assess the relative importance of the six factors in determining how much workload they experienced and list containing 6 workload factors and an unmarked 20-point, equal interval scale.

A second laptop to log participants' responses and RTs. The final piece of apparatus was a tape recorder in order to document participants' responses during a retrospective think-aloud.

2.5 Procedure:

2.5.1 Participants:

All participants were tested individually, seated at a desk in front of a computer monitor. Participants were informed that they would be presented with a series of images depicting possible collision scenarios and that these scenarios would be presented on two different types of display; the radar display and the CDS display.

Participants were briefed on the rules of the game which are as follows –

2.5.2 Rules common to Radar and CDS:

The common rules for the Radar and CDS displays stated that participants were not allowed to cross targets path. Participants were informed that their responses would be limited to five options. Their options would consist of a soft left or right turn which would rotate the own ship icon by 45 degrees and a hard left or right turn that would rotate their ship by 90 degrees. The final option is the straight up increase speed option. Participants were informed that the only point at which this option could be appropriate was if there was no possibility of crossing target's path.

The trajectories that target would be approaching would be 10-40 degrees from the vertical, where the soft turn response would be appropriate as this rotates the own ship icon by 45 degrees, which is enough to avoid a collision for all 10-40 degree approach trajectories. The hard turn 90 degree response was appropriate for the hard turn trajectories where target was approaching at 50-80 degrees from the vertical. In these cases the soft turn response of a 45 degree turn would never suffice to take the own ship out of a collision situation or a situation where own ship broke the rules by crossing target's path. Participants were also informed about the delay system, as mentioned above. The rules do not accurately reflect the actual COLREGS which state that the target vessel only has right of way if it approaching from the right hand side at a distance where the risk of a collision is imminent. In the current investigation the target and own vessel were positioned at a distance where a collision was imminent but instead of stipulating that only approaches from the right hand side required a avoidance response participants were instructed that approaches from the both left and right required a response. The reasons for this were firstly that it allowed an increase in the number of unique trials of interest used for data collection as responses from both the left and right side could be collected for analysis rather than just from the right. Secondly it was in the interest of the researcher to make the rules as simple as possible to minimize confusion for participants who had a small space of time to learn the appropriate rules as well as being familiarized with each display to a suitable level of proficiency.

2.5.3 Rules for Radar:

Target would be approaching own ship icon on either a constant bearing trajectory, which would result in a collision, or moving across the display on a path that posed no threat to the own ship. Participants were informed that as target was approaching on a constant bearing they had to make a response in order to avoid a collision situation.

2.5.4 Rules for CDS display:

If participants were presented with an image where the tip of their vector was inside the CDS they would be forced to choose one of the options to rotate the own ship icon so as to take the tip of the vector out of the CDS. In the case of the CDS-X trials the vector of the own ship icon would be crossing the CDS with the tip of the vector protruding out from the other side of the CDS. Although the tip of the vector in the CDS-X display is outside the CDS which means that own ship can continue on it path without danger of collision, remaining on the same path would result in crossing targets path which is against the rules.

2.5.5 Display familiarization Radar:

Training on the radar display consisted of images where all five options previously mentioned above would have to be utilised. Participants were therefore presented with images where a soft left and right turn, hard left and right turns, and the no turn speed up option were appropriate. Participants were shown one of each therefore training resulted in a total of five trials. The reasons as to why the particular responses were appropriate were communicated to participants and any queries were answered.

Participants were then presented with a further five training trials where they had to make a response unaided. After each response they were requested to give the reasons for their decision so as to ascertain whether they were using the rules of the game correctly to inform their decision making process. Any errors made were corrected and the reasons as to why their response was not appropriate were communicated. Out of the twenty four participants three were not able to reach proficiency by getting all the unaided training trials correct. In these cases the training session was run through a second time, with a repeat of the aided training session followed by the unaided training session. All participants after this point reached proficiency.

2.5.6 Display familiarization CDS:

Participants were presented with five training trials to cover all possible response options where the tip of the vector was inside the CDS and five trials when the tip of the vector was protruding out of the CDS, which are referred to as the CDS-X trials.

For CDS trials, participants were instructed that if the tip of the vector was inside the CDS they would have to choose one of the options that would rotate the tip of the vector out of the CDS and put the own ship icon on a path that would not cross targets trajectory.

In the case of the CDS-X trials participants were informed that although the tip of the vector was outside the display a response would still need to include one of the turning options, as the no turn speed up option would result in target crossing targets path which is against the rules.

The same process as the Radar display was conducted where five CDS and CDS-X trials were initially aided followed by five unaided CDS and CDS-X trials. Out of the twenty four participants seven were not able to reach proficiency at the end of both the CDS and CDS-X trials. In this case the aided and unaided training trials were repeated and all seven participants were able to reach proficiency by the end of the second round of training.

2.5.7 The experiment Radar:

Participants were presented with the two image sequences that made up each radar trial with target approaching at all eight angles ranging from 10-80 degrees from the vertical, from both the left and the right hand side. The total number of unique trials, where participants had to make a soft or hard turn response numbered 16 in total, and an additional four trials, where the no turn speed up response would be appropriate. The twenty unique trials were replicated to increase the statistical power of the experiment, resulting in the radar condition being comprised of forty trials in total and the order of presentation was randomised. Participants were informed that they had a time limit of four seconds to simulate the time pressure of resolving a collision situation. In these four seconds they would have to carry out a mental rotation of target and verbally give their response depending on whether they believed the degree of rotation fell within the soft turn category, hard turn category or no turn-speed up category.

2.5.8 The experiment CDS and CDS-X

Participants were presented with single images rather than a sequence of two. The number of unique CDS trials also numbered to twenty, as did the number of unique CDS-X trials. Both CDS and CDS-X were replicated to produce forty trials for each factor. The total number of trials in this stage of the investigation numbered at eighty and the order of presentation was again randomised. Participants were given the same instructions as above regarding mental rotation, response and time constraints.

2.5. 9 NASA-TLX

After completion of each factor participants were presented with the NASA-TLX question sheet. The NASA-TLX is comprised of six different items each with a rating scale developed for participants to rate their experiences during the task. Participants were required to place an X at any point on the 20 point interval of each scale for that particular item that best described their experience. Once completed participants were presented with a series of 15 cards containing pairs of the rating items and were asked to choose which of the items was most important to their experience of the workload in the tasks that they just performed. The number of times a particular item from a pair was selected was recorded and that tally was used to multiply its corresponding score on the scale. This process was carried out for all 6 items on the NASA-TLX and produced an adjusted personal task load score for each participant. The sum of the adjusted ratings was then divided by 15 which is the number of cards containing the pairs participants used to tally their weights this final weighted score was produced for each participant and were used as the dependent variable as deemed appropriate.

3 RESULTS

3.1 Crossing Error Scores:

The number of crossing errors made by a participant were subtracted from one. As each level contained 8 trials each trial was worth a value of 0.125. Therefore each error incurred a cost of 0.125.

3.1.1 Kolmogorov-Smirnov test for normality on data containing number of crossing errors made by participants.

The scores for both levels in all three conditions were significantly non-normal. Highest D(48) > .504, and lowest p<.000. All scores are highly positively skewed across the four levels of all three conditions demonstrating a ceiling effect of learning. Out of a possible score of 1, all scores < 0.98.

3.1.2 Friedmans ANOVA

Due to the positively skewed nature of the data it is appropriate to conduct a Friedmans ANOVA to compare the scores of the Radar display with the CDS display and CDS-X display. Friedmans ANOVA for crossing scores: D(2) F =2.8, $p>0.2$ – as the lowest p was still above the 0.2 mark it is possible to conclude that there was no effect of the three conditions on the number of crossing errors.

3.1.3 Radar – CDS and CDS-X Post Hoc Wilcoxon Signed Rank comparisons

The radar condition was utilised as the control condition to examine whether comparatively more turning errors were made in the advanced displays. Wilcoxon Signed Rank comparisons were carried out.
The comparisons comprised of the following.

Radar Normal turn vs CDS normal turn

Radar hard turn VS CDS hard turn

Radar Normal turn VS CDS-X normal turn

Radar hard turn VS CDS-X hard turn

No significant differences arose for the comparisons with a highest $Z > -1.155$, and lowest $p>.2$ as the result here demonstrates that the lowest p values was at the 0.2 mark.

3.1.3 Effect of Degree of turn on crossing errors Post Hoc Wilcoxon Signed Rank comparisons

It is also of interest to examine whether the degree of turn, be that a soft or a hard turn has any effect on the number of crossing errors made within each condition. To examine this, a Wilcoxon Signed –Rank pair test was utilised:

Comparisons comprised of:

Radar Normal Turn VS Radar hard turn

CDS normal turn VS CDS hard turn

CDS-X normal turn VS CDS-X hard turn

There was no effect of degree of turning on the number of crossing errors made

Highest $Z > -.707$ and lowest $p > 0.4$ within each condition.

3.1.4 Radar soft turn – CDS and CDS-X hard turns Post Hoc Wilcoxon Signed Rank comparisons

Lastly it was attempted to examine whether compared to the radar control condition the effect of display and turn could combine to increase the difference in number of turning errors made. In order to examine this possibility comparisons using the Wilcoxon signed rank pairs test would utilised. The radar control conditions normal turn scores were compared to both the advanced displays hard turn scores.

The comparisons comprised of:

Radar Normal turn vs Arrow in CDS hard turn

Radar Normal turn VS Arrow crossing CDS hard turn

No significant differences arose for the comparisons with highest $Z > -1.508$,and lowest $p > .132$. Although a slight decrease in p was found. A comparative decrease in scores between radar and the two advanced displays was not found. Inspection of the means further demonstrates this point.

Radar normal turn mean score: 0.98

Arrow in CDS hard turn mean score: 0.99

Arrow crossing CDS hard turn mean score : 0.98

3.2 Total Number of Errors:

Again the number of errors made by a participant was subtracted from one. As each level contained 8 trials each trial was worth a value of 0.125. Therefore each error incurred a cost of 0.125.

3.2.1 Kolmogorov-Smirnov test :

Kolmogorov-Smirnov test for normality on data containing number of total errors made by participants.

All scores are highly positively skewed across the four levels of all three conditions. Out of a possible score of 1, all scores < 0.72

As the data does not fulfil assumptions of normality non-parametric tests are appropriate. A Friedmans ANOVA was conducted to test for significant differences in total number of errors.

3.2.2 Friedmans ANOVA for total error scores:

D(2) F =5.06, $p>0.05$ – the results show that there is no significant effect of condition on the number of total errors although the p value lies at 0.079 thus lies close to significance.

On close inspection of the data containing a total of 2,880 trials, the total number of errors made was 443 of which 14 were crossing errors. This leaves the majority of errors made to be either over-rotating or under-rotating errors. In such cases participants chose an incorrect hard turn response as opposed to a correct soft turn response or an incorrect soft turn response as opposed to a correct hard turn response. It is thought that the placing of left and right labels on corresponding sides of the screen along with corresponding arrows can explain this high accuracy rate at least regarding participants rotating in the correct direction. For this reason the data will be discussed in terms of over-rotation and under-rotation.

3.2.3 Total Error Post Hoc Post Hoc Wilcoxon Signed Rank comparisons:

Post-hoc pair wise comparisons were conducted to investigate if a particular display was superior in its ability to decrease error. The first comparison Radar (Mean 0.77) and CDS (Mean 0.82) was significant with Z= -3.217, p<0.005. The higher the mean the fewer the number of incorrect trials made using each display therefore CDS (Mean 0.82) out-performed the Radar (Mean 0.77) display significantly.

The CDS-X factor (Mean 0.81) was not significantly superior in its ability to keep number of errors to a minimum compared to the Radar (Mean 0.77) display but the result was extremely close to significance Z= -1.954, p,= 0.051.

A final comparison between the two CDS displays CDS-X (Mean 0.81) and CDS (Mean 0.82) was not significant Z=-1.44, p>0.1.

The means indicate that both CDS and CDS-X allowed participants to be more accurate with their responses compared to Radar with participants performing best while using the CDS display.

3.2.5 Post Hoc Wilcoxon Signed Rank comparisons between the three factors for increase in proximity to the discrimination threshold

Firstly an explanation will be provided to explain what is meant by proximity to the discrimination threshold. A hard turn response meant that own vessel would be rotated by 90 degrees, the angle of approach trajectories where this is the appropriate response range from 50-80. The closest angle to the discrimination threshold of 45 degrees where the cut-off point between hard and soft turns was placed is when the target vessel is approaching at 50 degrees from the vertical. This proximity decreases for trials where the target ships approach trajectory increases to 60, 70 and finally 80 degrees from the vertical which is the lowest proximity angle. Similarly the highest proximity angle for soft turn response is when target approaches at 40 degrees from the vertical. Proximity decreases for 30, 20 and finally 10 degrees approach trajectory which has the lowest proximity. It is of interest to investigate the effects of the different displays as the angles approach the threshold of 45 degrees as it is closest to the discrimination threshold where the task of recognising a hard from a soft turn response becomes more challenging. Therefore the display that fares best closest to the discrimination threshold will demonstrate superior accuracy of response.

The same angle occurred twice in each of the four levels included in the analysis due to replication of trials. Further the two levels of soft left and soft right and hard left and hard right were merged so that the only categories remaining were either soft turns or hard turns.

Therefore every angle occurs four times. As every angle occurs four times it was possible to construct scores out of 1 giving each angle a value of 0.25 so that all correct angle responses would add up to 1. Participants erroneous responses within the soft turn range being 10-40 degrees will be discussed as over-rotations with the high numbers of over-rotations leading to a lower scores while erroneous responses within the hard turn range being 50-80 will be discussed as under-rotations ,the more under-rotations the lower the score.

Soft turn responses:

The scores for both levels in all three conditions were significantly non-normal. With the highest $D(24) > 0.2$, lowest $p<.001$. All scores are positively skewed across the four levels of all three conditions with the lowest mean 0.52.

Mean score for Radar at 10 degrees rotation was 0.93 and steadily decreased to a mean score of 0.63 by the 40 degrees rotation. Mean score for CDS began at 0.91 at 10 degree rotation and decreased to a mean score of 0.59 by 40 degree rotation while the mean score for CDS-X began at 0.91 at 10 degrees rotation and decreased to 0.5 by 40 degrees rotation. This decrease in mean scores was significant for all conditions lowest $Z=-3.479$, highest $p<0.001$. Inspection of the means also suggests that participants over-rotated more in the CDS and CDS-X conditions at 40 degrees than in the Radar condition. Further Inspection of the means indicates that he highest scoring display type at the highest proximity trajectory is the Radar display. As discriminating between the appropriateness of a hard or a soft turn response is most challenging at the highest proximity approach trajectory it could be suggested that the Radar display allows participants to be most accurate. No significant difference was found between the scores at this approach trajectory for Radar compared to CDS or CDS-X as will be demonstrated next when the three displays will be compared at all angles of approach.

Wilcoxon Signed Rank Pair test will be utilised to uncover any potential differences in the data. To give an example of the pairs, CDS-X 10 degrees was paired with CDS-10 degrees, CDS-X 10 degrees would also be paired with R-10 degrees and R-10 degrees would be paired up with CDS-10 degrees. This would occur for all 8 angles ranging from 10-80 degrees for all conditions.

Comparisons for angle range 10-40 degrees:
All except two pairs of the 10 to 40 degrees were not significant with the highest Z= -1.416 and the lowest p>0.1 of the non significant results. The two pairs where a significant difference in over-rotation were found was between CDS-X 20 (mean 0.75) and CDS-20 (mean 0.83) Z=-2.3, p< 0.05 and between CDS-40X (mean0.5) and CDS-40 (mean 0.59) Z= -2.714, p<0.01. The means indicate a higher degree of over-rotation in the case of the CDS-X compared to the CDS in both pairs. Otherwise there is not difference in over-rotation between any of the conditions in any of the other pairs.
This indicates that the radar was just as likely to over-rotate as the CDS displays in all angles of approach, even at the highest proximity to the threshold where accuracy was most tested.

Hard turn responses:
To investigate whether a significant difference was also found in the hard turn category as proximity to the discrimination threshold increased further comparisons were conducted. Significant differences were found for all displays, Radar (50 degree Mean 0.56) (80 degree Mean 0.93), CDS-X (50 degree mean 0.82) (80 degree mean 0.96) and CDS (50 degree mean

0.80) (80 degree mean 0.93) when comparing the difference between the mean score for each factor at 50 and at 80 degrees rotation Lowest Z=-2.07, highest $p<0.05$.

Increased proximity to the discrimination threshold again resulted in a decrease of mean scores across all three conditions with Radar having the largest change in mean score from mean 0.93 in the 80 degrees trials down to mean 0.56 in the 50 degrees trials. while these decreases are significant the mean Radar score (mean 0.56) at highest proximity to the discrimination threshold is lower than the CDS and CDS-X means at the same proximity of 50 degrees (mean0.8) and (mean 0.82) respectively . A dramatic change occurs where the Radar display which had the highest mean at the at the 40 degrees approach trajectory now at 50 degrees has the lowest.

Comparisons for angle range 50-80 degrees:

In the case of the 50-80 degrees pairs an increased number of significant differences were uncovered some of them highly significant demonstrating a significantly higher level of under rotation for the radar compared to both CDS and CDS-X displays at multiple angles of approach.

50 degrees Pairs:

In the CDS 50 (mean 0.8) - R-50 pair (mean 0.56) and the CDS-X50(mean 0.82) – R-50(mean 0.56) pair the highest Z=-3.617 and the lowest $p< 0.001$. In these pairs participants in the Radar condition significantly under-rotated compared to the CDS and CDX at the 50 degree angle. No significant difference was uncovered between the CDS-X50(0.82) and the CDS-50(0.8) Z=0.63, $p> 0.5$.

60 degree Pairs:

A significant difference was uncovered between the R-60 (mean 0.71) and CDS-60 (mean .86) Z= -2.627, p<0.01. As well as R-60 (mean 0.71) compared with CDS-X60 (mean 0.89) Z= 3.127, p<0.005. Scores for CDS 60(mean .86) and CDS-X 60(mean 0.89) were not significant with Z=-1.342, p>0.1.

70 Degree Pairs:

As the distance from the threshold increases under-rotations decrease and the differences between the three condition decrease in their significance and in some cases become highly non-significant. As is the case of R-70 (mean 0.89) and CDS-70 (mean 0.93) are now not significantly different in terms of under-rotations made Z=-0.905, p> 0.1. Due to CDS-X70 (mean 0.97) extremely high mean due to very few under-rotations the difference with R-70(0.89) is in this case significant Z=-2.126, p>0.05. There is no significant difference between CDS-70 (mean 0.92) and CDS-X70 (mean 0.97) Z= -1.667, p>0.96.

80 Degree Pairs:

There is no significant difference in between any of the factors R-80 (mean 0.92) and CDS-80 (mean 0.93), R-80(mean 0.92) and CDS-80X (mean 9.5), CDS-80(mean 0.93) and CDS-80X(mean 9.5) highest Z= -0.905 and lowest p>0.05.

A final comparison was conducted to investigate whether any of the displays performed significantly better at the highest proximity angles of 40 and 50 degrees approach where judging between a hard and a soft turn response would be most challegning. The means again indicate that participants when using the CDS scored highest in terms of accuracy (CDS Mean 0.69), (CDS-X Mean 0.66), (Radar Mean 0.59). The only significant difference found was between Radar and CDS, Z=-2.157, p<0.05

	SOFT TURN RESPONSE				
Radar 10	0.93	CDS 10	0.91	CDS-X 10	0.91
Radar 20	0.98	CDS 20	0.83	CDS-X 20	0.75
Radar 30	0.76	CDS 30	0.71	CDS-X 30	0.88
Radar 40	0.62	CDS 40	0.59	CDS-X 40	0.5
	HARD TURN RESPONSE				
Radar 50	0.56	CDS 50	0.8	CDS-X 50	0.82
Radar 60	0.71	CDS 60	0.86	CDS-X 60	0.89
Radar 70	0.89	CDS 70	0.92	CDS-X 70	0.97
Radar 80	0.92	CDS 80	0.92	CDS-X 80	0.95

Although the results of the comparisons did not reach significance the table above helps demonstrate the point that the Means for CDS and CDS-X compared to Radar indicate a higher degree of rotation in the majority of the trials.

Lower scores in the soft turn responses indicate that participants incorrectly chose the higher degree of rotation hard turn response. Higher scores in the hard turn response indicate that participants correctly chose the higher degree of rotation hard turn response.

The combined results from both the hard and soft turn responses indicate that Radar over-rotated when close to the proximity threshold for soft turn responses and under-rotated close to the proximity threshold in the hard turn responses. Both CDS and CDS-X displays under-rotated significantly less than radar near the hard turn proximity threshold indicating that while Radar, CDS and CDS-X are all equally likely to over-rotate CDS and CDS-X are not equally likely to under-rotate close the proximity threshold. In the current paradigm this demonstrates an increased accuracy for the CDS and the CDS-X displays over-all compared to Radar. Also in the case of CDS and CDS-X participants seemed to be more consistent in their direction of response, that being in the direction of harder turns. If the radar condition had caused similar behaviour then the highly significant differences in number of under-rotations between the Radar and CDS displays would not have existed. The final analysis demonstrates that the CDS is the most accurate out of the three displays which is in line with the initial comparison conducted immediately after the Friedmans ANOVA.

3.3 Reaction Time Results:

The mean reactions times for the angles within each condition were measured and compared. Only Reaction Times's for correct responses were utilised in the analysis.

Kolmogorov-Smirnov test for normality on data containing participant's reaction times. All conditions were not significantly different from being normally distributed lowest $D(24) = 0.113$, and highest $p > 0.1$

Means for each condition were

Radar – 1.9 s

CDS – 2.9 s

CDS-X – 3s

A three way ANOVA was conducted on the reaction times for each condition a significant effect of display on participants RT's was uncovered (D2) $F = 41.1$, $p < 0.001$

Post-Hoc comparisons were conducted demonstrating a significant difference between the mean RT's of CDS ($M = 2.9$, $SE = 0.11$)– Radar ($M = 1.9$, $SE = 0.7$, $t(23) = -8.66$, $p < 0.001$) CDS-X ($M = 3$, $SE = 0.9$) –Radar ($M = 1.9$, $SE = 0.7$, $t(23) = -10.37$, $p < 0.001$) whereby participants took significantly longer to complete the task using the CDS and CDS-X displays compared to the Radar display. No significant difference in the mean RT's of CDS($M = 2.9$, $SE = 0.11$) and CDS-X ($M = 3$, $SE = 0.9$ $t(23) = -0.916$, $p > 0.1$.

Post hoc comparisons were also conducted to ascertain whether there was a significant increase in RT between 10 degrees approach trajectories and 80 degrees approach trajectories. A significant result would provide evidence that participants were mentally rotating the own ship icon to gauge whether their response fell into the hard or soft turn response category.

Significant differences were found within all factors.

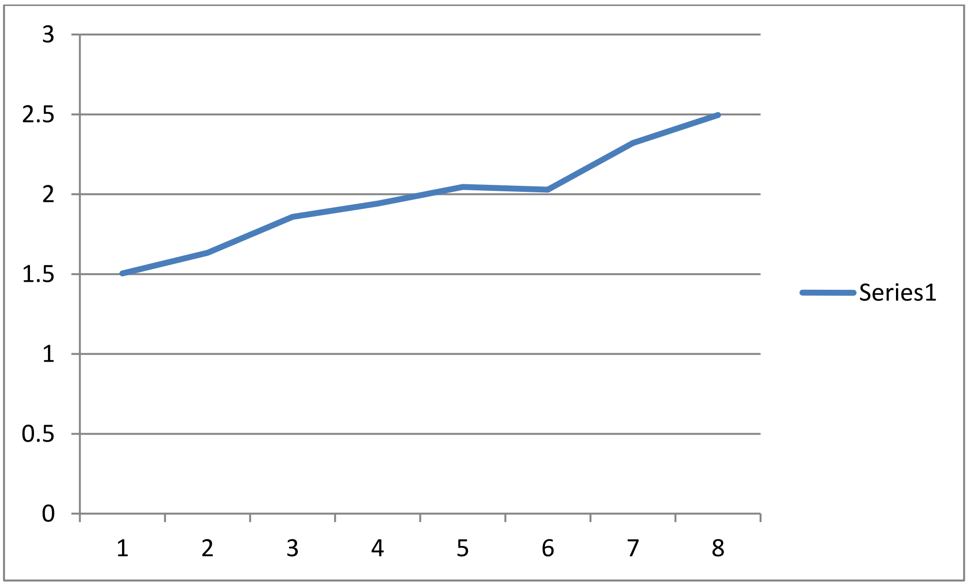
Radar 10 (M=1.5, SE=0.7) Radar 80 (M=2.5, SE=0.16, t(23) = - 5.471, p<0.001)

CDS 10 (M= 2.3, SE=0.18) – CDS 80 (M=3.4, SE=0.23, t(23) = -3.7, p<0.001)

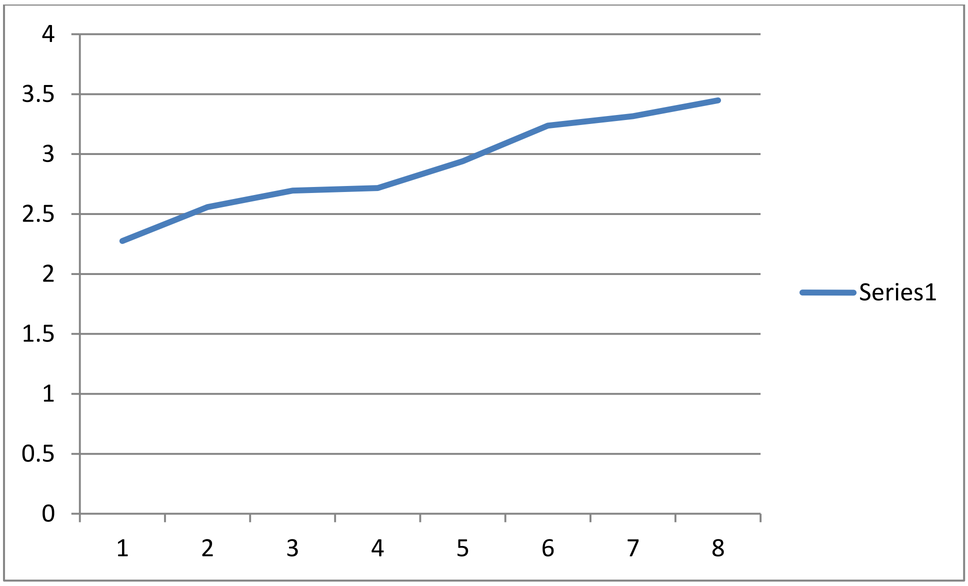
CDS-X 10 (M=2.4, SE=1.4) – CDS-X 80 (M3.7, SE=0.36, t(23) = -2.8, p<0.01)

The line graphs below all demonstrate the constant increase in reaction times as trajectory of targets angle of approach from the vertical increases.

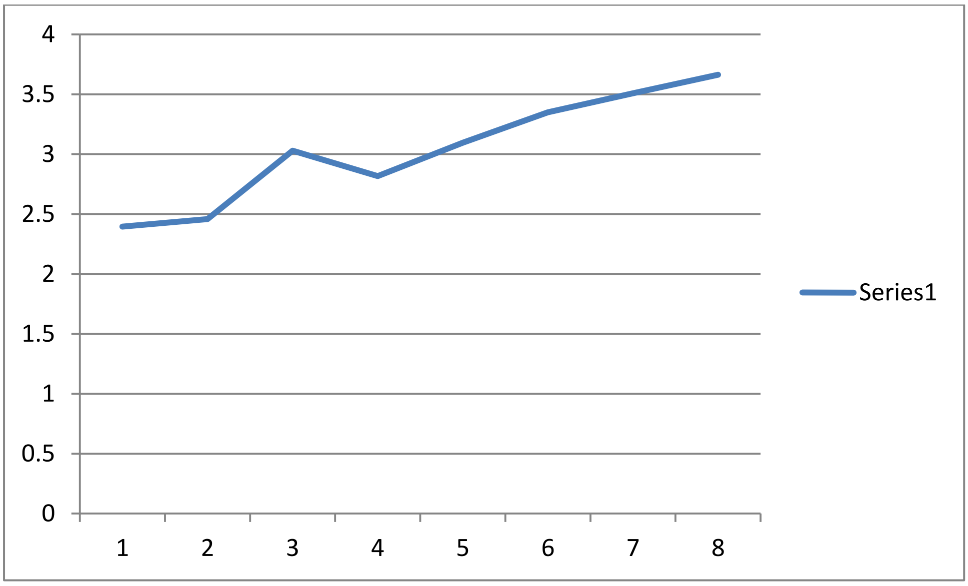
Reaction times for Radar over the 8 angles (10-80 degrees)



Reaction times for CDS over the 8 angles (10-80 degrees



Reaction times for CDS-X over the 8 angles (10-80 degrees)



3.4 NASA-TLX Results:

Kolmogorov-Sminov test for normality indicates a normally distributed data set $D(23)= 0.118$, $p>0.2$ allowing for the use of the parametric paired samples t-test.

The paired sample t-test compared the two sets of weighted scores for the Radar ($M=48.71$, $SE=2.55$) – CDS ($M=54.94$, $SE=2.87$, $t(23)= -1.99$, $p=0.59$) and a marginally non-significant result was indicated. Inspection of the means shows a slightly higher mean task load for the CDS compared to the Radar display.

3.5 Retrospective think-aloud:

Participants reported that the CDS-X display was harder to gauge the degree of appropriate turn as the arrow protruding out of the cone gave the impression that they should make a harder turn in order to ensure that the arrow was able to pass through and safely exit the other side of the CDS.

DISCUSSION

4.1 Crossing Errors

Trials in the CDS-X factor did not produce significantly different results from the CDS, or Radar factors in terms of number of crossing errors. The numbers of crossing errors made were extremely low across all three conditions indicating that participants had retained their ability to use their conceptual knowledge while interacting with all displays. This finding conflicts with the position taken by Lee and Sanquist (2000). Lee and Sanquist (2000) claimed that users were more likely to be drawn to the perceptual salience of an ARPA display and thus forgo all conceptual knowledge necessary to make the correct response and not break the COLREGS. The finding also conflict with the studies reported by Vincente and Rasmussen (1992) such as Hollagnel (1981), Smith (1989) and the Fischoff (1978) out of sight out of mind position. If the above authors' position was proven in this experiment the CDS and particularly the CDS-X trials would have produced a significantly higher number of crossing errors than the Radar condition. The reasons for this is that if the users were in fact drawn to the perceptual properties of the display and particularly the CDS-X display they would have chosen the no turn speed up option. This is because the display was presenting this option as one that was acceptable while ignoring conceptual information not presented in the form of the COLREGS. The CDS display was expected to also demonstrate an increase in participants crossing errors as they would attend to the perceptual salience of the display and only use simple heuristic of keep the arrow outside the danger area and so not incorporate conceptual knowledge such as COLREGS.

Possible explanations do exist for the current investigations findings. Vincente and Rasmussen (1992) discuss that other factors that have an effect on the level of cognitive control are experience and skill of the operator. The more experienced the operator the greater the likelihood that the operator will rely on the lower levels of cognitive control such as SBB and RBB. Though it could be argued that this is in fact a fairly questionable suggestion as expertise should surely involve learning appropriate level schemas in which the high level control is also embedded. Nonetheless it is therefore possible based on the evidence produced by the investigations referenced by Vincente and Rasmussen (1992) that the participants lack of skill and experience affected the level of their cognitive control. Specifically making them less likely to rely on the lower levels of cognitive control where the perceptual salience of the display may have had an increased effect. Participants therefore may have had to draw on their analytical abilities incorporating the rules of the game to make the correct response and so exhibited Knowledge based behaviour (KBB). It follows then that a significant limitation of the current experiment was lack of resources in the form of being able to use experienced navigators for the task that were familiar with using the CDS display.

Lee and Sanquist (2000) suggested that inexperienced mariners can be drawn to the salience and seemingly reliable information provided by a display due to the component input select which is based on selective attention and can be affected by the salience of the information source. This explanation could be another interesting avenue to explore whereby the question would be framed in terms of novelty. Participants would be presented with fewer CDS-X trials in order to increase their novelty. It is possible based on the suggestion made by Lee and Sanquist (2000) that by increasing the novelty of the CDS-X trials where the display communicates that no response is appropriate could further test participants' ability to use their conceptual knowledge in order to make the correct response. Participants may be drawn to the perceptual salience of the display communicating that a response that breaks the COLREGS is in fact appropriate. The current paradigm would be utilised, only the training stage of the experiment would be changed by simply taking out the CDS-X training trials. The information given to participants regarding the COLREGS using only the CDS trials would suffice to build a conceptual understanding of the rules that they are expected to adhere to. Further factors could also be investigated that may increase the likelihood of participants treating the information displayed as absolute. One such factor may be an increase in task difficulty by introducing a secondary loading task such as reciting a text or a list of numbers. This could possibly affect participants' ability to use their conceptual knowledge and thus make them more likely to make their decision purely based on the display.

4.2 Total errors

The accuracy of participants total responses including crossing errors were measured and compared between all the conditions. Again it was expected that both CDS and CDS-X displays would be inferior to the radar display due to the prediction that they would cause significantly higher crossing errors. As discussed above there was not difference in the number of crossing errors between any of the displays. Further the high means demonstrate that not only was there no significant difference but there was a strong ceiling effect of learning for all conditions. These differences end when looking for more than just crossing errors. The majority of errors were where participants rotated in the correct direction but either chose too hard a turn or too soft a turn. This error section was framed as errors described as under-rotations or over-rotations, meaning that participants selected inappropriate soft or inappropriate hard turn respectively.

The hypothesis that the higher the proximity to the cut-off the harder it would become to recognise which category the suitable response would fall into was supported. A significant decrease for all displays in the mean score for both the soft and hard response category was demonstrated as the proximity from the threshold increased.

Comparisons made at each angle of approach trajectory indicated that compared to Radar CDS and CDS-X did not rotate significantly more at the high proximity angles to the cut-off point for the soft turn response category though inspection of the means does indicate that participants were more likely to select a hard turn response when using the CDS and CDS-X displays than the Radar display.

Comparisons made at individual angles between the three factors showed no significant difference between the scores of the Radar compared to both the CDS displays at any of the 10-40 degree angles of approach. The only two significant differences demonstrated were between the two CDS displays at the 20 and 40 degree rotations where CDS-X scores were lower due to higher levels of over-rotation. This finding is in line with reports given by participants during the retrospective think-aloud. During the retrospective think-aloud participants reported that if the arrow is protruding out of the other side of the CDS then it was assumed that a hard turn was more appropriate compared to a soft turn to bring the tip of the vector back through the CDS and out of the other side in order to avoid crossing targets path. Participants at the hard turn 50-80 range under-rotated significantly more in the angles closer to the discrimination threshold in the radar condition as opposed to the CDS displays. This supports the idea that radar is equally likely to over-rotate when closer to the discrimination threshold on the soft turn side and under-rotate when near the discrimination threshold on the hard turn side. The table of means supports the suggestion that participants were more likely to over-rotate in the majority of angles of approach.

The current paradigm demonstrates an increased level of accuracy for the CDS and CDS-X displays though this could be misleading. Reasons for this are that although in the hard turn response category participants made less under-rotations using the CDS and CDS-X displays this does not necessarily mean that they were accurately able to gauge the appropriate angle of rotation but rather always aired on the side of a harder rotation which happened to fall in the appropriate response hard turn category. Therefore the current paradigm is relatively blunt in its ability to glean accurate information as to how accurate participants' responses for the CDS and CDS-X displays really were. This in itself is a limitation of this paradigm and improvements could be made. Improvements could include rather than asking participants to place their response in one of two categories to rather give a numerical value corresponding to the angel that they deem appropriate to rotate the tip of the vector outside the collision danger sector. It should be remembered that for the aim of the current investigation the paradigm would have been able to identify responses that would be classified as crossing errors, it is just from a usability point of view that this change to the paradigm would be beneficial in terms of increasing accuracy.

As the data for the CDS displays suggests that closer to the discrimination threshold in for both hard and soft turn responses, participants were more likely to consistently choose a higher angle of rotation than in the radar condition with this effect being increased for the CDS-X display. This result would have been further supported if Radar was found to over-rotate significantly less in the soft turn response condition than the CDS and CDS-X. Once again inspection of the means demonstrates an effect in this direction at the highest proximity angle of 40 degrees as both the means for the CDS and CDS-X are lower than the Radar display although the difference does not reach significance.

Again a change in the paradigm whereby a numerical response as opposed to a category response were given would be able to help verify this claim. If the claim was verified it would support the suggestion that the users are powerless to the perceptual salience of the display as the CDS display is undermining their judgement causing them to select a higher degree of rotation than they otherwise would. If the CDS and CDS-X displays do in fact increase the likelihood that participants choose a harder response as opposed to a soft one then this has implications for the use of the CDS displays in a maritime environment. In the context of a marine environment choosing a higher as opposed to a lower angle for the CDS can be seen as a plus as the navigator will choose the more risk averse option of the two when deciding on the angle by which he may deem appropriate to turn his vessel in order to avoid a collision with another vessel. In highly crowded waters this may become an issue as the navigator will want to be as accurate as possible. In the commercial context this risk averse behaviour may not be as favourable as a consistently higher turning rate to avoid collisions will mean that the vessel will have to travel an increased distance to its destination increasing the delay to destination and fuel consumption. The more optimal type of display could make use of an increased level of automation whereby a selection of optimal paths which take into consideration the COLREGS would be presented to the navigator allowing the navigator to select the most appropriate route. This route could then be communicated to other vessels and superimposed on their displays in order for all vessels within that area to have a complete understanding of other target vessels intentions.

4.3 Task Load: Reaction Time, NASA-TLX:

Liu and Pedersten (2004) claim that their perceptual display increases transparency of the task of collision avoidance by making the CDS visible and thus decreases the level of workload. Therefore it was expected that participants' RT would be lower in both the CDS and CDS-X factors as the CDS plays on the strength of rapid perceptual processing (Vincente and Rasmussen 1992). Radar has no representation aid to augment the participants' perceptual processing abilities. The results do not support this hypothesis as the difference between the results was in the opposite direction and significantly so. Reasons for the increase in time to task completion experienced by participants' when using the CDS display could lie in the increased detail of the CDS displays which contained own ships vector, the target ships vector and CDS. The radar was comparatively less rich with just a line representing target and a triangle shaped blip at the centre of the display representing own ship. On the other hand the simplicity of the display also could contribute to why the total error score for the Radar display was higher as participants were able to use the increased detail of the perceptual display in the CDS and CDS-X factors in order to increase the accuracy of their responses. Data from the NASA task load index did not show any significant difference in the level of task load that participants experienced when carrying out the task using the two different types of displays. Though the trend did indicate that task load was higher for the CDS displays than the radar display with a marginally non-significant result. Liu and Pedersen (2006) using the NASA-TLX measure on two experienced navigators found that one of the navigators struggled to use the CDS display due to lack of familiarisation. The navigator consequently rated the CDS as harder to use than the standard ARPA that he was familiar with. Participants' lack of extensive experience in the current studies increases the possibility that they found the task more challenging due to lack of familiarization as well. One issue with this claim is that participants were familiarized to the same degree for both Radar and CDS. Therefore an explanation based purely on familiarization does not fully explain this significant increase in reaction times and the trend where CDS loaded more than Radar. It is therefore possible that the combination of factors such as the increase in complexity of the CDS display along with lack of familiarization caused the longer reaction times and the increased difficulty as recorded by the NASA-TLX. Implications for these results in terms of design recommendations could include keeping a display as simple as possible so that participants can quickly become familiarized with its functions. An added bonus of keeping the display as simple as possible is that it reduces the level of clutter, in a high traffic condition where these displays are most needed to optimize the navigators' performance; clutter can severely undermine the navigators' ability to carry out the task.

4.4 Mental Rotation

The final hypothesis presented regarding the replication of the increase in reaction time as a linear function of the degree of mental rotation was supported. Participants demonstrated a steady increase in reaction times as the degree by which they had to rotate the own ship icon increased. Participants did not use recognition to guide their responses of either a hard or soft turn. It seems that the effect of mental rotation overcame recognition and participants were mentally rotating the own ship icon even at the most extreme angles of rotation such as 80 degree where it was clear that the particular collision example fell into the hard turn category. The effect of increased reaction times as a function of mental rotation could disadvantage the navigator when faced with a collision situation. Rather than allowing response time to be affected by mental rotation it would be more appropriate to introduce a function that would present the navigator with the required degree of rotation when they select a specific target vessel. This information could then be rapidly utilised to prevent a possible collision situation.

6 CONCLUSION

In conclusion while the effect of perceptual salience on participant's ability to abide by the COLREGS was not supported, the fact that participants consistently selected hard turn responses more in the CDS displays than the Radar suggest an effect of perceptual salience on participants. This effect of salience is not on the judgement of course chosen with respect to the COLREGS but rather with regard to participant's ability to judge the appropriate rotation that their vessel should complete. An interesting trade-off between display complexity and performance was also indicated as participants seemed to take longer to complete the task using the CDS displays but were more accurate in their responses compared to the Radar display. The findings reported above have produced further avenues for investigation into the critical issue regarding the negative effects that could be produced by the perceptual salience of a display. While the commercial fleet is using technology as a reason to increase the traffic at sea it is of the utmost importance that controlled studies such as this one be conducted in order that any failings in these systems be recognised and used to inform design, training and implementation.

7 REFERENCES

Cockroft, A. N. ‘Collision at sea”, *Safety at sea,* pp17-19, June, 1984

Fischhoff, B., Slovic, P., & Lichtenstein, S. (1978). Fault trees: Sensitivity of estimated failure probabilities to Problem Representation, *Exp. Psych.: Human Perception and Performance 4*(2), 330–344.

Hollnagel, E. “Report from the third NKAiKRU-experiment: The performance of control engineers in the surveillance of a complex process,” NKA/KRU-P2(81)36, Electron. Dept., Risa Nat. Lab., Roskilde, Denmark, 1981.

Kirlik, A.C. “The organization of perception and action in complex control skills,” unpublished dissertation, The Ohio State University, Dept. Ind. Syst. Eng., Columbus, OH, 1989.

Klein, G.A. “Recognition-primed decisions,” in W. B. Rouse, Ed. *Advances in Man-Machine Systems Research.* vol. *5,* Greenwich, CT: JAI Press, 1989, pp. 47-92.

Liu, Q., & Pedersen, E. (2004). Direct perception of collision danger information for safe marine navigation. *Systems, Man and Cybernetics, 2004 IEEE …*, 3898–3902.

Liu, Q., Pedersen, E., Okazaki, T., Fukuto, J., & Tanaka, K. (2006). Direct Perception Interface for Ship-Ship Collision Avoidance.

Nielsen, M., & Petersen, J. (n.d.). Collision avoidance at sea – practice and problems, 81–90.

Pedersen, E., Inoue, K., & Tsugane, M. "Evaluation of a radar plot and display technique for anti-collision assessment of multiple targets in true vector presentation by application of a environmental stress model," *Journal of Japan Institute of Navigation,* Vol.106 pp. 1-12, 2002a.

Pedersen, E., Inoue, K., & Tsugane, M. "On decision making in connection with optimal anti-collision manoeuvres by high speed craft in waters congested by conventional speed vessels," *Journal of Japan Institute of Navigation,* Vol. 107, pp. 177-183, 2002b.

Rasmussen, J. *Information processing and human machine interaction: an approach to cognitive engineering*, North-holland, New York, 1986.

Shepard, N., & Metzler, J. (1971). Mental rotation of Three-Dimensional Objects. *Science,* 171, 701-703.

Smith, G. F. (1989). Representational effects on the solving of an unstructured decision problem. *IEEE Transactions on Systems, Man, and Cybernetics, 19*(5), 1083–1090. doi:10.1109/21.44024

Tam, C. K., Bucknell, R., and Greig, A. (2009). Review of collision avoidance and path planning methods for ships in close range encounters. J. *Navigation*, 62, 455–476.

Vicente, K., & Rasmussen, J. (1992). Ecological interface design: Theoretical foundations. *Systems, Man and Cybernetics, …*, *22*(4).